Empirical Panel Data: Lecture 3

INSTRUCTOR: CHAOYI CHEN
NJE & MNB

March 21, 2023

Topic 2: Linear unobserved effects panel data models

- A standard setup for a linear unobserved (individual) effects panel data model:

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it},$$

where

- α_i is a scalar,
- $\beta = (\beta_1, \beta_2, \ldots, \beta_k)^T$ is a $k \times 1$ vector of parameter,
- $x_{it} = (x_{it,1}, \ldots, x_{it,k})^T$ is a $k \times 1$ vector of exogenous variables,
- ε_{it} is an error term and assumed to be $i.i.d.$ with, for all $i = 1, \ldots, n$, and all $t = 1, \ldots, T$, we have

$$E(\varepsilon_{it}) = 0, \; E(\varepsilon_{it}^2) = \sigma^2.$$

Topic 2: Individual effects

- There are many names for the scalars α_i:
  1. unobserved effects
  2. individual effects
  3. unobserved components
  4. latent variables (for random effects models)
- If α_i is treated as a random variable, the proposed panel model is called random effect model and α_i is called a random effect.
- If α_i is treated as a fixed as a parameter for each cross section observation i, the proposed panel model is called fixed effect model and α_i is called a fixed effect.
- However, in many (in particular microeconometric) applications, "fixed effect" does not always assume α_i has to be non-random. α_i can still be random but has to be correlated with x_{it}.

- Thus, the key to distinguish one from another involves whether or not the *unobserved* α_i is uncorrelated with the *observed* explanatory variables x_{it}.
- For fixed effect model, we assume $Cov(x_{it}, \alpha_i) \neq 0$ and $E(\varepsilon_{it}|\alpha_i) \neq 0$.
- For random effect, we assume α_i is independent to $\mathbf{x}_i = [x_{i1}^T, \ldots, {}_{iT}^T]^T$ such that we have $E(\alpha_i) = E(\varepsilon_{it}) = 0$, $E(\alpha_i \mathbf{x}_i) = E(\alpha_i \varepsilon_{it}) = 0$.
- Why? Assuming α_i is a random variable, we have

$$f(\mathbf{y}_i, \mathbf{x}_i) = f(\mathbf{y}_i|\mathbf{x}_i) f(\mathbf{x}_i) = \left[\int f(\mathbf{y}_i|\mathbf{x}_i, \alpha_i) f(\alpha_i|\mathbf{x}_i) d_{\alpha_i} \right] f(\mathbf{x}_i),$$

  where $\mathbf{y}_i = [y_{i1}, \ldots, y_{iT}]^T$, a $T \times 1$ vector.
- We need to know $f(\alpha_i|\mathbf{x}_i)$ and independence implies

$$f(\alpha_i|\mathbf{x}_i) = f(\alpha_i).$$

Topic 2: Fixed or random effects? An example

- Let us consider the simple example of the Cobb Douglass production function.

$$Output_{it} = \beta_1 Capital_{it} + \beta_2 Labor_{it} + \alpha_i + \varepsilon_{it}.$$

- In this case, α_i can be the unobserved effect on Total factor productivity (TFP) due to unobserved country-level omitted factor (climate, institutions, organization, etc..).
- If we expect that the more a country is productive, the more it invests in capital, in other words α_i is positively correlated with $Capital_{it}$, then we should consider a fixed effect panel model.
- We can also employ a Hausman's test (1978) to test the null hypothesis of $cov(x_{it}, \alpha_i) = 0$ for all i and t. This is a specification test (fixed or random) for the unobserved effects. We will discuss it more in Topic 3.

Topic 3: A vector form of a fixed effect model

- Define

$$\underset{T\times 1}{\boldsymbol{y_i}} = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \dots \\ y_{iT} \end{bmatrix} \quad \underset{T\times k}{\boldsymbol{x_i}} = \begin{bmatrix} x_{i1,1} & x_{i1,2} & \dots & x_{i1,k} \\ x_{i2,1} & x_{i2,2} & \dots & x_{i2,k} \\ \dots & & & \\ x_{iT,1} & x_{iT,2} & \dots & x_{iT,k} \end{bmatrix}$$

$$\underset{T\times 1}{\boldsymbol{e}} = \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} \quad \underset{T\times 1}{\boldsymbol{\varepsilon_i}} = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \dots \\ \varepsilon_{iT} \end{bmatrix} \quad \underset{k\times 1}{\boldsymbol{\beta}} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_k \end{bmatrix}$$

- We can rewrite our fixed model as

$$\boldsymbol{y_i} = \boldsymbol{e}\alpha_i + \boldsymbol{x_i}\boldsymbol{\beta} + \boldsymbol{\varepsilon_i}, \tag{1}$$

for all $i = 1, \dots, n$, where α_i is assumed to be a fixed term or a random variable satisfying $E(\alpha_i|\boldsymbol{x_i}) \neq 0$.

- Stacking all n observations, we can rewrite model (1) as

$$\begin{bmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \\ \dots \\ \boldsymbol{y_n} \end{bmatrix} = \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{0} \\ \dots \\ \boldsymbol{0} \end{bmatrix} \alpha_1 + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{e} \\ \dots \\ \boldsymbol{0} \end{bmatrix} \alpha_2 + \dots + \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{0} \\ \dots \\ \boldsymbol{e} \end{bmatrix} \alpha_n + \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \dots \\ \boldsymbol{x_n} \end{bmatrix} \beta + \begin{bmatrix} \boldsymbol{\varepsilon_1} \\ \boldsymbol{\varepsilon_2} \\ \dots \\ \boldsymbol{\varepsilon_n} \end{bmatrix} \tag{2}$$

- For model (2), we assume $E(\boldsymbol{\varepsilon_i}) = 0$, $E(\boldsymbol{\varepsilon_i}\boldsymbol{\varepsilon_i}^T) = \sigma^2 I_T$, and $E(\boldsymbol{\varepsilon_i}\boldsymbol{\varepsilon_j}^T) = 0$ if $i \neq j$.
- Under the assumed properties of $\boldsymbol{\varepsilon_i}$, we know that the ordinary least-squares (OLS) estimator of model (2) is the best linear unbiased estimator (BLUE).

Topic 3: LSDV estimator derivation

- The OLS estimators of α_i and $\boldsymbol{\beta}$ ar obatined by minimizing

$$SSR = \sum_{i=1}^{n} \left(\boldsymbol{y_i} - \boldsymbol{e}\alpha_i - \boldsymbol{x_i}\boldsymbol{\beta}\right)^T \left(\boldsymbol{y_i} - \boldsymbol{e}\alpha_i - \boldsymbol{x_i}\boldsymbol{\beta}\right). \tag{3}$$

- Taking partial derivatives of SSR w.r.t. α_i, the FOC gives

$$\widehat{\alpha}_i^{LSDV} = \bar{\boldsymbol{y}}_{\boldsymbol{i}} - \bar{\boldsymbol{x}}_{\boldsymbol{i}}\boldsymbol{\beta},\ \ i = 1, \ldots, n, \tag{4}$$

where $\bar{\boldsymbol{y}}_{\boldsymbol{i}} = \frac{1}{T}\sum_{t=1}^{T} y_{it},\ \ \bar{\boldsymbol{x}}_{\boldsymbol{i}} = \frac{1}{T}\sum_{t=1}^{T} x_{it}^T$.

- Substituting (4) into (3) and taking the partial derivative of SSR w.r.t. $\boldsymbol{\beta}$, we have

$$\widehat{\boldsymbol{\beta}}^{LSDV} = \left[\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it} - \bar{\boldsymbol{x}}_{\boldsymbol{i}}\right)\left(x_{it} - \bar{\boldsymbol{x}}_{\boldsymbol{i}}\right)^T\right]^{-1}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}\left(x_{it} - \bar{\boldsymbol{x}}_{\boldsymbol{i}}\right)\left(y_{it} - \bar{\boldsymbol{y}}_{\boldsymbol{i}}\right)\right]$$

Topic 3: LSDV remarks

- The computational procedure for estimating the slope parameters in this model does not require that the dummy variables for the individual (and/or time) effects actually be included in the matrix of explanatory variables.
- We only need to take the following steps:
  1. the sample average of time-series observations separately for each cross-sectional unit
  2. transform the observed variables by subtracting out these means
  3. apply the least squares method to the transformed data
- The above demeaned transformation is called a "within-group" transformation and, thus, the above estimator is also called "within-group estimator".

Topic 3: Within-group estimator

- The foregoing procedure is equivalent to premultiplying the i^{th} equation $\boldsymbol{y_i} = \boldsymbol{e}\alpha_i + \boldsymbol{x_i}\boldsymbol{\beta} + \boldsymbol{\varepsilon_i}$ by a $T \times T$ idempotent (covariance) transformation matrix $Q = I_T - \frac{1}{T}\boldsymbol{e}\boldsymbol{e}^T$ to "sweep out" the individual effect α_i such that, for $i = 1, \ldots, n$, we have

$$Q\boldsymbol{y_i} = Q\boldsymbol{e}\alpha_i + Q\boldsymbol{x_i}\boldsymbol{\beta} + Q\boldsymbol{\varepsilon_i} = \boxed{Q\boldsymbol{x_i}\boldsymbol{\beta} + Q\boldsymbol{\varepsilon_i}}. \tag{5}$$

- Applying the OLS procedure to (5), we have

$$\widehat{\boldsymbol{\beta}}^W = \left[\sum_{i=1}^{n} \boldsymbol{x_i}^T Q \boldsymbol{x_i}\right]^{-1} \left[\sum_{i=1}^{n} \boldsymbol{x_i}^T Q \boldsymbol{y_i}\right].$$

- $\widehat{\boldsymbol{\beta}}^W$ is identical to $\widehat{\boldsymbol{\beta}}^{LSDV}$!

Topic 3: Properties of the LSDV and within-group estimator

1. The LSDV ($\widehat{\boldsymbol{\beta}}^{LSDV}$) and within-group estimator ($\widehat{\boldsymbol{\beta}}^{W}$) of $\boldsymbol{\beta}$ is unbiased and consistent when either n, or T, or both tend to infinity.
2. The LSDV estimator ($\widehat{\alpha}_i^{LSDV}$) for the unobserved effects is unbiased but only consistent when $T \longrightarrow \infty$. It is inconsistent if T is fixed.
3. The asymptotic variance-covariance matrix of the $\widehat{\boldsymbol{\beta}}^{W}$ (or $\widehat{\boldsymbol{\beta}}^{LSDV}$) is give by

$$Var(\widehat{\boldsymbol{\beta}}^{W}) = \sigma^2 \left[\sum_{i=1}^{n} \boldsymbol{x_i}^T Q \boldsymbol{x_i} \right]^{-1}.$$

4. When $var(\varepsilon_{it}) = \sigma_i^2$, within-group (or LSDV) estimator will be no longer BLUE. However, it remains consistent.

Topic 3: A Stata example to estimate a fixed effects model

- Let us consider a simple panel regression model for the wages of young working women who had an age of 14–26 years in 1968. These data are collected within the "National Longitudinal Survey" over the years 1968-1988 (with gaps). There are 28534 observations in total.

- **Step 1**: use the "`use`" command to load a sample dataset, and then use the "`xtset`" command to declare the dataset as a panel dataset with a group variable and a time variable:

```
use http://www.stata-press.com/data/r14/nlswork.dta,
clear
xtset idcode year
```

- **Step 2**: run a fixed effect regression using the "`xtreg`" command with the "`fe`" option , which specifies fixed effects. In this example. Then, regress log wages on age and total experience, with fixed effects for individuals:

```
xtreg lnwage age grade ttlexp, fe.
```