# 局部处理效应

4.4—4.5

荆艺玮

2019.10.24

# 工具变量与异质性潜在结果

- 目前为止，我们对工具变量的讨论仅局限在因果效应为常数的假设下。
- 在虚拟变量表征是否有服役经历的例子中，这个假设意味着对所有的$i$，都有$Y_{1i}-Y_{0i}=\rho$。
- 因果效应为常数不仅假设了因果效应具有同质性，还假设了在我们关心的因果模型中因果效应是线性的。

# 因果效应异质性的重要性

- 研究设计的两种不同类型的效度:
- 内部效度（internal validity):给定的研究设计是否成功揭示了总体中令人感兴趣的因果效应。
- 外部效度（external validity）：当情况发生改变时研究结论的预测能力。

# 局部平均处理效应

- 记$Y_i(\mathrm{z},\mathrm{d})$为个体$i$ 在处理结果为$D_i=\mathrm{d}$，工具变量取值为$\mathrm{Z}_i=\mathrm{z}$时会选择的潜在结果。
- 令$D_{1i}$表示工具变量$Z_i=1$时个体$i$ 的处理状态，$D_{0i}$表示工具变量$\mathrm{Z}_i=0$时个体$i$ 的处理状态：

$$D_i=D_{0i}+(D_{1i}-D_{0i})Z_i=\pi_0+\pi_{1i}Z_i+\xi_i$$

- 如果用随机系数的记号表示，那么有$\pi_0\equiv E[\mathrm{D}_{0i}]$和$\pi_{1i}\equiv(\mathrm{D}_{1i}-\mathrm{D}_{0i})$，因此$\pi_{1i}$是工具变量带给$\mathrm{D}_i$的异质性因果效应。

# 异质性因果效应框架下的假设

- 假设一：独立性假设
- 工具变量因该起到随机分配的效果，即工具变量既于潜在结果无关，也与潜在处理状态无关。

$$[\{\mathrm{Y}_i(d,z);\ \forall\, d,z\}, D_{1i}, D_{0i}] \perp Z_i$$

# 异质性因果效应框架下的假设

- 具体而言：

$$E[Y_i|Z_i=1]-E[Y_i|Z_i=0]$$
$$=E[Y_i(\mathrm{D}_{1i},1)|Z_i=1]-E[Y_i(D_{0i},0)|Z_i=0]$$
$$=E[Y_i(\mathrm{D}_{1i},1)-Y_i(D_{0i},0)]$$

- 同时意味着：

$$E[D_i|Z_i=1]-E[\mathrm{D}_i|Z_i=0]$$
$$=E[D_{1i}|Z_i=1]-E[D_{0i}|Z_i=0]$$
$$=E[D_{1i}-D_{0i}]$$

# 异质性因果效应框架下的假设

- 假设二：排他性约束
- 工具变量只能通过唯一的已知途径对被解释变量产生影响，即$Y_i(d,z)$只是$d$的的函数。
- $Y_i(d,0)=Y_i(d,1)$，对于$d=0$，1
- 在因果效应为常数且在因果模型中是线性的时候，我们把工具变量排除在等式便是的因果模型之外，并且宣称$E[Z_i\eta_i]=0$，以此来实现排他性约束。

# 异质性因果效应框架下的假设

- 使用排他性约束后，我们可以用不同的处理状态来定义潜在结果：

$$Y_{1i} \equiv Y_i(1,0) = Y_i(1,0)$$
$$Y_{0i} \equiv Y_i(0,1) = Y_i(0,0)$$

- 于是观察到的结果$Y_i$可以写成潜在结果的组合：

$$Y_i = Y_i(0, Z_i) + [Y_i(1, Z_i) - Y_i(0, Z_i)]\mathrm{D}_i$$
$$= Y_{0i} + (Y_{1i} - Y_{0i})D_i$$

- 用随机参数表示就是：

$$Y_i = \alpha_0 + \rho_i D_i + \eta_i;$$

其中$\alpha_0 \equiv E[Y_{0i}]$和$\rho_i \equiv Y_{1i} - Y_{0i}$

# 异质性因果效应框架下的假设

- 假设三：单调性假设
- 对所有的$i$，要么$\pi_i \geqq 0$，要么$\pi_i \leqq 0$。
- 该假设允许一些人不被工具变量影响，但对于那些受工具变量影响的人，这种影响必须是以相同方式发生的。换言之，对所有的$i$，或者$D_{1i} \geqq D_{0i}$或者$D_{1i} \leqq D_{0i}$。在接下来的讨论中，假设$D_{1i} \geqq D_{0i}$。
- 在用随机数决定参军资格的例子中，单调性假设意味着虽然对某些人而言参军资格不影响他入伍服役的概率，但是不存在获得参军资格而不允许其参军服役的事情。

# 局部平均处理效应定理(The LATE Theorem)

- 假设：

1.独立性：$[\{Y_i(d,z);\ \forall\, d,z\}, D_{1i}, D_{0i}] \perp Z_i$；

2.排他性：$Y_i(d,0) = Y_i(d,1)$，对于$d = 0$，$1$；

3.第一阶段存在：$E[D_{1i} - D_{0i}] \neq 0$;

4.单调性：$D_{1i} - D_{0i} \geqq 0\ \forall\, i$，当然这个假设中的不等号反过来也可；

- 于是：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]} = E[Y_{1i}-Y_{0i}|D_{1i}>D_{0i}]$$
$$= E[\rho_i|\pi_i>0]$$

# 局部平均处理效应定理(The LATE Theorem)

- 证明：

$$
\begin{aligned}
&E[Y_i|Z_i=1]-E[Y_i|Z_i=0]\\
&=E[Y_{0i}+(Y_{1i}-Y_{0i})D_{1i}|Z_i=1]-E[Y_{0i}+(Y_{1i}-Y_{0i})D_{0i}|Z_i=0]\\
&=E[Y_{0i}+(Y_{1i}-Y_{0i})D_{1i}]-E[Y_{0i}+(Y_{1i}-Y_{0i})D_{0i}]\\
&=E[(Y_{1i}-Y_{0i})(D_{1i}-D_{0i})]\\
&=E[Y_{1i}-Y_{0i}|D_{1i}>D_{0i}]P[D_{1i}>D_{0i}]
\end{aligned}
$$

$$
\begin{aligned}
&E[D_i|Z_i=1]-E[D_i|Z_i=0]\\
&=E[D_{1i}-D_{0i}]=P[D_{1i}>D_{0i}]
\end{aligned}
$$

# 局部平均处理效应定理(The LATE Theorem)

- 这个定理指出如果工具变量能够像随机分配好那样，能够通过唯一的已知途径影响潜在结果，存在第一阶段，只通过一个方向影响因果性，那么这个工具变量就可以用来估计被其影响的群体的平均因果效应。
- 在用随机抽取的参军资格做工具变量研究服役对收入的影响的例子中，我们得到的瓦尔德估计值实际上是对于那些仅仅因为在随机抽取中获得了参军资格而接受服役的人，服役对其收入的影响。这个估计值没有包括志愿参军对个人收入的影响，也没有在计算平均因果效应时考虑因病而免于服役的那些人。

# 单调性假设的意义

- 不满足单调性意味着工具变量使一些人的处理状态从0变成1，却使得另外一些人的处理状态从1变成0。

$$E[Y_i|Z_i=1]-E[Y_i|Z_i=0]$$
$$=E[(Y_{1i}-Y_{0i})(D_{1i}-D_{0i})]$$
$$=E[Y_{1i}-Y_{0i}|D_{1i}>D_{0i}]P[D_{1i}>D_{0i}]$$
$$-E[Y_{1i}-Y_{0i}|D_{1i}<D_{0i}]P[D_{1i}<D_{0i}]$$

- 由此我们发现即使所有人的因果效应都是正的，但是因为依从工具变量的个体的因果效应被对抗者的因果效应给抵消掉了，简约式估计值也还可能为0。

# 依从工具变量的子集

- 在局部平均处理效应框架下，依照总体中个体对工具变量作出的反应，任何总体可被分为三类子集：

1）依从工具变量者：满足$D_{1i}=1$和$D_{0i}=0$的子集；

2）始终接受者：满足$D_{1i}=1$和$D_{0i}=1$的子集；

3）从不接收者：满足$D_{1i}=0$和$D_{0i}=1$的子集。

- 局部平均处理效应便是依从工具变量者的平均因果效应。

# 依从工具变量的子集

- 依从工具变量者的平均因果效应往往不等于接受处理的人（也就是$D_i = 1$）的平均处理效应。

$$\underbrace{E[Y_{1i} - Y_{0i} | D_i = 1]}_{\text{接受处理的效应}}$$

$$= \underbrace{E[Y_{1i} - Y_{0i} | D_{0i} = 1]}_{\text{始终接受者的效应}} P[D_{0i} = 1 | D_i = 1]$$

$$+ \underbrace{E[Y_{1i} - Y_{0i} | D_{1i} > D_{0i}]}_{\text{依从接收者的效应}} P[D_{1i} > D_{0i}, Z_i = 1 | D_i = 1]$$

- 接受处理的个体的因果效应时依从工具变量者和始终接受者的因果效应的加权平均值。

# 依从工具变量的子集

- 类似的，未接受处理的个体的平均因果效应也不等于局部平均处理效应。

$$\underbrace{E[Y_{1i} - Y_{0i}|D_i = 0]}_{\text{未接受处理的效应}}$$

$$= \underbrace{E[Y_{1i} - Y_{0i}|D_{0i} = 0]}_{\text{从不接受者的效应}}P[D_{1i} = 0|D_i = 0]$$

$$+\underbrace{E[Y_{1i} - Y_{0i}|D_{1i} > D_{0i}]}_{\text{依从接收者的效应}}P[D_{1i} > D_{0i}, Z_i = 1|D_i = 0]$$

- 未受处理的个体的平均因果效应等于依从工具变量者和从不接收者的因果效应的加权平均。

# 依从工具变量的子集

- 最后，将两个等式加权平均后可得：

$$E[Y_{1i}-Y_{0i}]$$
$$=E[Y_{1i}-Y_{0i}|D_i=1]P[D_i=1]+E[Y_{1i}-Y_{0i}|D_i=0]P[D_i=0]$$

- 上式指出无条件的加权的平均因果效应时对依从工具变量者，始终接受者和从不接受者的因果效应的加权平均值。

# 两个特例

- 使用双胞胎做工具变量研究生育三个孩子对收入的影响。

原因在于第二胎生育双胞胎的所有女性最终都会有三个孩子，不存在从不接受者。

- 使用义务教育法中发生的变化（英国义务教育法将强制接受教育的年龄从14岁提高到15岁）作为工具变量对英国教育回报进行估计。

原因在于当义务教育法加强时，所有人都接受了新增加一年的教育，所以样本中不存在从不接受者。

# 随机实验中的工具变量

- 工具变量产生于随机实验，由于个体选择是否接受随机分配的处理本身就存在自选择问题，所以简单地比较处理组合实验组之间状况的差距并以此作为因果效应，可能存在较大的误导性。
- 将随机分配的处理记作$Z_i$，将个体接受的处理记为$D_i$，用工具变量就可以解决自选择问题，因为存在不依从随机分配的处理的情况，所以$Z_i$和$D_i$之间可能不相等。

# 随机实验中的工具变量

表 4.5　从 JTPA 实验中得到的结果：对培训项目的最小二乘估计值和工具变量估计值

| | 用接受培训状况进行比较 (OLS) | | 用接受处理状况进行比较 (ITT) | | 工具变量估计值 (IV) | |
|---|---|---|---|---|---|---|
| | 没有协变量 (1) | 存在协变量 (2) | 没有协变量 (3) | 存在协变量 (4) | 没有协变量 (5) | 存在协变量 (6) |
| A. 男性 | 3 970 | 3 754 | 1 117 | 970 | 1 825 | 1 593 |
| | (555) | (536) | (569) | (546) | (928) | (895) |
| B. 女性 | 2 133 | 2 215 | 1 243 | 1 139 | 1 942 | 1 780 |
| | (345) | (334) | (359) | (341) | (560) | (532) |

# 估计被处理者的处理效应

- Bloom Result。假设局部平均处理效应定理所要求的假设都成立，并且有$E[D_i|Z_i=0]=P[D_i=1|Z_i=0]=0$。那么有：

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i=1|Z_i=1]}$$

$$=\frac{E[Y_{0i}|Z_i=1]+E[(Y_{1i}-Y_{0i})D_i|Z_i=1]-E[Y_{0i}|Z_i=0]}{E[D_i=1|Z_i=1]}$$ 独立性

$$=\frac{E[(Y_{1i}-Y_{0i})D_i|Z_i=1]}{E[D_i=1|Z_i=1]}$$

$$=\frac{E[(Y_{1i}-Y_{0i})D_i=1,Z_i=1]P[D_i=1|Z_i=1]}{P[D_i=1|Z_i=1]}$$

$$=E[Y_{1i}-Y_{0i}|D_i=1]$$

# 依从工具变量者的特征

- 对依从工具变量者所成集合的规模进行度量，给定单调性，它正好是瓦尔德估计的第一阶段值。

$$P[D_{1i} > D_{0i}] = E[D_{1i} - D_{0i}]$$
$$= E[D_{1i}] - E[D_{0i}]$$
$$= E[D_{1i}|Z_i = 1] - E[D_{0i}|Z_i = 0]$$

- 我们还可以指出被处理组中有多少人是依从变量者：

$$P[D_{1i} > D_{0i}|D_i = 1] = \frac{P[D_i=1|D_{1i}>D_{0i}]P[D_{1i}>D_{0i}]}{P[D_i=1]}$$
$$= \frac{P[Z_i=1|D_{1i}>D_{0i}]P[D_{1i}>D_{0i}]}{P[D_i=1]}$$ 独立性
$$= \frac{P[Z_i=1]P[D_{1i}>D_{0i}]}{P[D_i=1]}$$
$$= \frac{P[Z_i=1](E[D_i|Z_i=1])-E[D_i|Z_i=0]}{P[D_i=1]}$$

# 依从工具变量者的特征

表 4.6 在工具变量研究中响应工具变量的概率

| 来源 | 内生变量 | 工具变量 | 样本 | $P[D=1]$ | 第一阶段 $P[D_1>D_0]$ | $P[Z=1]$ | 依从概率 $P[D_1>D_0 \mid D=1]$ | 依从概率 $P[D_1>D_0 \mid D=0]$ |
|---|---|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Angrist (1990) | 服役状态 | 随机抽取的参军资格 | 生于 1950 年的白人男性 | 0.267 | 0.159 | 0.534 | 0.318 | 0.101 |
| | | | 生于 1950 年的非白人男性 | 0.163 | 0.060 | 0.534 | 0.197 | 0.033 |
| Angrist 和 Evans(1998) | 超过两个孩子 | 第二胎中的双胞胎 | 年龄在 21—35 岁的已婚女性，在 1980 年生育两个以上的孩子 | 0.381 | 0.603 | 0.008 | 0.013 | 0.966 |
| | | 前两个孩子性别相同 | | 0.381 | 0.060 | 0.506 | 0.080 | 0.048 |
| Angrist 和 Krueger (1991) | 高中毕业 | 在第三和第四季度出生 | 生于 1930 年到 1939 年的男性 | 0.770 | 0.016 | 0.509 | 0.011 | 0.034 |
| Acemoglu 和 Angrist (2000) | 高中毕业 | 州要求接受至少 11 年的学校教育 | 年龄在 40—49 岁的男性 | 0.617 | 0.037 | 0.300 | 0.018 | 0.068 |

# 依从工具变量者的个体特征

- 令$x_i$是一个遵循伯努利分布的个体特征，比如表征大学毕业的虚拟变量，现考虑用性别组成作工具变量，对于那些依从工具变量的女性而言，生育两个孩子和生育两个以上孩子的女性，他们在教育水平方面有没有不同。

- $$\frac{P[x_{1i}=1|D_{1i}>D_{0i}]}{P[x_{1i}=1]}=\frac{P[D_{1i}>D_{0i}|x_{1i}=1]}{P[D_{1i}>D_{0i}]}$$

$$=\frac{E[D_i|Z_i=1,x_{1i}=1]-E[D_i|Z_i=0,x_{1i}=1]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}$$

依从工具变量者时大学生的可能性等与针对大学生群体做的第一阶段回归的结果与总体回归结果之比。

# 对局部平均处理效应的推广

- 局部平均处理效应定理适用于无协变量，只用单个虚拟变量做工具变量去估计只存在单一处理时的因果效应，时进行因果推断的最基本模型。
- 三个方向推广：

1）存在多个工具变量的情况（用一组虚拟变量来表示出生季度）

2）模型中存在协变量的情况（比如存在用以控制出生年份的控制变量）

3）因果效应取多个值甚至是连续值的情况（比如接受一年教育和接受两年教育对收入的影响不同）

# 多工具变量下的局部处理效应

- 考虑一对以虚拟变量身份出现的工具变量$Z_{1i}$和$Z_{2i}$，假设他们互斥，分别使用他们，可以构造出两个瓦尔德估计值，通过2SLS得到的局部平均处理效应的线性组合是一种加权平均值。令：

$$\rho_j = \frac{cov(Y_i, Z_{ji})}{cov(D_i, Z_{ji})}, j = 1,2$$

表示分别使用工具变量$Z_{1i}$和$Z_{2i}$后得到的两个工具变量估计值。

- 如果同时使用工具变量$Z_{1i}$和$Z_{2i}$，我们在2SLS第一阶段得到的总体拟合值就是$\widehat{D}_i = \pi_{11}Z_{1i} + \pi_{12}Z_{2i}$，其中$\pi_{11}$和$\pi_{12}$都是正数。

# 多工具变量下的局部处理效应

$$\rho_{2SLS} = \frac{cov(Y_i, \widehat{D}_i)}{cov(D_i, \widehat{D}_i)} = \frac{\pi_{11} cov(Y_i, Z_{1i})}{cov(D_i, \widehat{D}_i)} + \frac{\pi_{12} cov(Y_i, Z_{2i})}{cov(D_i, \widehat{D}_i)}$$

$$= \left[\frac{\pi_{11} cov(D_i, Z_{1i})}{cov(D_i, \widehat{D}_i)}\right]\left[\frac{cov(Y_i, Z_{1i})}{cov(D_i, Z_{1i})}\right] + \left[\frac{\pi_{12} cov(D_i, Z_{2i})}{cov(D_i, \widehat{D}_i)}\right]\left[\frac{cov(Y_i, Z_{2i})}{cov(D_i, Z_{2i})}\right]$$

$$= \psi\rho_1 + (1-\psi)\rho_2$$

- $\rho_1$是使用工具变量$Z_{1i}$得到的局部处理效应，$\rho_2$是使用工具变量$Z_{2i}$得到的局部处理效应，并且

$$\psi = \frac{\pi_{11} cov(D_i, Z_{1i})}{\pi_{11} cov(D_i, Z_{1i}) + \pi_{12} cov(D_i, Z_{2i})}$$

是介于0和1之间的一个数字，它的大小依赖于第一阶段中每个工具变量的相对重要性。

# 存在协变量的异质性因果模型

- 当我们将工具变量看作某种随机实验时，它和协变量相互独立，协变量不会发挥作用，但并非所有的工具变量都有这种性质。将其纳入使用工具变量进行因果分析的模型中，其主要原因为控制了协变量后条件独立性和排他性约束会更可能成立。

- 存在协变量时工具变量估计需要的假设变为条件独立假设：

$$[\{Y_i(d,z);\ \forall\, d,z\}, D_{1i}, D_{0i}] \perp Z_i | X_i$$

# 存在协变量的异质性因果模型

- 将协变量纳入模型的第二个理由是它可以降低被解释变量的变化，这可以带来更精确的2SLS估计值。
- 存在协变量的常因果效应的总体方方程的函数形式：

$$E[Y_{0i}|X_i] = X_i{}'\alpha^*, \text{其中}\alpha^*\text{是}\kappa \times 1\text{的参数向量}$$

$$Y_{1i} - Y_{0i} = \rho$$

将其与条件独立假设结合起来，就得到2SLS估计方程。

- 对常数因果效应模型直接推广得：

$$Y_{1i} - Y_{0i} = \rho(X_i)$$

其中$\rho(X_i)$是关于协变量$X_i$的确定型函数。

# 存在协变量的异质性因果模型

- 通过将$Z_i$和$X_i$的交互项加入第一阶段回归，将$D_i$和$X_i$的交互项加入第二阶段回归，就可以开始估计了。

所以第一阶段回归方程组写成：

$$D_i = X_i{}'\pi_{00} + \pi_{01}Z_i + Z_iX_i{}'\pi_{02} + \xi_{0i}$$
$$D_iX_i = X_i{}'\pi_{10} + \pi_{11}Z_i + Z_iX_i{}'\pi_{12} + \xi_{1i}$$

第二阶段方程是：

$$Y_i = \alpha' X_i + \rho_0 D_i + D_iX_i{}'\rho_1 + \xi_{1i}$$

- 于是得$\rho(X_i) = \rho_0 + \rho_1{}'X_i$，可以选择在关于$X_i$饱和的子集中用2SLS对$\rho(X_i)$进行非参数估计。

# 存在协变量的异质性因果模型

- 我们还可以在条件独立假设下求解异质性因果效应模型的局部平均处理效应定理，对于每一个特定的值，定义与该协变量相对应的局部平均处理效应为：

$$\lambda(X_i) \equiv E[Y_{1i} - Y_{0i}|X_i, D_{1i} > D_{0i}]$$

- 其中$\pi_x$和$\alpha_x$代表协变量得饱和模型（表示$X_i$所有可能取值得虚拟变量组成的集合），$\pi_{1x}$表示给定$X_i$得每个取值，$Z_i$在第一阶段的影响。

# 饱和加权定理

- 设给定后局部平均处理效应定理所要求的假设都满足。也就是说：

1.独立性：$[\{\mathrm{Y}_i(d,z);\ \forall\, d,z\}, D_{1i}, D_{0i}] \perp Z_i\ |X_i$；

2.排他性：$\mathrm{P}[Y_i(d,0)=Y_i(d,1)|X_i]=1$，对于$d=0$，1；

3.第一阶段存在：$E[D_{1i}-D_{0i}|X_i]\neq 0$;

4.单调性：$D_{1i}-D_{0i}\geqq 0\ \forall\, i$，当然这个假设中的不等号反过来也可；

我们基于第一阶段方程：

$$D_i=\pi_x+\pi_{1x}Z_i+\xi_{1i}$$

以及第二阶段方程：

$$Y_i=\alpha_x+\rho D_i+\eta_i$$

来估计2SLS估计值。

# 饱和加权定理

然后$\rho = E[\omega(X_i)\lambda(X_i)]$，其中：

$$\omega(X_i) = \frac{V\{E[D_i|X_i, Z_i]X_i\}}{E[V\{E[D_i|X_i, Z_i]X\}]}$$

并且

$$V\{E[D_i|X_i, Z_i]X_i\} = \mathrm{E}\{E[D_i|X_i, Z_i](E[D_i|X_i, Z_i] - E[D_i|X_i\ ])|X_i\}$$

- 这个定理指出在第一阶段使用完全饱和模型，在第二阶段使用饱和模型得到的2SLS估计值就是对每个协变量取值下估计出的局部平均处理效应得加权平均，在$X_i$可以取到的每个值上，权重与第一阶段拟合值$E[D_i|X_i, Z_i]$的条件方差均值成正比。

# 存在协变量的异质性因果模型

- 在实际中，不可能对所有$X_i$的可能取值都估计出一个系数。首先可能会带来偏误，其次大量不精确的第一阶段拟合值看上去也不好。在Abadie的研究中，就给出一个$E[Y_i|D_i,X_i,D_{1i}>D_{0i}]$最小均方误意义下的逼近。
- 对于依从工具变量者，给定$X_i$后处理组和控制组之间的平均差别等于给定$X_i$后的局部平均处理效应：

$$E[Y_i|D_i=1,X_i,D_{1i}>D_{0i}]-E[Y_i|D_i=0,X_i,D_{1i}>D_{0i}]$$
$$=E[Y_{1i}-Y_{0i}|X_i,D_{1i}>D_{0i}]$$
$$=\lambda(X_i)$$

- 问题在于如何识别出依从变量者。

# Abadie Kapper Theorem

- 考虑给定$X_i$后局部平均因果处理效应所要求的假设都满足。记$g(Y_i, D_i, X_i)$为定义在$(Y_i, D_i, X_i)$上的任何可测函数，其期望值有限，定义：
- 依从工具变量者就是除去始终接收者和从不接受者剩下的那部分人。

$$k_i = 1 - \frac{D_i(1 - Z_i)}{1 - P(Z_i = 1|X_i)} - \frac{(1 - D_i)Z_i}{P(Z_i = 1|X_i)}$$

那么：

$$E[g(Y_i, D_i, X_i)|D_{1i} > D_{0i}] = \frac{E[k_i g(Y_i, D_i, X_i)]}{E[k_i]}$$

# 存在协变量的异质性因果模型

- 在线性回归中逼近$E[Y_i|D_i,X_i,D_{1i}>D_{0i}]$，令

$$(\alpha_c,\beta_c)=\arg\min_{a,b}E\{(E[Y_i|D_i,X_i,D_{1i}>D_{0i}]-aD_i-X_i'b)^2|D_{1i}>D_{0i}\}$$

运用Abadie Kapper定理指出可以通过下式来逼近函数$E[Y_i|D_i,X_i,D_{1i}>D_{0i}]$：

$$(\alpha_c,\beta_c)=\arg\min_{a,b}E\{k_i(Y_i-aD_i-X_i'b)^2\}$$

- 其中$E\{k_i(Y_i-aD_i-X_i'b)^2\}$是相应的最小二乘的最小化元，这里Abadie Kapper函数被用来对普通最小而二乘中的最小化元$(Y_i-aD_i-X_i'b)^2$进行加权。

# 存在多种处理强度时的平均因果响应

- 取值为{0,1}的虚拟变量带来的因果效应与取值为{0,1,2, … }的因果变量带来的因果效应之间有很大的不同。
- 在研究教育带来的经济回报时，令：

$$Y_i \equiv f_i(s)$$

表示个体$i$在接受$s$年教育后应该获得的收入。

- 假设$s_i$可以在集合$\{0,1,\ldots,\bar{s}\}$，就存在$\bar{s}$个因果效应$Y_{si}-Y_{s-1,i}$，2SLS提供了对单位因果效应计算加权的工具。
- 记$s_{0i}$表示$Z_i=0$时个体$i$接受的教育水平，$s_{1i}$表示$Z_i=1$时个体$i$接受的教育水平。

# 平均因果响应定理

- 假设：

ACR1，独立性假设和排除性约束

$$\{Y_{0i}, Y_{1i}, \ldots, Y_{ni}; s_{0i}, s_{1i}\} \perp Z_i \ ;$$

ACR2，第一阶段 $E[s_{1i} - s_{0i}] \neq 0$;

ACR3,单调性 $s_{1i} - s_{0i} \geqq 0 \ \forall i$, 或者反之，

那么

$$\frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[s_i|Z_i = 1] - E[s_i|Z_i = 0]} = \sum_{s=1}^{\bar{s}} \omega_s E\left[Y_{si} - Y_{s-1,i} | s_{1i} \geqq s > s_{0i}\right]$$

- 其中 $\omega_s = \frac{P[s_{1i} \geqq s > s_{0i}]}{\sum_{s=1}^{\bar{s}} P[s_{1i} \geqq s > s_{0i}]}$，这个权重是非负的而且相加为1的。

# 因果变量是连续的

- 如果我们感兴趣的因果变量是连续的，那么我们可以将因果响应函数看作微分。令$q_i(p)$表示在市场中价格为$p$时的需求量。
- 研究者们使用虚拟变量$stormy_i$来表示出现大风大浪的天气，以此估计对鱼类的需求。使用$stormy_i$做工具变量得到的瓦尔德估计值可以用下面的方程来解释：

$$\frac{E[q_i|stormy_i=1]-E[q_i|stormy_i=0]}{E[p_i|stormy_i=1]-E[p_i|stormy_i]}$$
$$=\frac{\int E[q_i(t)'|P_{1i}\geqq t>P_{0i}]P[P_{1i}\geqq t>P_{0i}]dt}{\int P[P_{1i}\geqq t>P_{0i}]dt}$$

- 其中$p_i$是在市场$i$中的日观测价格，$P_{1i}$和$P_{0i}$是用$stormy_i$表示的潜在价格，可以用$P[P_{1i}\geqq t>P_{0i}]$作为权重函数，来估计价格为$t$时需求量的倒数的平均值。