potential outcome $= \begin{cases} y_{1i}, & \text{if } X_i = 1 \\ y_{0i}, & \text{if } X_i = 0 \end{cases}$

y_{1i}, y_{0i} 是假想的 outcome，不管实际上他是否真的参与了。

observed outcome, $y_i = \begin{cases} y_{1i}, & \text{if } X_i = 1 \\ y_{0i}, & \text{if } X_i = 0 \end{cases}$

$$= y_{0i} + \underbrace{(y_{1i} - y_{0i})}_{\text{causal effect}} X_i$$

Naive Comparison:

$$E[y_i | X_i = 1] - E[y_i | X_i = 0] = \underbrace{E[y_{1i} | X_i = 1] - E[y_{0i} | X_i = 1]}_{\text{ATE on the treated}}$$

$$+ \underbrace{E[y_{0i} | X_i = 1] - E[y_{0i} | X_i = 0]}_{\text{selection bias}}$$

如果是看去医院的作用，$E[y_{0i} | X_i = 1] < E[y_{0i} | X_i = 0]$（很可能）

使 ATE 变小，也即医院对"去了医院的人"(treated) 的平均作用 (ATE) 被低估了。

如果实验是随机的，则

$$E[y_i|x_i=1]-E[y_i|x_i=0]$$
$$=E[y_{1i}|x_i=1]-E[y_{0i}|x_i=0]$$
$$=E[y_{1i}|x_i=1]-E[y_{0i}|x_i=1]$$
$$=E[y_{1i}-y_{0i}|x_i=1]$$
$$=E[y_{1i}-y_{0i}]$$

Regression:

$$y_i=\alpha+\beta x_i+\varepsilon_i$$

$$E(y_{0i})+(y_{1i}-y_{0i})x_i+y_{0i}-E(y_{0i})$$

假设对所有人都一样，也即常数

$$E[y_i|x_i=1]=\alpha+\beta+E[\varepsilon_i|x_i=1]$$

$$E[y_i|x_i=0]=\alpha+E[\varepsilon_i|x_i=0]$$

$$E[y_i|x_i=1]-E[y_i|x_i=0]$$
$$=\beta+E[\varepsilon_i|x_i=1]-E[\varepsilon_i|x_i=0]$$ (x_i与ε_i之间的相关性)

$\varepsilon_i=y_{0i}-E[y_{0i}]$

$$=\beta+E[y_{0i}|x_i=1]-E[y_{0i}|x_i=0]$$

selection bias.

例 x_i为服用药量，比如，$x_i = \{0, 1, 2\}$，而y_i为病情康复情况。

例 $x_i = \{0, 1\}$表示是否参加过某一就业培训项目(job training program)，y_i为未来的就业状态。

如果$x_i = \{0, 1\}$为虚拟变量，则方程的 OLS 估计量为

$$\hat{\beta}_{\text{OLS}} = \bar{y}_{\text{treat}} - \bar{y}_{\text{control}}$$

$\bar{y}_{\text{treat}}$为实验组的样本均值，$\bar{y}_{\text{control}}$为控制组的样本均值。

在回归方程中加入虚拟变量的效果就相当于给予实验组与控制组不同的截距项。而当$\{y_i\}$对常数项回归，系数估计值就是$\bar{y}$。因此，

$$b = (X'X)^{-1}X'y, \quad X = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{bmatrix} \begin{matrix} \} n_c \\ \} n_t \end{matrix}$$

$$X'X = \begin{bmatrix} 1 & \cdots & 1 \\ 0 & \cdots & 0 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} n & n_t \\ n_t & n_t \end{bmatrix}$$

$$(X'X)^{-1} = \frac{1}{n_t(n-n_t)} \begin{bmatrix} n_t & -n_t \\ -n_t & n \end{bmatrix}$$

$$X'y = \begin{bmatrix} \sum^{n} y_i \\ \sum^{n_t} y_i \end{bmatrix}$$

$$(X'X)^{-1}X'y = \frac{1}{n_t(n-n_t)} \begin{bmatrix} n_t \sum^{n} y_i - n_t \sum^{n_t} y_i \\ -n_t \sum^{n} y_i + n \sum^{n_t} y_i \end{bmatrix}$$

$$\frac{1}{n_t(n-n_t)} \left(-n_t \sum^{n} y_i + n \sum^{n_t} y_i\right)$$

$$= \frac{-1}{n-n_t} \left(\sum^{n-n_t} y_i + \sum^{n_t} y_i\right) + \frac{1}{n_t(n-n_t)} n \sum^{n_t} y_i$$

$$= \frac{-n_t + n}{n_t(n-n_t)} \sum^{n_t} y_i - \frac{1}{n-n_t} \sum^{n-n_t} y_i$$

$$= \bar{y}_t - \bar{y}_c$$