

面板数据

司继春

2024年12月



不可观测异质性

- 为了识别以上模型，一个非常常用的假设是假设 α_{it} 可以被分解为可加的两部分：

$$\alpha_{it} = \alpha_i + \lambda_t$$

得到

$$y_{it} = \alpha_i + \lambda_t + x'_{it}\beta + u_{it}$$

其中 α_i 为不随时间变化 (time-invariant) 的个体效应 (individual effects) 或者个体异质性 (individual heterogeneity) ，而 λ_t 则代表了不同时间截距项的不同，也就是时间异质性。

时间固定效应

- 对于 λ_t ，一种处理方法是直接设定 t 的函数形式，比如线性趋势：

$$\lambda_t = \lambda t$$

或者二次趋势：

$$\lambda_t = \lambda_1 t + \lambda_2 t^2$$

等。

- 然而这些设定都有很强的函数形式假定：对于任意的非线性的、偶发的冲击都无法很好的建模。

时间固定效应

- 由于我们在这里假设短面板数据，面临的是 T 固定但是 $N \rightarrow \infty$ 的情况，因而 λ_t 通常比较容易处理，可以直接通过引入时间固定效应（time fixed effects），即当前观测属于哪一期的虚拟变量来解决。
- 在引入了时间固定效应后，所有的不随个体 i 而变化的变量（比如国家的GDP、利率等）都与 λ_t 共线，从而不能再加入到回归方程中。
- 不失一般性，我们通常将时间固定效应 λ_t 看做是 x_{it} 中的一部分，因而我们可以将模型写为

$$y_{it} = \alpha_i + x'_{it}\beta + u_{it}$$

其中现在的 x_{it} 中包含着时间固定效应 λ_t 。

面板的向量形式

- 接下来，在本章中，我们都以

$$y_{it} = \alpha_i + x'_{it}\beta + u_{it}$$

为基础进行讨论。

- 为了方便起见，我们通常记

$$X_i = \begin{bmatrix} x'_{i1} \\ \vdots \\ x'_{iT} \end{bmatrix}_{T \times K}, Y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix}_{T \times 1}, U_i = \begin{bmatrix} u_{i1} \\ \vdots \\ u_{iT} \end{bmatrix}_{T \times 1}$$

从而上式可以写为

$$Y_i = \alpha_i + X_i\beta + U_i$$

由于 T 的固定的，从而以上 X_i, Y_i, U_i 的维数都是固定的。

两种不同的模型

- 虽然对于时间固定效应 λ_t 的处理相对简单，然而，由于 $N \rightarrow \infty$ ，对 α_i 的处理相对困难。
- 注意 α_i 也是一个随机变量，根据对 α_i 的假设不同，通常来讲有两类不同的模型可以使用：
 - 随机效应 (random effects) 模型
 - 固定效应 (fixed effects) 模型
- 两者的关键区别在于 α_i 与解释变量 x_{it} 之间相关性假设。在介绍两种方法之前，我们首先介绍使用传统的最小二乘法及其所需要的假设。

外生性假设

- 注意以上仅仅假设了当期的 u_{it} 和当期的 x_{it} 不相关
- 但是由于不可观测的个体异质性 α_i 出现在了每一期的方程中，因而需要 α_i 和所有期的 x_{it} 都不相关，可以简写为 α_i 和 X_i 不相关。
- 由于本章中主要考虑个体异质性 α_i 的问题，因而我们暂且忽略 u_{it} 与 x_{it} 的相关性。
- 当然在实证中，如果 u_{it} 与 x_{it} 相关，那么这种外生性的违背是需要重点考虑的。

POLS

- 在以上假设条件下，我们可以将 v_{it} 视为新的误差项，此时我们可以简单的使用OLS对其进行估计。
- 若以上假设成立：

$$\mathbb{E}(x_{it}v_{it}) = \mathbb{E}[\mathbb{E}(x_{it}v_{it} | x_{it})] = \mathbb{E}[x_{it}\mathbb{E}(\alpha_i + u_{it} | x_{it})] = 0$$

从而 $\mathbb{E}(v_{it}x_{it}) = 0$ ，由此得到矩条件：

$$\mathbb{E}[x_{it}(y_{it} - x'_{it}\beta)] = 0$$

得到

$$\beta = [\mathbb{E}(x_{it}x'_{it})]^{-1} \mathbb{E}(x_{it}y_{it})$$

POLS

如果记

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}_{NT \times K}, Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix}_{NT \times 1}$$

那么POLS估计量可以写为：

$$\hat{\beta}_{POLS} = (X'X)^{-1} X'Y$$

从而获得POLS估计量的一个重要前提是 $(X'X)^{-1}$ 可逆。据此，我们可以引入如下识别条件：

识别条件

假设 $\mathbb{E}(X_i'X_i)$ 可逆。

POLS统计性质

- 如果记 $V_i = \alpha_i + U_i$ ，那么回归式可以写为 $Y_i = X_i\beta + V_i$
- 将其带入到混合最小二乘的估计式中，有：

$$\begin{aligned} \hat{\beta}_{POLS} &= \left(\sum_{i=1}^N X_i' X_i \right)^{-1} \left[\sum_{i=1}^N X_i' (X_i\beta + V_i) \right] \\ &= \beta + \left(\sum_{i=1}^N X_i' X_i \right)^{-1} \left(\sum_{i=1}^N X_i' V_i \right) \\ &= \beta + \left(\sum_{i=1}^N X_i' X_i \right)^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} v_{it}) \right] \end{aligned}$$

POLS统计性质

- 从而在外生性假设以及一定可积性条件下，根据大数定律，有：

$$\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^T (x_{it}v_{it}) \xrightarrow{P} \mathbb{E} \left(\sum_{t=1}^T (x_{it}v_{it}) \right) = \sum_{t=1}^T \mathbb{E} (x_{it}v_{it}) = 0$$

进而易得，POLS估计量为一致估计量。

- 进一步，由于

$$\begin{cases} \frac{1}{N} \sum_{i=1}^N X_i' X_i \xrightarrow{P} \mathbb{E} (X_i' X_i) \\ \sqrt{N} \frac{1}{N} \sum_{i=1}^N X_i' V_i \overset{a}{\sim} N(0, \mathbb{V} (X_i' V_i)) \end{cases}$$

而其中 $\mathbb{V} (X_i' V_i) = \mathbb{E} (X_i' V_i V_i' X_i)$ ，从而

$$\sqrt{N} (\hat{\beta}_{POLS} - \beta) \overset{a}{\sim} N \left(0, [\mathbb{E} (X_i' X_i)]^{-1} \mathbb{E} (X_i' V_i V_i' X_i) [\mathbb{E} (X_i' X_i)]^{-1} \right)$$

POLS的标准误

- 为了估计混合最小二乘估计量的标准误，类比异方差稳健标准误的计算方法。
- 记 $\hat{V}_i = Y_i - X_i \hat{\beta}_{POLS}$ 为个体 i 的残差向量，那么使用插入法：

$$\mathbb{V}(\widehat{\beta}_{POLS}) = (X'X)^{-1} \left[\sum_{i=1}^N (X_i' \hat{V}_i \hat{V}_i' X_i) \right] (X'X)^{-1}$$

以上就得到了POLS估计系数的标准误的一致估计。注意其中：

$$\begin{aligned} \sum_{i=1}^N (X_i' \hat{V}_i \hat{V}_i' X_i) &= \sum_{i=1}^N \left(\begin{bmatrix} x_{i1} & \cdots & x_{iT} \end{bmatrix} \begin{bmatrix} \hat{v}_{i1} \\ \vdots \\ \hat{v}_{iT} \end{bmatrix} \begin{bmatrix} \hat{v}_{i1} & \cdots & \hat{v}_{iT} \end{bmatrix} \begin{bmatrix} x'_{i1} \\ \vdots \\ x'_{iT} \end{bmatrix} \right) \\ &= \sum_{i=1}^N \sum_{t=1}^T \sum_{s=1}^T (\hat{v}_{it} \hat{v}_{is} x_{it} x'_{is}) \\ &= \sum_{i=1}^N \sum_{t=1}^T (\hat{v}_{it}^2 x_{it} x'_{it}) + \sum_{i=1}^N \sum_{t=1}^T \sum_{s=1, s \neq t}^T (\hat{v}_{it} \hat{v}_{is} x_{it} x'_{is}) \end{aligned}$$

- 聚类标准误：至少聚类到个体 i 层面

POLS实例

俄罗斯大选

在之前的例子中，我们展示了Enikolopov, Petrova和Zhuravskaya（2011）的截面数据分析结果，除此之外文章中还使用了如下的面板数据设定：

$$vote_{s,t} = \beta_0 + \beta_1 NTV_{s,t} + x'_{st}\beta + \lambda_t + \alpha_i + \epsilon_{s,t}$$

其中 $t = 1995, 1999$ ，即一个两年的面板数据， s 为选区， x_{st} 为其他控制变量。如果使用POLS对以上模型进行回归，那么可以使用如下代码：

POLS实例

俄罗斯大选

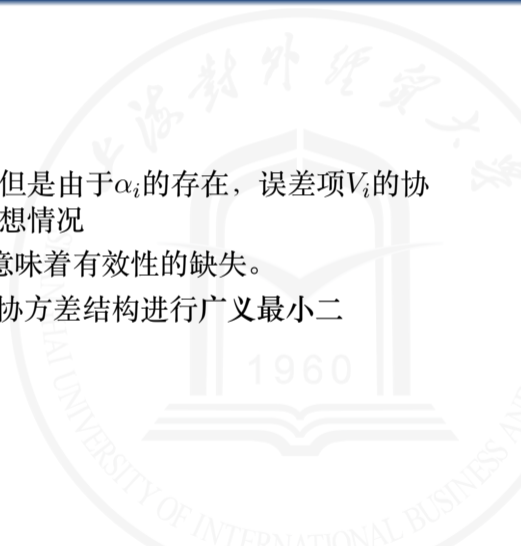
```

1 use datasets/NTV_Aggregate_Data_reshaped.dta
2 // 年1995并没有成立NTV
3 gen Watch_probit_p=0
4 replace Watch_probit_p=Watch_probit if year==1999
5 // 回归POLS
6 reg Votes_SPS_ Watch_probit_p i.region i.year if year!=2003,
   cluster(tik_id)
7 // reghdfe
8 reghefe Votes_SPS_ Watch_probit_p if year!=2003, a(i.region i.year
   ) cluster(tik_id)

```

随机效应模型

- 虽然我们可以使用混合最小二乘得到估计，但是由于 α_i 的存在，误差项 V_i 的协方差矩阵不会是一个同方差、无自相关的理想情况
- 而这种理想情况是BLUE所需要的，这也就意味着有效性的缺失。
- 为了达到有效性，我们可以使用误差项 V_i 的协方差结构进行广义最小二乘（generalized least squares, GLS）。



随机效应的假设

协方差结构和外生性假设

假设：

1. α_i 具有同方差，即 $\mathbb{V}(\alpha_i | X_i) = \sigma_\alpha^2$
2. u_{it} 具有同方差且无自相关，即假设 $\mathbb{V}(U_i | X_i) = \sigma_u^2 \cdot I_T$ 。
3. $\mathbb{E}(\alpha_i u_{it} | X_i) = 0$

随机效应模型的估计

- 如果我们知道 Ω ，就可以得到最有效的估计量。
- 如果我们在方程的两边同时乘以 $\Omega^{-1/2}$ ，得到：

$$\Omega^{-1/2}Y_i = \Omega^{-1/2}X_i\beta + \Omega^{-1/2}V_i$$

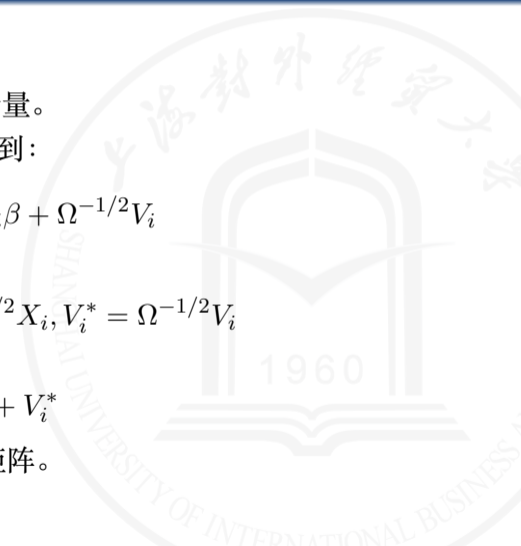
如果记

$$Y_i^* = \Omega^{-1/2}Y_i, X_i^* = \Omega^{-1/2}X_i, V_i^* = \Omega^{-1/2}V_i$$

那么以上方程可以写为：

$$Y_i^* = X_i^*\beta + V_i^*$$

其中 Y_i^* 为 $T \times 1$ 的向量，而 X_i^* 为 $T \times K$ 的矩阵。



随机效应模型的估计

- 此外，有：

$$\mathbb{V}(V_i^*) = \mathbb{V}(\Omega^{-1/2}V_i) = \Omega^{-1/2}\Omega\Omega^{-1/2} = I$$

因而我们可以直接使用 Y_i^* 对 X_i^* 做OLS，得到：

$$\begin{aligned} \hat{\beta}_{GLS} &= \left(\sum_{i=1}^N X_i^{*'} X_i^* \right)^{-1} \left(\sum_{i=1}^N X_i^{*'} Y_i^* \right) \\ &= \left(\sum_{i=1}^N X_i' \Omega^{-1} X_i \right)^{-1} \left(\sum_{i=1}^N X_i' \Omega^{-1} Y_i \right) \\ &= [X' (I_N \otimes \Omega^{-1}) X]^{-1} [X' (I_N \otimes \Omega^{-1}) Y] \end{aligned}$$

随机效应的假设

- 值得注意的是，虽然在当期外生性的条件下，混合最小二乘是一致的，但是这并不代表随机效应估计量一定一致。
- 注意到

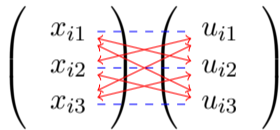
$$\begin{aligned}\hat{\beta}_{GLS} &= \left(\sum_{i=1}^N X_i' \Omega^{-1} X_i \right)^{-1} \left[\sum_{i=1}^N X_i' \Omega^{-1} (X_i \beta + V_i) \right] \\ &= \beta + \left(\sum_{i=1}^N X_i' \Omega^{-1} X_i \right)^{-1} \left(\sum_{i=1}^N X_i' \Omega^{-1} V_i \right)\end{aligned}$$

其中

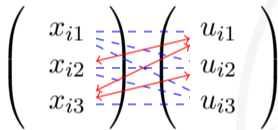
$$\frac{1}{N} \sum_{i=1}^N X_i' \Omega^{-1} V_i \xrightarrow{p} \mathbb{E}(X_i' \Omega^{-1} V_i)$$

- 然而，如果 Ω 不是对角矩阵（从而 Ω^{-1} 也不是对角阵），在 $X_i' \Omega^{-1} V_i$ 中就必然存在着 $x_{is} u_{it}, s \neq t$
- 而当期外生性假设仅仅要求当期不相关，即 $\mathbb{E}(x_{it} u_{it}) = 0$ ，而没有对 $\mathbb{E}(x_{is} u_{it})$ 做任何假定
- 因而如果仅仅假设当期不相关，可能会存在着不同期的 x 和 u 之间的相关性，此时 $\mathbb{E}(X_i' \Omega^{-1} V_i) = 0$ 可能不成立。

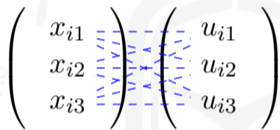
随机效应的假设



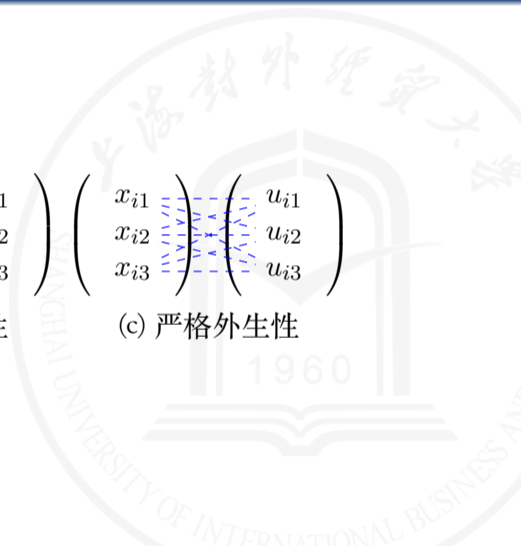
(a) 当期外生性



(b) 序贯外生性



(c) 严格外生性



序贯外生性

- 比当期外生性更加严格一点的是序贯外生性：

序贯外生性, sequential exogeneity

假设对于所有的 i, t , 有 $\mathbb{E}(u_{it} | x_{it}, x_{i,t-1}, \dots, x_{i1}) = 0$ 。

- 上图(b)展示了序贯外生性。在当期外生性的基础上, 序贯外生性不仅要求当期的 u_{it} 与 x_{it} 不相关, 而且还额外要求 u_{it} 和过去的所有 x_{it} ($x_{it}, x_{i,t-1}, \dots, x_{i1}$) 都不相关。
- 注意此时我们允许 u_{it} 和未来的 x ($x_{i,t+1}, \dots, x_T$) 相关。
- 此时, 看起来好像 x_{it} 是由滞后的误差项 ($u_{i,t-1}, \dots, u_1$) 所“决定”的, 我们称这些 x 为前定变量 (predetermined variables)。

严格外生性

- 以上使用了条件期望定义严格外生性，实际上在这里我们只需要一个稍微弱一点的假设：

严格外生性2

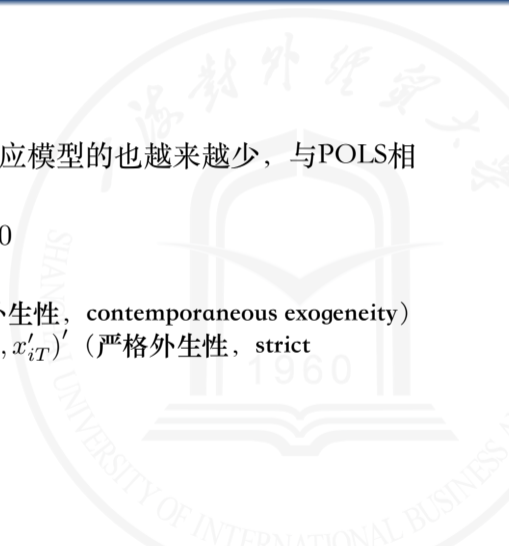
假设

$$\mathbb{E}(u_i \otimes X_i) = \mathbb{E} \begin{bmatrix} u_{i1} X_i \\ \vdots \\ u_{iT} X_i \end{bmatrix}_{T^2 \times K} = \mathbb{E} \begin{bmatrix} u_{i1} x'_{i1} \\ \vdots \\ u_{i1} x'_{iT} \\ \vdots \\ u_{iT} x'_{i1} \\ \vdots \\ u_{iT} x'_{iT} \end{bmatrix} = 0$$

POLS v.s. RE

经常会看到，即使面板数据，文献中使用随机效应模型的也越来越少，与POLS相比，其在假设上有一点细微差别：

- 对 α_i 的假设上：两者都需要假设 $\mathbb{E}(\alpha_i|x_i) = 0$
- 对 u_{it} 的假设上：
 - POLS只需要假设 $\mathbb{E}(u_{it}|x_{it}, \alpha_i) = 0$ （当期外生性，contemporaneous exogeneity）
 - RE需要假设 $\mathbb{E}(u_{it}|X_i, \alpha_i) = 0, X_i = (x'_{i1}, \dots, x'_{iT})'$ （严格外生性，strict exogeneity）



随机效应实例

俄罗斯大选

在俄罗斯大选的例子中，如果需要使用随机效应估计量，可以使用：

```
1 xtreg Votes_SPS_Watch_probit_p i.region i.year if year!=2003, re  
cluster(tik_id)
```

其中xtreg命令为面板数据回归命令，而re选项则指定使用随机效应模型进行估计，此外，标准误被聚类到个体层次。

- 在xtreg命令中，robust选项与cluster(panel_id)是等价的，由于该例该面板本身为选区层面面板，此时使用cluster(tik_id)与直接加入robust选项是等价的。
- 但是在其他命令如reg、reghdfe等并没有这一设定，仍然需要手动使用cluster()进行聚类。

一阶差分方法

- 为了处理可能的 α_i 与 X_i 之间的相关性，一个更加聪明的办法是想办法在回归方程中将 α_i 消掉，从而可以完全无需对 α_i 和 X_i 之间的关系进行建模。
- 将 α_i 消掉最简单的做法是直接对回归方程式

$$y_{it} = \alpha_i + x'_{it}\beta + u_{it}$$

进行一阶差分（first difference），从而得到

$$\Delta y_{it} = \Delta x'_{it}\beta + \Delta u_{it}$$

其中 $\Delta y_{it} = y_{it} - y_{i,t-1}$ 为一阶差分，其他类似。

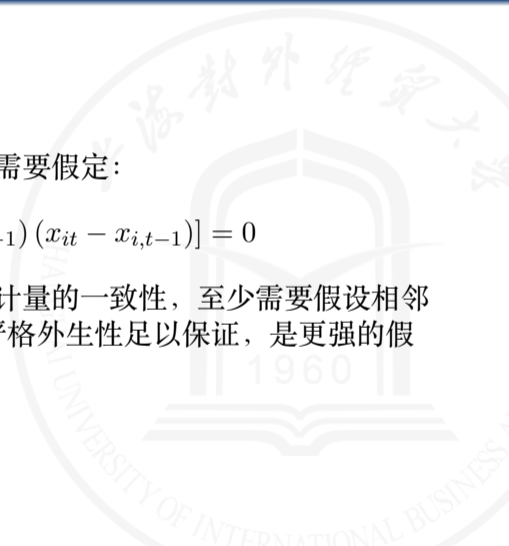
- 由于 α_i 是不随时间变化而变化的，所以经过一阶差分之后，自然消掉了 α_i ，从而也就无需考虑 α_i 与 x_{it} 之间的相关性。
- 当然，解释变量中不随时间变化的所有特征也被一阶差分所消掉，所以使用该方法时， x_{it} 中将不能包含不随时间而变化的解释变量。

一阶差分方法

- 为了使得以上方程可以通过OLS直接估计，需要假定：

$$\mathbb{E}(\Delta u_{it} \Delta x_{it}) = \mathbb{E}[(u_{it} - u_{i,t-1})(x_{it} - x_{i,t-1})] = 0$$

- 因而序贯外生性等假设并不足以保证OLS估计量的一致性，至少需要假设相邻两期的 u_{it} 与 x_{it} 之间不相关才可以，当然，严格外生性足以保证，是更强的假设。



一阶差分的应用

美国媒体与选举

Gentzkow, Shapiro and Sinkinson(2011)研究了美国各县的报纸数量对美国选举投票率的影响。由于考虑到数据中冲击具有较强的持续性影响，从而选择了一阶差分模型，其模型设定为：

$$y_{ct} = \beta \times n_{ct} + x'_{ct}\delta + \alpha_c + \gamma_{st} + \epsilon_{ct}$$

其中 y_{ct} 为 c 县 t 年的总统选举投票率， n_{ct} 为报纸数量， α_c 为县固定效应，而 γ_{ct} 为州 \times 年份固定效应。经过一阶差分后，得到：

$$\Delta y_{ct} = \beta \times \Delta n_{ct} + \Delta x'_{ct}\delta + \Delta \gamma_{ct} + \Delta \epsilon_{ct}$$

经过一阶差分后使用线性回归即可。 (linear_panel_fd_np.do)

LSDV

- 注意到回归方程式中将 α_i 视为个体固定效应：

$$y_{it} = \alpha_i + x'_{it}\beta + u_{it}$$

如果使用最小二乘回归，等价于：

- 使用将 x_{it} 对 N 个虚拟变量 $\alpha_i, i = 1, \dots, N$ 做回归，得到残差。根据之前的结论，该残差为 $\ddot{x}_{it} = x_{it} - \bar{x}_i$ 其中 \bar{x}_i 为第 i 个个体的所有时间段的 x 的平均值： $\bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}$ 以上过程即每个观测减去个体的平均，也成为去平均 (demean)；
- 使用 y_{it} 对 N 个虚拟变量 $\alpha_i, i = 1, \dots, N$ 做回归，得到残差。根据之前的结论，该残差为： $\ddot{y}_{it} = y_{it} - \bar{y}_i$ 即被解释变量去平均；
- 使用OLS进行估计：

$$\hat{\beta}_{LSDV} = \left(\sum_{i=1}^N \sum_{t=1}^T \ddot{x}_{it} \ddot{x}'_{it} \right)^{-1} \left(\sum_{i=1}^N \sum_{t=1}^T \ddot{x}_{it} \ddot{y}_{it} \right)$$

固定效应

- 实际上，以上估计量也可以由如下步骤得到：

1. 对于回归方程式 $y_{it} = \alpha_i + x'_{it}\beta + u_{it}$ 按照时间进行平均，得到：

$$\bar{y}_i = \alpha_i + \bar{x}'_i\beta + \bar{u}_i$$

2. 相减，得到：

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)' \beta + (u_{it} - \bar{u}_i)$$

从而减去了个体异质性 α_i

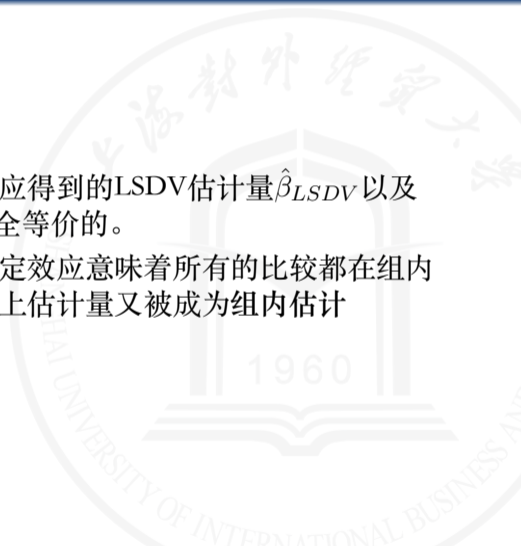
3. 使用 $\check{y}_{it} = y_{it} - \bar{y}_i$ 对 $\check{x}_{it} = x_{it} - \bar{x}_i$ 做回归，得到估计量：

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^N \sum_{t=1}^T \check{x}_{it} \check{x}'_{it} \right)^{-1} \left(\sum_{i=1}^N \sum_{t=1}^T \check{x}_{it} \check{y}_{it} \right)$$

- 传统上，在定义以上估计量时，通常将 α_i 视为固定的、不随机的参数，因而也被成为固定效应（fixed effects）模型。

固定效应

- 可见，以上两种方法：通过加入个体固定效应得到的LSDV估计量 $\hat{\beta}_{LSDV}$ 以及通过去平均得到的固定效应估计量 $\hat{\beta}_{FE}$ 是完全等价的。
- 实际上，根据我们之前的结论，加入个体固定效应意味着所有的比较都在组内（个体内部的不同时间之间）进行，因而以上估计量又被成为组内估计量（within-group estimator）。



固定效应统计性质

- 记 $M_0^T = I - \frac{1}{T} \iota_T \iota_T'$, 那么:

$$\begin{aligned} \hat{\beta}_{FE} &= \left(\sum_{i=1}^N \ddot{X}'_i \ddot{X}_i \right)^{-1} \left(\sum_{i=1}^N \sum_{t=1}^T \ddot{X}'_i \ddot{Y}_i \right) \\ &= \left(\sum_{i=1}^N X'_i M_0^T X_i \right)^{-1} \left(\sum_{i=1}^N X'_i M_0^T Y_i \right) \\ &= \beta + \left(\sum_{i=1}^N X'_i M_0^T X_i \right)^{-1} \left(\sum_{i=1}^N X'_i M_0^T U_i \right) \end{aligned}$$

- 由于 M_0^T 不是对角阵, 因而此时, 我们需要严格外生性此时, 有:

$$\frac{1}{N} \sum_{i=1}^N X'_i M_0^T U_i \xrightarrow{p} \mathbb{E} (X'_i M_0^T U_i) = 0$$

从而一致性满足

固定效应的应用

俄国选举与NTV

在上面俄国选举的例子中，作者还额外使用1995年的选举结果进行了比对，其模型设定如下：

$$vote_{st} = \beta_0 + \beta_1 NTV_{st} + \delta_t + \alpha_s + \epsilon_{st}$$

由于以上只有两期，实际上以上做法与一阶差分是等价的。

(linear_panel_fe_ntv.do)

可用的variation

- 组内估计量：估计

$$\dot{y}_{it} = \dot{x}'_{it}\beta + \dot{u}_{it}$$

得到 $\hat{\beta}^{WG}$

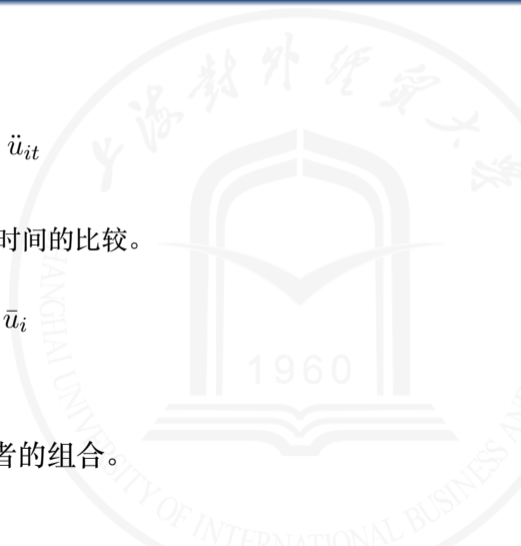
- 等价于LSDV，因而只包含了同一个个体的不同时间的比较。

- 组间估计量：估计

$$\bar{y}_i = \bar{x}'_i\beta + \bar{u}_i$$

得到 $\hat{\beta}^{BG}$

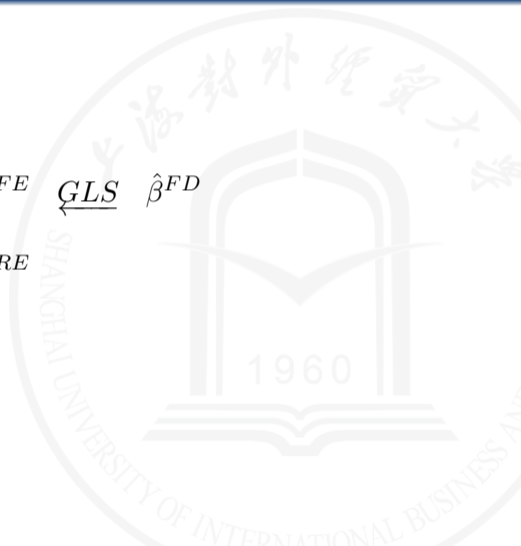
- 只比较了不同个体的均值。
- 固定效应为组内估计量，而POLS、RE为两者的组合。



不同估计量之间的关系

$$\begin{array}{c}
 \hat{\beta}^{WG} \\
 \searrow \\
 \hat{\beta}^{POLS} \\
 \nearrow \\
 \hat{\beta}^{BG}
 \end{array}
 = \hat{\beta}^{LSDV} = \hat{\beta}^{FE} \xleftarrow{GLS} \hat{\beta}^{FD}$$

$$\hat{\beta}^{POLS} \xrightarrow{GLS} \hat{\beta}^{RE}$$



固定效应v.s.随机效应、混合最小二乘

固定效应估计量与随机估计量的优缺点：

- 从假设强弱来看：
 - 固定效应所需假设更弱，更稳健
 - 随机效应、POLS假设更强：需假设 $E(\alpha_i|x_i) = 0$
- 另一个重要区别：
 - 固定效应不能估计不随时间变化的变量的系数
 - 随机效应、POLS可以



面板数据：标准误的问题

线性面板数据的标准误与OLS的标准误一样，同样存在Cluster的问题：

- 至少要Cluster到个体层面
 - xtreg命令加robust选项即可
 - 或者xtreg, cl(id)
- 当然，可以进行更高层次的Cluster
- 特别是使用reg、reghdfe等手动作固定效应时，一定不要忘了加cluster

固定效应与随机效应：检验

- 固定效应与随机效应的关键区别： $\mathbb{E}(\alpha_i|x_i) = 0$ 是否成立：

| | $\mathbb{E}(\alpha_i x_i) = 0$ | $\mathbb{E}(\alpha_i x_i) \neq 0$ |
|--------------------|--------------------------------|-----------------------------------|
| $\hat{\beta}^{RE}$ | 一致，最有效 | 不一致 |
| $\hat{\beta}^{FE}$ | 一致 | 一致 |

- 理论上：如果 $\mathbb{E}(\alpha_i|x_i) = 0$ 成立优先使用随机效应。
 - 实际上，如果该检验不显著，随机效应和固定效应的估计应该区别不大
- 检验：Hausman检验
- 实践中：固定效应最稳健，应优先考虑。
- 实践中：xtreg命令、reghdfe命令

