ANALYSIS OF PANEL DATA

Fixed-Effect and Random-Effect Models

datascience@berkeley

Fixed-Effect Model

An Introduction to Fixed-Effect Models

A More General Form

β_0

Fixed-effect, <u>potentially correlated with explanatory variables</u>

$$y_{it} = \beta_1 x_{it1} + \cdots + \beta_k x_{itk} + a_i + u_{it},\; i = 1, \ldots, N, t = 1, \ldots, T$$

$$\bar{y}_i = \beta_1 \bar{x}_{i1} + \cdots + \beta_k \bar{x}_{ik} + \bar{a}_i + \bar{u}_i$$

Form time averages for each individual.

$$\Rightarrow \boxed{[y_{it} - \bar{y}_i]} = \beta_1 \boxed{[x_{it1} - \bar{x}_{i1}]} + \cdots + \beta_k \boxed{[x_{itk} - \bar{x}_{ik}]} + \boxed{[u_{it} - \bar{u}_i]}$$

Because $a_i - \bar{a}_i = 0$ (the fixed effect is removed)

- **Estimate time-demeaned equation by OLS.**
  - Uses time variation within cross-sectional units (= within-estimator).

Within Estimation

- The *fixed effect transformation* is also called *within transformation*, in which *within* can be read as within each of the subjects in the dataset: each of the cross-sectional subjects' data is demeaned leveraging on time variation in both y and x and more importantly, the *unobserved individual heterogeneity* is eliminated within each of the individuals.

- The *fixed effect estimator* is also called *within estimator*.

- Because the transformation relies on time variation within each of the cross-sectional subjects, those variables without much variation to begin with will become (almost) a constant after the transformation, resulting in imprecise estimates.

Between Estimation

- **Between Estimators:** The OLS estimators on the cross-sectional average equation introduced above, from which we substract the cross-sectional equation, is called the $.

- We will not discuss the $, as it is biased when the observed explanatory variables, $x's$, are correlated with the unobserved fixed effect, a_i . Also, it does not utilize the panel data efficiently: it does not leverage on the time varying information.

- If we think that the observed explanatory variables and the unobserved fixed effect are not correlated, then we should us **random effect model**, which we will discuss later in this lecture.

Berkeley SCHOOL OF INFORMATION