

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。

两种不同的模型

- 对于面板数据，如果忽略其面板结构，我们可以直接通过OLS进行估计。
- 然而，注意到由于 α_i 不可观测，从而模型可以写为

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

其中 v_{it} 包含所有不可观测的因素，包括误差项 u_{it} 以及个体异质性 α_i 。

- 针对以上的设定，我们首先可以忽略面板结构，将每个个体的每一期都当做是一个独立的个体，从而，我们将 $N \times T$ 个观测视为独立的个体。

外生性假设

根据之前OLS一致性的讨论，为了使得一致性成立，我们需要假设 $\mathbb{E}(v_{it}|x_{it}) = 0$ ，而为了保证 $\mathbb{E}(v_{it}|x_{it}) = 0$ 成立，我们通常将其分解为两个单独的假设：

当期外生性， contemporaneous exogeneity

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

以及：

个体异质性的外生

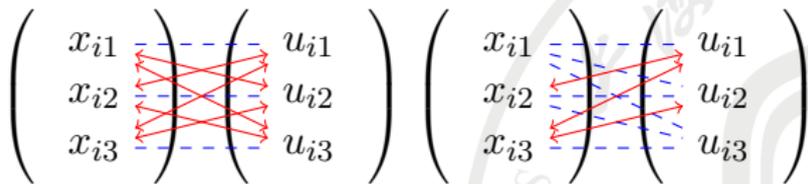
假设对于所有的 i ，有 $\mathbb{E}(\alpha_i|X_i) = 0$ 。

POLS统计性质

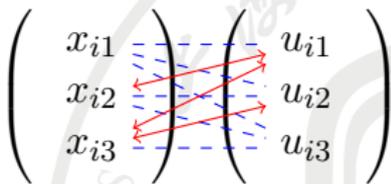
- 如果记 $V_i = \alpha_i + U_i$, 那么回归式可以写为 $Y_i = X_i\beta + V_i$
- 将其带入到混合最小二乘的估计式中, 有:

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

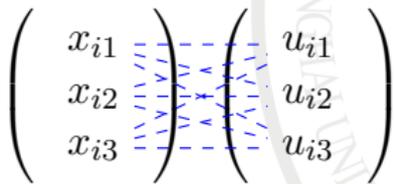
随机效应的假设



(a) 当期外生性



(b) 序贯外生性



(c) 严格外生性

严格外生性

- 然而以上序贯外生性还是无法保证 $\mathbb{E}(X_i' \Omega^{-1} V_i) = 0$ 成立，实际上为了保证该条件成立，我们需要更严格的外生性，即：

严格外生性， strict exogeneity

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

- 上图(c)展示了严格外生性，即需要假设所有期的 x 和所有期的 u 都不相关，这是一个最严格的外生性假设。
- 在严格外生性的条件下，我们可以得到 $\mathbb{E}(X_i' \Omega^{-1} V_i) = 0$ 成立，从而GLS估计量或者随机效应估计量 $\hat{\beta}_{GLS}$ 是 β 的一致估计。

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

- 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**）。

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

