


0. Intuition

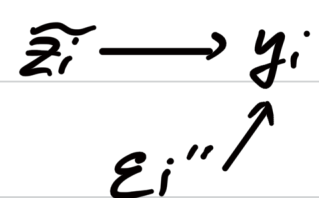
一、情境：解释变量 x or 工具变量 IV (本质一样都用 z_i)

1. 解释变量: $z(x) \rightarrow Y$ $y_i = \beta z_i + \varepsilon_i$, $z_i = f_i(g, w)$


✓ 期望处理效应: $\mu_i = E(z_i | w_i)$ ($G(\cdot)$ 分布已知)

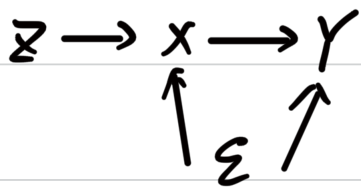
• (1) Control: $y_i = \beta z_i + \rho \mu_i + \varepsilon_i'$: 

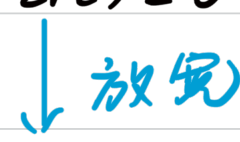
本质一样
(除 ε_i)

(2) OLS: $\tilde{z}_i = z_i - \mu_i$, $y_i = \beta \tilde{z}_i + \varepsilon_i''$: 

2. 工具变量: 更方便.

(1) 传统 IV Assumption: (1) relevance; (2) exclusion;



(3) exogeneity: $\text{Cov}(z, \varepsilon) = 0$
 放宽

(3) shock exogeneity: $z_i = f_i(g, w)$

$\text{Cov}(g, \varepsilon) = 0$, $G(\cdot)$ 已知
 $\text{Cov}(w, \varepsilon) \neq 0$

(2) 期望工具变量: $\mu_i = E[z_i | w_i]$ ($G(\cdot)$ known)

(3) $\tilde{z}_i = z_i - \mu_i$ 满足 Exogeneity.

3. Data Requirement.

放松了 Assumption 中的外生性要求, 找 z_i 更容易

提升了数据要求: 要求 $G(\cdot)$ 已知, 即反事实情境冲击已知

如: 铁路开通时间; 天气数据(天气); 网络等一定随机性