Empirical Panel Data: Lecture 9

Instructor: Chaoyi Chen
NJE & MNB

May 2, 2023

Topic 4: Dynamic panel bias

- The LSDV for the dynamic individual-effects model is biased due to the presence of $Cov(\bar{\boldsymbol{y}}_{i,-1}, \bar{\varepsilon}_i) \neq 0$ if T is small.
- This is known as an **endogenous** problem as the conditional mean of error is no longer zero in this case.
- To fix this issue, we need to tackle the endogeneity issue.
- Solutions: Find a valid instrument variable and use IV or GMM approach.
- We will have a review lecture today to review/study the endogenous problem and the IV method in the classical linear regression model.

Review: The endogeneity

- Consider the classical linear regression model as studied in Lecture 1:

$$\boldsymbol{y} = \boldsymbol{x}\beta + \mu.$$

- Assume that Gauss-Markov Assumption 1-3 hold, but Assumption 4 does **not** hold (i.e., $E(\boldsymbol{\mu}|\boldsymbol{x}) \neq 0$).
- Under this setup, we call $\boldsymbol{x}$ is an **endogenous** regressor.
- If the regressors are endogenous, the OLS estimator of β is biased. This bias is also called endogenity bias
- Why? Recall $E(\widehat{\beta}^{OLS}) = \beta + E\left[\left(\boldsymbol{x}^\top \boldsymbol{x}\right)^{-1} \boldsymbol{x}^\top \mu\right]$ and the second term is no longer zero.

Review: Instrument variable

- What is the solution to the endogeneity issue? A standard approach in the existing literature uses the instruments.
- Assume now we can find a set of variable $\boldsymbol{z}$, which is of $n \times m$. The variable $\boldsymbol{z}$ is called **instruments** or **instrumental variable** if $\boldsymbol{z}$
  1. is **uncorrelated** with μ, $E(\mu|\boldsymbol{z}) = 0$. This is called Exogeneity condition.
  2. is **correlated** with the independent variables $\boldsymbol{x}$, $E(\boldsymbol{xz}) \neq 0$. This is called the relevance condition.

Review: Check the exogeneity and relevance condition

- The validity of the IV relies on the exogeneity and relevance condition.
- How to check? In most cases intuition is enough!
- For example, consider an empirical growth model for a country

$$y_t = \beta_0 + \beta_1 L_t + \beta_2 K_t + \varepsilon_t, \tag{1}$$

  where y is the output, L is labor and K is capital.
- Apparently, L_t and K_t cause y_t. However, there is also an inverse causality, which implies y_t causes L_t and K_t as well - the presence of contemporaneous bidirectional causality! If that is true, model (1) suffers from the endogenous issue.
- How to propose a Valid IV? Lag K and L would be perfect in this case! Why? 1. the current output will not affect the past labour and capital. The exogeneity condition holds. 2. the past labour and capital correlate with the present ones. The relevance condition holds.

Review: IV identification

- How do we use IV to solve the issue? First, we need to ensure the model is identified.
- **Definition:** The system is identified if there exists a unique vetor of β such that

$$E\left(z_i\left(y_i - x_i^\top \beta\right)\right) = 0$$

where $z_i = [z_{i1}, \ldots, z_{im}]^\top$ and $x_i = [x_{i1}, \ldots, x_{ik}]^\top$. For that, we have the following conditions:

1. if $m < k$, the model is not identified;
2. if $m = k$, the model is just-identified or exact-identified;
3. if $m > k$, the model is over identified.

Review: IV identification remarks

- Remark 1: With **under-identification**, we can not use z to estimate the model due to the number of equations being less than the number of unknowns (unique solution).

- Remark 2: With **just-identification**, the number of equations equals the number of unknowns (unique solution). We can use IV estimator.

- Remark 3: With **over-identification**, the number of equations being greater than the number of unknowns (unique solution). Two equivalent solutions:

  1. Using m choices to select k linear combinations of the instruments to have a unique solution. This is called Two-Stage Least Squares (2SLS).

  2. Set the sample analog of the moment conditions as close as possible to zero, i.e. minimize the distance between the sample analog and zero given a metric (optimal metric or optimal weighting matrix?). This is called Generalized Method of Moments (GMM).

Review: Just-identification and IV estimator

- Now, we consider the $m = k$ and derive the IV estimator
- We start from the conditional moment condition $E(\boldsymbol{\mu}|\boldsymbol{z}) = 0$

$$E(\boldsymbol{z}^\top \boldsymbol{\mu}) = E(\boldsymbol{z}(\boldsymbol{y} - \boldsymbol{x}\beta)) = 0 \Longrightarrow E(\boldsymbol{z}^\top \boldsymbol{y}) = E(\boldsymbol{z}^\top \boldsymbol{x})\beta$$

$$\Longrightarrow \beta = E\left(\boldsymbol{z}^\top \boldsymbol{x}\right)^{-1} E\left(\boldsymbol{z}^\top \boldsymbol{y}\right)$$

- Remark 1: The **relevance condition** ensures the existence of the inverse of $E\left(\boldsymbol{z}^\top \boldsymbol{x}\right)$. If $E\left(\boldsymbol{z}^\top \boldsymbol{x}\right) \longrightarrow 0$, we will have a weak instrument problem.
- Therefore, if $m = k$, the **instrumental variable (IV) estimator** $\widehat{\beta}_{IV}$ of β is defined as

$$\widehat{\beta}_{IV} = \left(\boldsymbol{z}^\top \boldsymbol{x}\right)^{-1} \boldsymbol{z}^\top \boldsymbol{y}.$$

- Remark 2: $\widehat{\beta}_{IV}$ is an unbiased estimator of β. Try to prove unbiasedness at home, and from the derivation, you will use and know why we need the **exogeneity condition**!

Review: Asymptotic results of IV estimator

- Under some regularity conditions, we can show

  1. Consistency:

$$\widehat{\beta}_{IV} \xrightarrow{p} \beta$$

  2. Asymptotic normality:

$$\sqrt{n}\left(\widehat{\beta}_{IV} - \beta\right) \xrightarrow{d} N\left(0_{k\times 1}, \Omega_z\right)$$

where $\Omega_z = \sigma_\mu^2 E\left[z_i x_i^\top\right]^{-1} E\left[z_i z_i^\top\right] E\left[z_i x_i^\top\right]^{-1}$.

- Remark 1: If IV is weak, the quantity of $E\left[z_i x_i^\top\right]$ goes to zero and its inverse goes to infinity. As a result, we will have a huge variance.

- Remark 2: As usual, the estimator of the variance of the error terms is:

$$\widehat{\sigma}^2 = \frac{\widehat{\mu}^\top \widehat{\mu}}{n-k} = \frac{1}{n-k}\sum_{i=1}^{n}\left(y_i - x_i^\top \widehat{\beta}_{IV}\right)^2.$$

Review: Over-identification and 2SLS

- If $m > k$, $\boldsymbol{z}$ contains more variables than $\boldsymbol{x}$. Then, the preceding derivation is unusable since $E(\boldsymbol{z}^\top \boldsymbol{x})$ is of $m \times k$ but $E(\boldsymbol{z}^\top \boldsymbol{y})$ is of $m \times 1$. It is impossible to derive $k \times 1$ vector of β.
- Solution:
  Step 1: find $k \times 1$ linear combinations from m variables.
  Step 2: Use the $k \times 1$ linear combinations in the preceding derivation.
- This method splits the estimation into two steps, where the first step regress $\boldsymbol{z}$ on $\boldsymbol{x}$ and the second step regress the linear combinations on $\boldsymbol{y}$. Thus, we call it a two-step least squares estimator (2SLS).

Review: Over-identification and linear combination

- In the first step, which is a proper linear combination to choose? **The projection**!
- We regress $\boldsymbol{x}$ on the space of $\boldsymbol{z}$ and obtain

$$\widehat{\boldsymbol{x}} = \boldsymbol{P}_z \boldsymbol{x} = \boldsymbol{z}(\boldsymbol{z}^\top \boldsymbol{z})^{-1} \boldsymbol{z}^\top \boldsymbol{x},$$

  where $\boldsymbol{P}_z$ is the projection matrix for space $\boldsymbol{z}$.

- Then, in the second step, we use $\widehat{\boldsymbol{x}}$ as instruments to have

$$\begin{aligned} \widehat{\beta}_{2SLS} &= \left(\widehat{\boldsymbol{x}}^\top \boldsymbol{x}\right)^{-1} \widehat{\boldsymbol{x}}^\top \boldsymbol{y} \\ &= \left(\boldsymbol{x}^\top \boldsymbol{z} \left(\boldsymbol{z}^\top \boldsymbol{z}\right)^{-1} \boldsymbol{z}^\top \boldsymbol{x}\right)^{-1} \boldsymbol{x}^\top \boldsymbol{z} \left(\boldsymbol{z}^\top \boldsymbol{z}\right)^{-1} \boldsymbol{z}^\top \boldsymbol{y}. \end{aligned}$$

Topic 4: Dynamic panel bias and IV approach

- Now, we turn back to a generalized dynamic panel model.

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

- The Intrumental Variable (IV) approach was first proposed by (Anderson and Hsiao 1981).
- They propose a 4-step estimation given 2 choices of instruments.
  1. First step: first difference transformation
  2. Second step: given choice of instruments apply IV method to estimate γ and β.
  3. Third step: estimate ρ_t. (can be ignored if no time-fixed effects)
  4. Fourth step: estimate of variance σ_α^2 and σ_ε^2.

Topic 4: Anderson and Hsiao (1982) first step

- W.l.o.g., we assume $\rho_t = 0$.
- First step is to take the first difference of the model, we obtain for $t = 2, \ldots, T$

$$(y_{it} - y_{i,t-1}) = \gamma(y_{i,t-1} - y_{i,t-2}) + \beta^\top(x_{it} - x_{it-1}) + \varepsilon_{it} - \varepsilon_{it-1}.$$

- The first difference transformation leads to "lost" one observation.
- But, it allows eliminating the individual effects (as the Within-transformation).

Topic 4: Anderson and Hsiao (1982) second step

- Second step: IV estimation
- Given a valid instruments z_{it}, where it satisfies

$$\begin{aligned} E\left(z_{it}(\varepsilon_{it} - \varepsilon_{it-1})\right) &= 0 \text{ Exogeneity property} \\ E\left(z_{it}(y_{it-1} - y_{it-2})\right) &\neq 0 \text{ Relevance property} \end{aligned}$$

- Then, we can apply the IV approach (if just identified) or the 2SLS approach (if over-identified) to estimate the model.
- Anderson and Hsiao (1982) propose two valid instruments:
  1. $z_{it} = y_{it-2}$.
  2. $z_{it} = y_{it-2} - y_{it-3}$.

Reference

Anderson, T. W. and Cheng Hsiao (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association* **76**(375), 598–606.