MGRECON 548Q: Empirical Economic Analysis

Fuqua School of Business, Duke University
Fall 2020

Help Section Week#5[1]

Simultaneity Bias:

- Consider regression:

$$Y_1 = \alpha_1 Y_2 + \beta_1 Z_1 + u_1 \tag{1}$$
$$Y_2 = \alpha_2 Y_1 + \beta_2 Z_2 + u_2 \tag{2}$$

We cannot elicit a causal effect of Y_1 upon Y_2 by running OLS on eq(2):

(i) There is a feedback: $Y_1 \to Y_2 \to Y_1 \to \cdots$
(ii) The error term is not exogenous

- Consider the reduce form:

$$Y_1 = \pi_{11} Z_1 + \pi_{12} Z_2 + v_1 \tag{3}$$
$$Y_2 = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2 \tag{4}$$

where $v_1 = \frac{\alpha_1 u_2 + u_1}{1-\alpha_1\alpha_2}$ and $v_2 = \frac{\alpha_2 u_1 + u_2}{1-\alpha_2\alpha_1}$.
We still cannot apply OLS to the reduced form equations and extract causal effects:

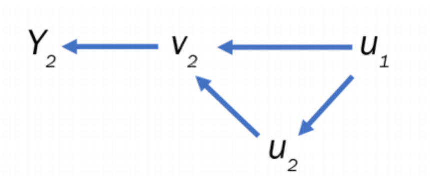
(i) The error term is exogenous now
(ii) But the reduced form parameters are nonlinear functions of the structural parameters

- Consider structural equation (1) and reduced form equation (4):

$$Y_1 = \alpha_1 Y_2 + \beta_1 Z_1 + u_1$$
$$Y_2 = \pi_{21} Z_1 + \pi_{22} Z_2 + v_2$$

Think about the exogeneity assumption again: $Cov(Y_2, u_1) = 0$ and $Cov(Z_1, u_1) = 0$. The first may be violated if:

(i) u_1 is directly related to v_2, or
(ii) u_1 is related to v_2 through u_2



(case i) If we assume u_1 is not related to u_2:

$$Cov(Y_2, u_1) = Cov(v_2, u_1) = \frac{\alpha_2}{(1-\alpha_2\alpha_2)} E(u_1^2)$$

By this equation, even if we can't obtain the exact magnitude in practice, we can interpret the bias sign: if $\alpha_2 > 0$ and $\alpha_1\alpha_2 < 0$, then the bias is positive.

[1] Please watch for typos and errors

(case ii)If we think Y_2 is related to u_1 indirectly via u_2:

this relationship would imply that the structural errors are related to one another. A typical example might be an omitted variable that is common to both structural equations

- IV Solution

  Consider the structural equations:

$$Y_1 = \beta_{10} + \alpha_1 Y_2 + \beta_{11} Z_1 + u_1 \tag{5}$$
$$Y_2 = \beta_{20} + \alpha_2 Y_1 + u_2 \tag{6}$$

  We can use Z_1 as an instrument variable: by changing the value of Z_1, we can have Y_1 changed without affecting Y_2 directly(see eq(5)). By analysis the change of Y_2, we can estimate the marginal effect of Y_1, that is α_2 (see eq(6)).

  Identification (Lec9, p. 27):

  (i) Order Condition ≈ Instrument Exogeneity

  (ii) Rank Condition ≈ Instrument Relevance.

- IV Estimator

  To implement IV (Z_1) on the equation (6) we need an instrument that is:

  (i) Instrument Relevance: Correlated with Y_1, i.e. $\beta_{11} \neq 0$

  (ii) Instrument Exogeneity: Not correlated with u_2, i.e. $Cov(Z_1, u_2) \neq 0$

  (iii) Instrument Exogeneity: Does not appear in equation (6)

$$\begin{aligned}
\hat{\alpha_2}^{IV} &= \frac{Cov(Z_1, Y_2)}{Cov(Z_1, Y_1)} \\
&= \frac{Cov(Z_1, \beta_{20} + \alpha_2 Y_1 + u_2)}{Cov(Z_1, Y_1)} \\
&= \alpha_2 (\text{unbiased})
\end{aligned}$$

  We can overcome the bias in our 2 estimate via IV by using Z_1 as an instrument for Y1 in equation(6), that is equation (6) is identified. The parameter α_1 cannot be estimated without bias, equation (5) is not identified.

Data Types:

- Cross section data, panel data and pooled cross section data (Lec10, p. 9)
  - (i) Cross section data: does not have the time dimension
  - (ii) Panel data: add time dimension to individual in cross section data. We have balance and unbalanced panel data
  - (iii) Pooled cross section data: also has time dimension but every year(t) we are observing different individuals

Pooled Cross Sections:

- Trade-off of using pooled cross section data:
  - (i) Pooled data has a larger sample size
  - (ii) But pooling might overlook time varying marginal effects, assumes cross sectional stability of the estimates, and assumes the errors across cross sectional units are uncorrelated over time
- Time Varying Marginal Effect

$$Y_i = \beta_0 + \beta_2 D_i^{time} + \beta_1 X_i + \beta_3 D_i^{time} X_i + e_i$$

By including the dummy variable and interaction term, we are able to estimate the time varying marginal effect:

  - (i) β_3 provides a time varying marginal effect of X upon Y
  - (ii) This model can be estimated by standard OLS
  - (iii) We can easily extend to multiple periods, but use caution. Too many dummy variables will use of degrees of freedom

- Difference-in-Difference (special case of time varying marginal effect)

Simple 2×2 table with 2 period (before and after) and 2 states (control and treatment):

	Before	After	After-Before
Control	$Y_{C,t=1}$	$Y_{C,t=2}$	$Y_{C,t=2}$ -$Y_{C,t=1}$
Treatment	$Y_{T,t=1}$	$Y_{T,t=2}$	$Y_{T,t=2}$ -$Y_{T,t=1}$
Treatment-Control	$Y_{T,t=1}$-$Y_{C,t=1}$	$Y_{T,t=2}$-$Y_{C,t=2}$	($Y_{T,t=2}$-$Y_{T,t=1}$)-($Y_{C,t=2}$-$Y_{C,t=1}$)

Example:

$$Y = \beta_0 + \delta_0 D^{After} + \beta_1 D^{Treat} + \delta_1 D^{After} \times D^{Treat} + e$$

	Before	After	After-Before
Control	β_0	$\beta_0 + \delta_0$	δ_0
Treatment	$\beta_0 + \beta_1$	$\beta_0 + \delta_0 + \beta_1 + \delta_1$	$\delta_0 + \delta_1$
Treatment-Control	β_1	$\beta_1 + \delta_1$	δ_1

About DiD:

  - (i) Cannot permit inference

(ii) Having only two periods doesn't permit us to assess the common trends assumption

(iii) These methods can be adapted to panel data settings, panel data model permits us to examine causality beyond policy experiments

Panel Data:

- Panel as A Solution to OVB

  If we observe the same person over time, many of their relevant characteristics (gender, etc..) will remain constant, which helps us control for those unobserved factors

$$
\begin{aligned}
Y_{it} &= \beta_0 + \delta_0 D_{it}^{After} + \beta_1 X_{it} + e_{it} \\
e_{it} &= a_i + u_{it}
\end{aligned}
$$

  where

  e_{it}: Composite error

  a_i: Entity-specific fixed effect; unobserved individual heterogeneity

  u_{it}: Idiosyncratic error that may vary through time

  If we treat the model like a pooled sample and run via OLS:

  (i) We would require that e_{it} is uncorrelated with X_{it}, i.e. exogeneity assumption. However, it is often the case that a_i will be correlated to X_{it}, thereby violating exogeneity

  (ii) This generates a heterogeneity bias

- First Difference Estimator to Avoid Heterogeneity bias

$$
\begin{aligned}
Y_{i,t=2} &= \beta_0 + \delta_0 * 1 + \beta_1 X_{i,t=2} + a_i + u_{i,t=2} \\
Y_{i,t=1} &= \beta_0 + \delta_0 * 0 + \beta_1 X_{i,t=1} + a_i + u_{i,t=1}
\end{aligned}
$$

  If we subtract these two equations we get

$$
\begin{aligned}
Y_{i,t=2} - Y_{i,t=1} &= \delta_0 + \beta_1 (X_{i,t=2} - X_{i,t=1}) + (u_{i,t=2} - u_{i,t=1}) \qquad (7) \\
\Delta Y_i &= \delta_0 + \beta_1 \Delta X_i + \Delta u_i \qquad (8)
\end{aligned}
$$

  (i) The fixed effect (a_i) does not appear in the difference equation

  (ii) Assume assumptions of cross sectional regression hold $\implies$ $\mathrm{Cov}(\Delta u_i, \Delta X_i) = 0$, we can estimate the first difference equation via OLS

  (iii) We still allow $\mathrm{Cov}(X_{it}, a_i) \neq 0$

  (iv) Recall $se(\hat{\beta}) = \frac{\hat{\sigma_e}}{SST_{\Delta X}(1-R^2_{\Delta X})}$, we need a large $SST_{\Delta X} = \sum_i^N (\Delta X_i - \overline{\Delta X_i})^2$ to decrease type II error $\implies$ we must have sufficient cross sectional variation. This can be solved by using a larger cross section and/or time changes over longer periods, but neither may be possible (major cost)

  (v) We can extend the 2-period data to multiple-period data and do first difference. We can estimate via pooled OLS, but we may have serial correlation problem (think about the assumptions that pooled OLS make)(Lec10, p. 43)

(vi) We are able to use FD to estimate the time varying marginal effect. Assume X_i is a time invariant regressor(think about 2-period):

$$Y_{it} = \beta_0 + \delta_0 D_{it}^{After} + \beta_1 X_i + \gamma D_{it}^{After} X_i + a_i + u_{it}$$

$$\Delta Y_i = \delta + \gamma X_i + \Delta u_i$$

we cannot estimate marginal effect of X on Y in a given period (β_1), but we can estimate how that marginal effect changes over time (γ)

Fixed Effect:

- Fixed Effect

$$Y_{it} = \beta_0 + \beta_1 X_{it} + a_i + u_{it}, t = 1, 2, \ldots T$$

To eliminate the time invariant effects a_i, we can:

(i) Do difference (FD)

(ii) De-meaning over time (within transformation)

$$\begin{aligned}
\bar{Y}_i &= \beta_1 \bar{X}_i + a_i + \bar{u}_{it} \\
\tilde{Y}_{it} &= Y_{it} - \bar{Y}_i \\
&= \beta_1 (X_{it} - \bar{X}_{it}) + (u_{it} - \bar{u}_{it}) \\
&= \beta_1 \tilde{X}_{it} + \tilde{u}_{it}
\end{aligned}$$

where $\bar{Y}_i = \frac{1}{T} \sum_{t=1}^{T} Y_{it}$, $\bar{X}_i = \frac{1}{T} \sum_{t=1}^{T} X_{it}$, and $\bar{u}_i = \frac{1}{T} \sum_{t=1}^{T} u_{it}$

(iii) Fixed Effect as dummy variable regression

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \sum_{i=1}^{N} \alpha_i D_i + u_{it}, t = 1, 2, \ldots T$$

This approach produces the same estimates as the fixed effect estimator, but we may reduces degrees of freedom significantly since the new regression has N more parameters need to be estimated

For this de-meaning regression:

(i) We can estimate via pooled OLS

(ii) The estimator $\hat{\beta}_1$ is referred to as a Fixed Effects estimator since it leverages the fact that a_i is fixed over time

(iii) This time de-meaning is also referred to as a Within Transformation since it makes a transformation within each cross sectional observation

(iv) Just as with the First Difference Estimator, the Fixed Effect estimator permits $\mathrm{Cov(X_{it}, a_i)} \neq 0$

(v) And just as with First Difference, the (within) transformation prevents us from estimating the marginal impact of these time invariant regressors (i.e. X_i)

(vi) $R^2 = 1 - \frac{\sum_{it} \hat{\tilde{u}}_{it}^2}{\sum_{it} (\tilde{Y}_{it} - \bar{\tilde{Y}}_{it})^2}$

(vii) $\bar{y}_i - \hat{\beta}_1 \bar{X}_i$ is not a good estimator of the time invariant unobserved effect(a_i) a for person i

- Fixed Effect vs First Difference (Lec10, p. 53, 54)

Random Effect

$$Y_{it} = \beta_0 + \beta_1 X_{it,1} + \beta_2 X_{it,2} + \cdots + \beta_K X_{it,K} + a_i + u_{it}, t = 1, 2, \ldots T$$

We use FD or FE to eliminate a_i because we were concerned it was correlated with X_{it}. If a_i is NOT correlated with X_{it} eliminating it might impact our standard errors.

- In a Random Effects model we assume

$$Cov(a_i, X_{it,k}) = 0, t = 1, 2, \ldots, T; k = 1, 2, \ldots, K$$

  Rewrite model:

$$Y_{it} = \beta_0 + \beta_1 X_{it,1} + \beta_2 X_{it,2} + \cdots + \beta_K X_{it,K} + v_{it}, t = 1, 2, \ldots T$$

  where $v_{it} = a_i + u_{it}$

  (i) Under the assumption, we can estimate this via typical cross sectional or pooled OLS methods and generate unbiased estimates

  (ii) However, if this (serial) correlation among the errors exists, it can impact our standard errors (heterogeneity assumption is violated)

$$\mathrm{corr}(\mathrm{v_{it}}, \mathrm{v_{is}}) = \frac{\sigma_\mathrm{a}^2}{\sigma_\mathrm{a}^2 + \sigma_\mathrm{u}^2}, \mathrm{t} \neq \mathrm{s}$$

- Fixing the Standard Errors(Lec10, p. 62)

  Apply a quasi-time demeaning

$$Y_{it} - \lambda \bar{Y}_i = \beta_0(1 - \lambda) + \beta_1(X_{it,1} - \lambda \bar{X}_{i,1}) + \cdots + \beta_K(X_{it,K} - \lambda \bar{X}_{i,K}) + (v_{it} - \lambda \bar{v}_i)$$

  where $\lambda = 1 - [\frac{\sigma_u^2}{(\sigma_u^2 + T\sigma_a^2)}]^{\frac{1}{2}}$