Topics in Applied Econometrics for Public Policy

TA Session 6

Sergi Quintana

Barcelona School of Economics

May 30, 2023

Plan for Today

- Problem Set Correction.

Non linear panel data

Non linear panel data is problematic because of the individual specific effects.

- Depending on it's nature, we will use some estimation methods or others.
- Under the fixed effect assumption, the incidental parameter problem will arise.

Non-linear panel data - Conditional Mean

The individual specific term can enter in different ways:

$$E[y_{it}|x_{it}] = g(\alpha_i, x_{it}, \beta)$$

The effect can be:

- Additive:

$$g(\alpha_i, x_{it}, \beta) = \alpha_i + g(x_{it}, \beta)$$

- Multiplicative:

$$g(\alpha_i, x_{it}, \beta) = \alpha_i g(x_{it}, \beta)$$

- Single-index:

$$g(\alpha_i, x_{it}, \beta) = g(\alpha_i + x_{it}'\beta)$$

Fixed Effects - Incidental Parameters Problem

Under the fixed effects models the individual parameter α_i might be correlated with the regressors x_{it}.

The Indicidental Parameters Problem

For the estimation of the mean, we have the common parameters β and the incidental parameters α, and we are only interested in consistently estimating β.

The incidental parameters are inconsistently estimated provided that T is fixed.

Since the model is non-linear, this will also bias the common parameters.

Fixed Effects - Incidental Parameters Problem

There are multiple ways to address this problem, depending on the nature of the $g()$ function.

Mean-differenced transformation:

$$E[(y_{it} - \bar{y}_i) - (g(x_{it}'\beta) - \bar{g}_i(\beta))|x_{i1,\ldots,x_{iT}}]$$

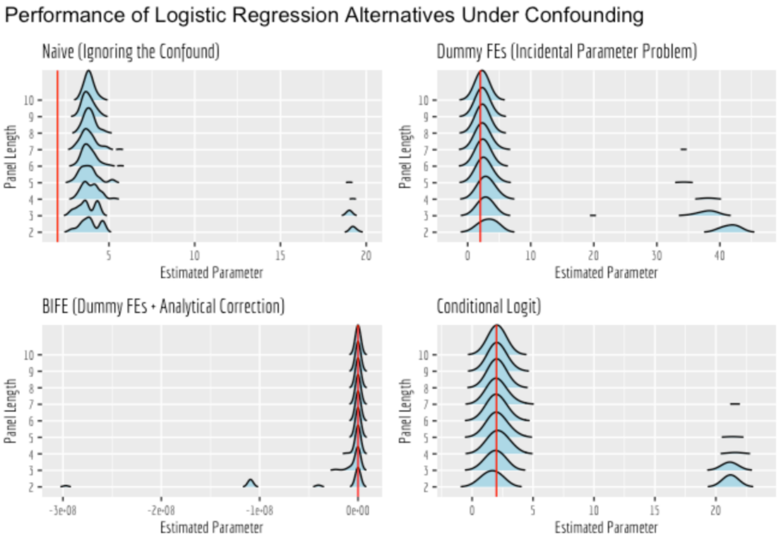
where $\bar{g}_i(\beta) = 1/T \sum_T g(x_{it}'\beta)$.

First-differences transformation:

$$E[(y_{it} - y_{it-1}) - (g(x_{it}'\beta) - g(x_{it-1}'\beta))|x_{i1,\ldots,x_{iT}}]$$

Dummy variable estimation.

Fixed Effects- Incidental Parameters Problem



Binary Outcome Models - Random Effects Model

The idea is to treat the individual specific term as a random variable from a specific distribution.

The individual term is eliminated by integrating over its distribution. For example, if we assume α_i follows a $N(0, \sigma^2)$ then:

$$f\left(\mathbf{y}_i|\mathbf{X}_i, \beta, \sigma_\alpha^2\right) = \int f(\mathbf{y}_i|\mathbf{X}_i, \alpha_i, \beta) \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(\frac{-\alpha_i}{2\sigma_\alpha^2}\right)^2 d\alpha_i$$

Notice this integral has no analytical solution, but numerical integration works well since we only need to integrate over one variable.

Binary Outcome Models - Fixed Effects Model

Under the logit framework, one way to eliminate α_i is the conditional logit, which conditions on the total number of outcomes of an invididual that equal to 1. Consider the example of two periods:

Binary Outcome Models - Fixed Effects Model

We demonstrate this in the simplest case of two time periods. Condition on $y_{i1} + y_{i2} = 1$, so that $y_{it} = 1$ in exactly one of the two periods. Then, in general,

$$\Pr(y_{i1} = 0, y_{i2} = 1 | y_{i1} + y_{i2} = 1) = \frac{\Pr(y_{i1} = 0, y_{i2} = 1)}{\Pr(y_{i1} = 0, y_{i2} = 1) + \Pr(y_{i1} = 1, y_{i2} = 0)} \quad (18.8)$$

Now $\Pr(y_{i1} = 0, y_{i2} = 1) = \Pr(y_{i1} = 0) \times \Pr(y_{i2} = 1)$, assuming that y_{1i} and y_{2i} are independent given α_i and $\mathbf{x}_{it}$. For the logit model (18.6), we obtain

$$\Pr(y_{i1} = 0, y_{i2} = 1) = \frac{1}{1 + \exp(\alpha_i + \mathbf{x}_{i1}'\beta)} \times \frac{\exp(\alpha_i + \mathbf{x}_{i2}'\beta)}{1 + \exp(\alpha_i + \mathbf{x}_{i2}'\beta)}$$

Similarly,

$$\Pr(y_{i1} = 1, y_{i2} = 0) = \frac{\exp(\alpha_i + \mathbf{x}_{i1}'\beta)}{1 + \exp(\alpha_i + \mathbf{x}_{i1}'\beta)} \times \frac{1}{1 + \exp(\alpha_i + \mathbf{x}_{i2}'\beta)}$$

Binary Outcome Models - Fixed Effects Model

$$
\begin{aligned}
\Pr(y_{i1} = 0, y_{i2} = 1 | y_{i1} + y_{i2} = 1) \\
&= \exp(\alpha_i + \mathrm{x}_{i2}'\beta) / \{\exp(\alpha_i + \mathrm{x}_{i1}'\beta) + \exp(\alpha_i + \mathrm{x}_{i2}'\beta)\} \\
&= \exp(\mathrm{x}_{i2}'\beta) / \{\exp(\mathrm{x}_{i1}'\beta) + \exp(\mathrm{x}_{i2}'\beta)\} \\
&= \exp\{(\mathrm{x}_{i2} - \mathrm{x}_{i1})'\beta\} / [1 + \exp\{(\mathrm{x}_{i2} - \mathrm{x}_{i1})'\beta\}]
\end{aligned}
$$

Notice that time invariant regressors will not be identified. Even if we overcome the incidental parameters problem, we might not be able to estimate the marginal effects if individual fixed effects are multiplicative:

$$\partial E[y_{it} | x_{it}, \alpha_i] / \partial x_{it} = \alpha_i \beta$$

Binary Outcome Models - Pooled Logit

Is the usual cross-section model, assuming $\alpha_i = \alpha$.

$$p(y_{it} = 1 | x_{it}) = \Delta(x_{it}'\beta)$$

We will just correct for individual correlation over time using cluster at the individual level.