Dalia GSR Week 1 June 26, 2024

**No table of contents entries found.**

## Overview

Understanding the variable construction in the Cengiz et al 2019 Paper

Specifically how the outcome variable (employment share) in the equation (1) is created

## Issues

* ~~I cant find the totransfer.do or totransfer.dta . These seem to be one the first codes that you create the data with.~~ Found the totransfer.dta in data files.

## Emails

### My Email on June 24, 2024

*Here I am thinking loudly, if you liked we can talk about these on Friday or any time that suits you.*

*1. I looked at the do-file that Alvaro mentioned (clean\_admin\_cps\_workingdata\_allind\_5cents\_nominal.do) and I am still not fully clear on how they create the dependent variable in regression 1 E\_{sjt}/N\_{st} . They are not doing the regressions in that do-file. I looked at the code that creates the Table 1 where they do the regressions (Table1\_for\_QJE.do). I believe they created the dependent variable there but that file was also hard to read and I still can't understand exactly the denominator N\_{st}. ... let me see what I can do.*

*2. We still have a mass at 0 in figure 2 while we are omitting all the observations with non-positive wage. This is weird. Do you know which of the cleaned data files we use to create Figure 2? Is it important to think about where the 0 wages come from?*

### My Email on June 26, 2024

Hi Dalia,

This is my latest understanding of how they create wage bin counts in each state-quarter EjstEjst​ and the total population in each state-quarter NstNst​, which are used to create the outcome variable in Equation 1 in the Cengiz et al. paper. Everything is weighted by the variable earnwt in the CPS data. This is the description in the “CPS Labor Extracts, 1979-2006”:

"Earnings weight for all races. Used for tabulating earnings-related items. Since the CD-ROM includes all persons asked earning questions, this sums to the total population each month and 12 times the population for each MORG file."

Based on their dofile “state\_panels\_cents\_new\_QJE.do”, where they create the main employment counts data, I replicated the totalpopulation variable. It is almost equivalent to the non-institutional population for each state (i.e., population aged 16 years and above). For example, this number for California (statenim == 6) is around 31 million, which is consistent with what I found on the internet.

Dalia Email

My Email