Pooling Cross Sections across Time

Lecture 2 EE426 - 2/2013

Chayanee Chawanote

Pooling Cross Sections Across Time

2 kinds of data sets:

- Independently pooled cross section: sampling randomly from a larger population at different points in time
  - Observations are not identically distributed.
  - To increase the sample size, more precise estimators and test statistics with more power
- Panel data or longitudinal data: collecting the <u>same</u> individuals, families, firms, cities, states, etc, across time
  - Can't assume that observations are independently distributed across time: there are time-constant, unobserved attributes of the units being studied.

Year dummy variables

- The population may have different distributions in different time period. We need the intercept to differ across periods >> including year dummy variables
  - The coefficients on year dummy variables represent changes in the dependent variable for reasons that are not captured in the explanatory variables.
  - Compared to a based year
- We can interact a year dummy variable with key explanatory variables to see if the effect of that variable has changed over a certain time period.
  - Without interactions, the effect of each explanatory variables has remained constant
  - What happens if we interact all independent variables with a year dummy variable?

Year dummy variables

- Interactions:

$$\log(wage) = \beta_0 + \delta_0 y85 + \beta_1 educ + \delta_1 y85 * educ + \beta_2 exper + \beta_3 exper^2 + \beta_4 union + \beta_5 female + \delta_5 y85 * female + u$$

  - Intercept for 1985 =
  - Return to education in 1978, and in 1985 =
  - The gender gap in log(wage) in 1985 =
  - Test that nothing has happened to the gender differential over this 7-year period?

- Cautions:
  – Monetary term needs to adjust to real value since we compare across years
  – But if using logarithmic form, real or nominal wage only affects the coefficient on the year dummy, y85

Chow test for structural change across time

- Chow test: whether a multiple regression differs across 2 groups

$$F = \frac{SSR_p - (SSR_1 + SSR_2)}{SSR_1 + SSR_2} \cdot \frac{n - 2(k+1)}{k+1}$$

- Alternative Chow test for 2 periods: interacting each variable with a year dummy for one of the two years and testing for joint significance of the year dummy and all of the interaction terms
- If more than 2 periods?
  1. Estimate the restricted model by pooled regression allowing for different time intercepts >> give SSR_r
  2. Run a regression for each of T time periods >> get SSR_{ur}

     $SSR_{ur} = SSR_1 + SSR_2 + \ldots + SSR_T$
  3. If there are k explanatory variables (not including the intercept or time dummies), we are testing (T-1)k restrictions, and there are T + Tk parameters in the unrestricted model

$$F = \frac{SSR_r - SSR_{ur}}{SSR_{ur}} \cdot \frac{n - T - Tk}{(T-1)k}$$

Policy analysis with pooled cross sections

- Evaluating the impact of a certain event or policy with the data collected before and after the occurrence of an event
- A natural experiment occurs when some exogenous event (a change in government policy) changes the environment in which individuals, families, firms operate.
  - control group (C): not affected by the policy change
  - treatment group (T): affected by the policy change
  - These two groups are not randomly assigned. To control for systematic differences between the two groups, we need 2 years of data: before and after the change
- Let dT = 1 if in the treatment group, 0 otherwise

  Let d2 = 1 if post-policy change time period, 0 otherwise

$$y = \beta_0 + \delta_0 d2 + \beta_1 dT + \delta_1 d2 \cdot dT + other\ factors$$

Policy analysis with pooled cross sections

- Without other factors in the regression, $\hat{\delta}_1$ will be the difference-in-differences estimator:

$$\hat{\delta}_1 = (\bar{y}_{2,T} - \bar{y}_{2,C}) - (\bar{y}_{1,T} - \bar{y}_{1,C})$$

- Sometimes, $\hat{\delta}_1$ is called the "average treatment effect", measuring the effect of the treatment or policy on the average outcome of y

	Before	After	After - Before
Control			
Treatment			
Treatment - Control			