Lecture Module 11 - Economics 113

- The Big Picture
- Labor Markets
  - Labor Force Participation
  - The "Wage Gap"

The Big Picture

- What have we been studying, and why do we care?
  - The world is comprised of things we see, and things we do not
  - Econometrics is only meaningful when we have some sense of the things we do not see.
  - Put differently, if we have a reasonably understanding of what we are controlling for, we can make sensible recommendations based on the data.
- In recent lectures, "Fixed Effects" help carve down the world into something manageable
  - We've mostly thought of fixed effects as "individual effects"
  - But, they can technically represent any group of dummy variables for which we'd like to de-mean the data.
  - We will do so today using a common cross-sectional dataset from the "Current Population Survey"

Current Population Survey

- The CPS is a widely used and influential dataset from the US Bureau of Labor Statistics
- The Center for Economics and Policy Research (CEPR) has cleaned this data for use in a way which is consistent over time
  - http://ceprdata.org/cps-uniform-data-extracts/
  - Outgoing Rotational Group: 1979-2013.
  - Data on Wages, Demographics, locations, Industries, Occupations, Labor Force Participation, Education, etc...
  - We'll combine surveys from 1983, 1988,...2008, 2013 to a "pooled cross-section".
- We will focus on two outcomes, and the use of fixed effects in establishing "facts" for these outcomes
  - Labor force participation rates
  - The "wage gap" across groups.

One more Stata "trick"

- Suppose we estimate the following:

$$\log(wage_i) = \beta_0 + \beta_1 educ_i + u_i$$

- But, we'd also like to control for the fixed effects of individuals being in some "group".
- Using a trick with xtreg, one can estimate a (potentially large) group of fixed effects.
- For example, to estimate 50 state fixed effects, write:
  - xtreg lwage educ, fe i(state)
- To look within state-year pairs, write:
  - egen state_year = group(state year)
  - xtreg lwage educ, fe i(state_year)
- This type of regression is very important - controls for state specific shocks, such as recessions, tech booms, natural disasters, etc...

Labor force participation

- Labor Force Participation is an important part of the calculation of unemployment.
  - The denominator of the unemployment rate is the "labor force", which is the population that participates in employment or job search.
  - The participation rate is a measure of "discouraged workers"
- We will use the CPS ORG pooled cross sections to study the basic characteristics of labor force participation:

$$\begin{aligned} NLF_i &= \beta_0 + \beta_1 educ_i + \beta_2 age_i + \beta_3 age_i^2 \\ &\quad \beta_4 female_i + \beta_5 black_i + \beta_6 hispanic_i + \beta_7 other_i + u_i \end{aligned}$$

- Run with year and with state-year fixed effects
  - Within Years (that is, the year individuals were surveyed)
  - Within State-Year groups

Labor force participation

- First, run the within year model
  - xtreg nlf educ age age2 female black hispanic other, fe i(year)
  - Where might not controling for states produce an omitted variable problem for these estimates?
- Next, run the within state-year model
  - egen state_year = group(state year)
  - xtreg nlf educ age age2 female black hispanic other, fe i(state_year)
- Next, run on the unemployment indicator, but for only those searching for work:
  - xtreg unem educ age age2 female black hispanic other if nlf==0, fe i(year)
  - xtreg unem educ age age2 female black hispanic other if nlf==0, fe i(state_year)

The Wage Gap

- One of the most hotly debated issues of policy
  - Most estimates put the wage gap at around 80 cents on the dollar for women vs. men.
  - Even the Obama administration, who actively champion for its elimination, have wage gap.
  - Are wages different across gender and other groups for otherwise similar workers that do the same task?
- The key issues are pretty obvious:
  - What defines "otherwise similar workers"?
  - What defines "same task"?
  - How does one assess productivity within the workplace?
- This question speaks to the core of econometrics - how to we eliminate all relevant unobserved factors that may complicate the assessment of the wage gap.

The Wage Gap

- Data rarely allows us to disentangle these things, but we'll try with the CPS ORG dataset.

- Estimate the following basic equation:

$$\begin{aligned} \ln(rw_i) &= \beta_0 + \beta_1 educ_i + \beta_2 age_i + \beta_3 age_i^2 \\ &\quad \beta_4 female_i + \beta_5 black_i + \beta_6 hispanic_i + \beta_7 other_i + u_i \end{aligned}$$

- $\ln(rw_i)$ is the natural log of the real wage.
- Within Years
  - xtreg ln_rw educ age age2 female black hispanic other, fe i(year)
- Within State-Years
  - xtreg ln_rw educ age age2 female black hispanic other, fe i(state_year)

The Wage Gap

- Evaluate how the gender gap has changed over time
- It's helpful to dummy all this out explicitly

$$\begin{aligned}\ln(rw_i) = \; & \beta_0 + \beta_1 educ_i + \beta_2 age_i + \beta_3 age_i^2 + \beta_4^{83} d_i^{83} female_i \\ & + \beta_4^{88} d_i^{88} female_i + \beta_4^{93} d_i^{93} female_i + \beta_4^{98} d_i^{98} female_i \\ & + \beta_4^{03} d_i^{03} female_i + \beta_4^{08} d_i^{08} female_i + \beta_4^{13} d_i^{13} female_i \\ & + \beta_5 black_i + \beta_6 hispanic_i + \beta_7 other_i + \cdots + u_i\end{aligned}$$

- Eg. d_i^{88}: respondent was surveyed in 1988.
- Since $d_i^{83} + d_i^{88} + d_i^{93} + d_i^{98} + d_i^{03} + d_i^{08} + d_i^{13} = 1$, do not need to include $female_i$ on its own.
- Interpretation: β_4^t measures the wage gap in year t
- See detailed notes for the code required to do this.

The Wage Gap

- Let's try to get at the "same task" assumption by adding industries and occupations to our "within" analysis.
- Industry and occupation questions were only asked for 2003, 2008, and 2013 for our sample.
- Within Industry-State-Years
  - egen ind_state_year = group(ind_2d state year)
  - xtreg ln_rw educ age age2 female black hispanic other, fe i(ind_state_year)
- Within Occupation-Industry-State-Years
  - egen occ_ind_state_year = group(docc03 ind_2d state year)
  - xtreg ln_rw educ age age2 female black hispanic other, fe i(ind_state_year)
- Can do this using even more detailed industry classifications and occupation classifications. And we probably should.