Empirical Panel Data: Lecture 8

Instructor: Chaoyi Chen
NJE & MNB

April 25, 2023

Topic 4: Introduction to dynamic panel data model

- Before, the regressors **exclude** the lagged dependent variable. Therefore, this type of panel data model is called "static" panel data model.

- The LSDV estimator is consistent for the static panel data model with fixed or random effects when n increases for fixed T.

- We now consider a dynamic panel data model (it contains (at least) one lagged dependent variables). For simplicity, let us consider an augmented panel AR(1) process

$$y_{it} = \gamma y_{i,t-1} + \beta^\top x_{it} + \alpha_i + \varepsilon_{it}$$

- $i = 1, \ldots, n$ and $t = 1, \ldots T$
- α_i is the (unobserved) individual effects
- ε_{it} is the error term with $E(\varepsilon_{it}) = 0$, $E(\varepsilon_{it}\varepsilon_{js}) = \sigma_\varepsilon^2$ for $i = j$ and $t = s$ and $E(\varepsilon_{it}\varepsilon_{js}) = 0$ otherwise.

Topic 4: Dynamic panel issues

- If lagged dependent variables appear as explanatory variables, strict exogeneity of the regressors no longer holds. The LSDV is no longer consistent when n tends to infinity and T is fixed.
- The bias of the LSDV estimator in a dynamic model is generally known as dynamic panel bias or Nickell's bias (Nickell 1981).
- The **initial values of a dynamic process** raise another problem. It turns out that with a random-effects formulation, the interpretation of a model depends on the assumption of initial observation.
- The consistency property of the maximum likelihood estimator (MLE) and the GLS estimator also depends on the way in which T and n tend to infinity.

Topic 4: LSDV estimator of an AR(1) example

- Without loss of generality, we consider a simple dynamic panel data model

$$y_{it} = \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it}$$

- Assume $\gamma < 1$ ensures the stationarity of y_{it} over time
- Assume that the initial condition y_{i0} is observed for all i.
- The LSDV estimator is given by

$$\widehat{\alpha}_i^{LSDV} = T^{-1} \sum_{t=1}^{T} \left(y_{it} - \widehat{\gamma} y_{i,t-1} \right) = \bar{\boldsymbol{y}}_i - \widehat{\gamma} \bar{\boldsymbol{y}}_{i,-1}$$

$$\widehat{\gamma}^{LSDV} = \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it-1} - \bar{\boldsymbol{y}}_{i,-1} \right) \left(y_{it} - \bar{\boldsymbol{y}}_i \right)}{\sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it-1} - \bar{\boldsymbol{y}}_{i,-1} \right)^2}$$

where $\bar{\boldsymbol{y}}_{i,-1} = T^{-1} \sum_{t=1}^{T} y_{i,t-1}$

Topic 4: Nickell's bias

- By simple calculation, the bias of the LSDV estimator can be written as:

$$\widehat{\gamma}^{LSDV} - \gamma = \frac{(nT)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\left(y_{it-1} - \bar{\boldsymbol{y}}_{i,-1}\right)\left(\varepsilon_{it} - \bar{\boldsymbol{\varepsilon}}_i\right)\right]}{(nT)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it-1} - \bar{\boldsymbol{y}}_{i,-1}\right)^2} \equiv \frac{A_{nT}}{B_{nT}}$$

- For A_{nT}, by using $\varepsilon_{it} - \bar{\boldsymbol{\varepsilon}}_i = (1 - T^{-1})\varepsilon_{it} - T^{-1}\sum_{s\neq t}\varepsilon_{is}$ and doing some algebra, we have

$$\begin{aligned} A_{nT} &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\left(y_{it-1} - \bar{\boldsymbol{y}}_{i,-1}\right)\left(\varepsilon_{it} - \bar{\boldsymbol{\varepsilon}}_i\right)\right] \\ &= \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} y_{i,t-1}\varepsilon_{it} - \frac{1}{n}\sum_{i=1}^{n}\bar{\boldsymbol{y}}_{i,-1}\bar{\boldsymbol{\varepsilon}}_i \end{aligned}$$

- As ε_{it} is $i.i.d.$ and is uncorrelated with μ_i, by applying LLN, we have

$$\boxed{\underset{n\longrightarrow\infty}{plim}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} y_{i,t-1}\varepsilon_{it} = E\left(y_{i,t-1}\varepsilon_{it}\right) = 0} \quad (1)$$

Topic 4: Decomposing A_{nT}

- By substitution method, we can show

$$
\begin{aligned}
y_{i,t-1} &= \gamma y_{i,t-2} + \alpha_i + \varepsilon_{i,t-1} \\
&= \gamma \left[\gamma y_{i,t-3} + \alpha_i + \varepsilon_{i,t-2}\right] + \alpha_i + \varepsilon_{i,t-1} \\
&= \gamma^2 y_{i,t-3} + (1+\gamma)\alpha_i + \varepsilon_{i,t-1} + \gamma\varepsilon_{i,t-2} \\
&= \ldots \\
&= \gamma^{t-1} y_{i,0} + \left(1 + \gamma + \gamma^2 + \ldots + \gamma^{t-2}\right)\alpha_i + \varepsilon_{i,t-1} + \gamma\varepsilon_{i,t-2} + \ldots + \gamma^{t-2}\varepsilon_{i,1} \\
&= \boxed{\gamma^{t-1} y_{i,0} + \frac{1-\gamma^{t-1}}{1-\gamma}\alpha_i + \varepsilon_{i,t-1} + \gamma\varepsilon_{i,t-2} + \ldots + \gamma^{t-2}\varepsilon_{i,1}}
\end{aligned}
$$

- Note that, above equation implies

$$
T\bar{\boldsymbol{y}}_{i,-1} = \sum_{t=1}^{T} y_{i,t-1} = \frac{1-\gamma^{T-1}}{1-\gamma} y_{i,0} + \frac{T-1-T\gamma+\gamma^T}{(1-\gamma)^2}\alpha_i
$$

$$
+ \varepsilon_{i,T-1} + \frac{1-\gamma^2}{1-\gamma}\varepsilon_{i,T-2} + \ldots + \frac{1-\gamma^{T-1}}{1-\gamma}\varepsilon_{i,1}
$$

Topic 4: Decomposing A_{nT} Cont.

- Therefore, for the second term of A_{nT}, we have

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\bar{\boldsymbol{y}}_{i,-1}\bar{\boldsymbol{\varepsilon}}_i \\
&= \frac{1}{nT}\sum_{i=1}^{n}\left(\frac{1-\gamma^{T-1}}{1-\gamma}y_{i,0}+\varepsilon_{i,T-1}+\frac{1-\gamma^2}{1-\gamma}\varepsilon_{i,T-2}+\ldots+\frac{1-\gamma^{T-1}}{1-\gamma}\varepsilon_{i,1}\right) \\
&\times\frac{1}{T}\left(\varepsilon_{i,1}+\ldots,\varepsilon_{i,T}\right) \\
&= \frac{1}{nT^2}\sum_{i=1}^{n}\left(\varepsilon_{i,T-1}^2+\frac{1-\gamma^2}{1-\gamma}\varepsilon_{i,T-2}^2+\ldots+\frac{1-\gamma^{T-1}}{1-\gamma}\varepsilon_{i,1}^2\right) \\
&\xrightarrow{p}\frac{\sigma_\varepsilon^2}{T^2}\left(1+\frac{1-\gamma^2}{1-\gamma}+\ldots+\frac{1-\gamma^{T-1}}{1-\gamma}\right)=\frac{\sigma_\varepsilon^2}{T^2}\frac{T-1-T\gamma+\gamma^T}{(1-\gamma)^2}
\end{aligned}
$$

- Thus, combine the above result with (1), we have

$$
\boxed{A_{nT}\xrightarrow{p}\frac{\sigma_\varepsilon^2}{T^2}\frac{T-1-T\gamma+\gamma^T}{(1-\gamma)^2}} \tag{2}
$$

Topic 4: Decomposing B_{nT}

- If you trust me :), applying similar arguments, for B_{nT}, we can show

$$\boxed{B_{nT} \xrightarrow{p} \frac{\sigma_\varepsilon^2}{(1-\gamma)^2}\left[1 - \frac{1}{T} - \frac{2\gamma}{(1-\gamma)^2}\left(\frac{T-1-T\gamma+\gamma^T}{T}\right)\right]} \tag{3}$$

- Try at home!

- Combining (2) and (3) we have

$$\underset{n\longrightarrow\infty}{plim}\left(\widehat{\gamma}^{LSDV}-\gamma\right) = -\frac{\frac{\sigma_\varepsilon^2}{T^2}\frac{T-1-T\gamma+\gamma^T}{(1-\gamma)^2}}{\frac{\sigma_\varepsilon^2}{(1-\gamma)^2}\left[1-\frac{1}{T}-\frac{2\gamma}{(1-\gamma)^2}\left(\frac{T-1-T\gamma+\gamma^T}{T}\right)\right]}$$

$$= -\frac{1+\gamma}{T-1}\left(1-\frac{1}{T}\frac{(1-\gamma^T)}{(1-\gamma)}\right)\left(1-\frac{2\gamma}{(1-\gamma)(T-1)}\left(1-\frac{1-\gamma^T}{T(1-\gamma)}\right)\right)^{-1}$$

Topic 4: Nickell's bias remarks

- As $n \longrightarrow \infty$ and $T \longrightarrow \infty$, we obtain

$$A_{nT} \xrightarrow{p} 0 \text{ and } B_{nT} \xrightarrow{p} \frac{\sigma_\varepsilon^2}{(1-\gamma)^2}$$

- Therefore, LSDV estimator is still consistent
- However, if T is fixed, LSDV estimator is biased asymptotically.
- What causes the dynamic bias? Remember that the LSDV estimator is equivalent to the within-group (or FE) estimator and is the OLS estimator from the transformed model

$$y_{i,t} - \bar{\boldsymbol{y}}_i = \gamma\left(y_{i,t-1} - \bar{\boldsymbol{y}}_{i,-1}\right) + \varepsilon_{i,t} - \bar{\boldsymbol{\varepsilon}}_i$$

- There is a correlation between the regressor and the error term of an order T^{-1}.

Topic 4: Intuition of the dynamic bias

- Intuition: Both $\bar{\boldsymbol{y}}_{i,-1}$ and $\bar{\varepsilon}_i$ depends on past value of ε_{it}.

$$Cov(\bar{\boldsymbol{y}}_{i,-1}, \bar{\varepsilon}_i) = Cov\left(\frac{1}{T}\sum_{t=1}^{T} y_{i,t-1}, \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}\right)$$
$$= \frac{1}{T^2} Cov\left((y_{i,1} + \ldots + y_{iT-1}), (\varepsilon_{i1} + \ldots \varepsilon_{iT})\right).$$

- Recall, in time-series scenario, AR(1) model equals MA(∞) model. Applying similar tenique here and assuming $y_{i,0} = 0$ for all i, we can show

$$y_{i,t} = \frac{1-\gamma^t}{1-\gamma}\alpha_i + \sum_{j=1}^{t}\gamma^{j-1}\varepsilon_{i,t-j+1}.$$

- Apparently, for all $t = 1, \ldots, T-1$, y_{it} correlates with $\varepsilon_{i,t-1}, \ldots, \varepsilon_{i,1}$ while the correlation is fading in longer period. This implies $\bar{\boldsymbol{y}}_{i,-1}$ is **not** orthogonal to $\bar{\varepsilon}_i$ given T is fixed.

Monte Carlo simulations: Design

- To have a visual impression on the dynamic panel bias, we consider the following data generation process (DGP)

$$y_{it} = \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it}$$

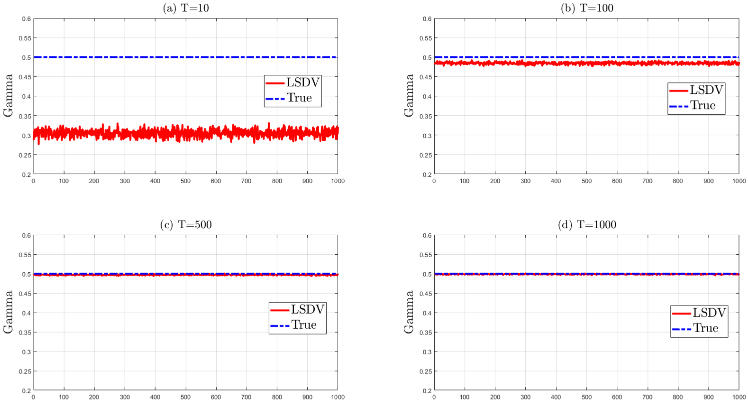
- $y_{i0} = 0$ for all i
- $\alpha_i \sim U(-1, 1)$.
- $\varepsilon_{it} \sim iidN(0, 1)$

- We repeat 1000 times.
- We fix $n = 1000$ but vary T.
- We compute average bias of the LSDV estimator based on 1000 simulations

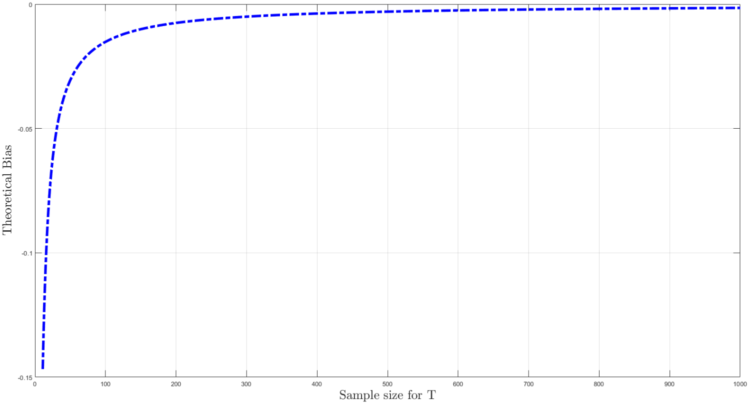
$$\text{average bias} = \frac{1}{1000} \sum_{s=1}^{1000} \left(\widehat{\gamma}_s^{LSDV} - \gamma \right)$$

- $\widehat{\gamma}_s^{LSDV}$ is the LSDV estimator from s^{th} simulation.

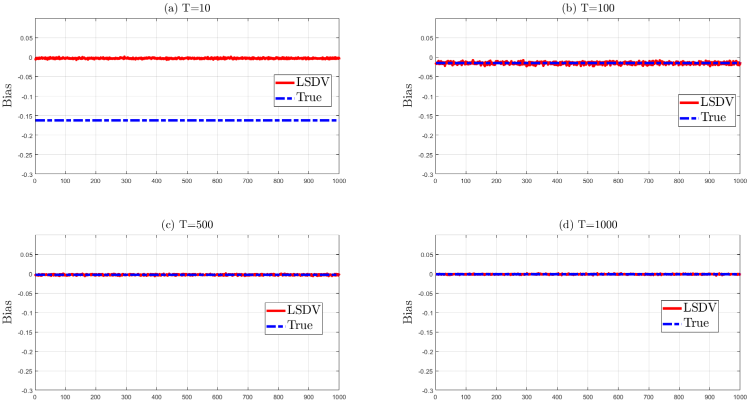
Monte Carlo simulations: LSDV estimates and ture parameter



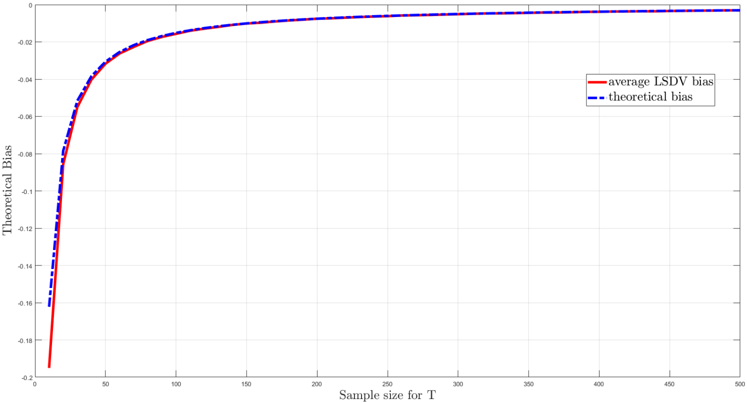
Monte Carlo simulations: Theoretical bias



Monte Carlo simulations: LSDV bias and theoretical bias



Monte Carlo simulations: Average LSDV bias and theoretical bias



Reference

Nickell, Stephen (1981). Biases in dynamic models with fixed effects. *Econometrica* **49**(6), 1417.