Recalling IV Regression

- In the econometric course we learned something about IV
- or something called “LATE”
- used to deal with “endogeneity” and selection, etc.
- What’s the relationship with Roy model?

Review of the IV Framework

$$Y_i = \beta D_i + \epsilon_i$$
$$D_i = \gamma Z_i + \eta_i$$

- Assumptions:
  - Exogeneity: $Z_i \perp \epsilon_i$
  - Relevance: $Cov(Z_i, D_i) \neq 0$
  - Exclusion: Z_i affects Y_i only through D_i
- Results:
  - $\beta = \frac{Cov(Y,Z)}{Cov(D,Z)}$
- Implication:
  - Find some instrument to identify the "causal effect"

Causal Implications

- What does the above IV framework imply about potential outcomes?

$$Y_1 - Y_0 = \beta$$

- No effect heterogeneity
- ATE = ATT = ...
- Against the core idea of Roy model or potential outcome framework

A More Flexible Framework (LATE)

- What if β is heterogeneous?
- As in potential outcome framework, write $\{D_z\}_{z \in \mathcal{Z}}$
- D_z means the treatment status D when $Z = z$
- Z is still the instrument
- Now assume D and Z are both binary

LATE Assumptions

1. Exogeneity: $(Y_1, Y_0, D_1, D_0) \perp Z | X$
2. Relevance: $P(D = 1 | X, Z = 1) \neq P(D = 1 | X, Z = 0)$ a.s.
3. Monotonicity: $P(D_1 \geqslant D_0 | X) = 1$ a.s.
   - also known as no defier
4. Overlap: $P(Z = 1 | X) \in (0, 1)$ a.s.

Results

- Result 1:

$$\beta_{IV} \equiv \frac{Cov(Y, Z)}{Cov(D, Z)} = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]}$$

- Result 2:

$$\beta_{IV} = E[Y_1 - Y_0|D_1 = 1, D_0 = 0] = \text{LATE}$$

- LATE means local average treatment effect
- The "local" means effects for compliers

Proof: Result 1

Note that for any random variable W and binary Z

$$
\begin{aligned}
Cov(W,Z) &= E[W(Z-E(Z))] \\
&= E[W(Z-E(Z))|Z=1]Pr(Z=1) + E[W(Z-E(Z))|Z=0]Pr(Z=0) \\
&= E(W(1-E(Z))|Z=1)E(Z) - E(WE(Z)|Z=0)(1-E(Z)) \\
&= E(W|Z=1)(1-E(Z))E(Z) - E(W|Z=0)(1-E(Z))E(Z) \\
&= [E(W|Z=1) - E(W|Z=0)]E(Z)(1-E(Z)) \\
\frac{Cov(Y,Z)}{Cov(D,Z)} &= \frac{[E(Y|Z=1) - E(Y|Z=0)]E(Z)(1-E(Z))}{[E(D|Z=1) - E(D|Z=0)]E(Z)(1-E(Z))} \\
&= \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]}
\end{aligned}
$$

Proof: Result 2

$$\begin{aligned} E[Y|Z=z] &= E[DY_1 + (1-D)Y_0|Z=z] \\ &= E[D_z Y_1 + (1-D_z)Y_0|Z=z] \\ &= E[D_z Y_1 + (1-D_z)Y_0], \text{(Exogeneity)} \end{aligned}$$

Therefore for any two z and z', we have:

$$\begin{aligned} E[Y|Z=z] - E[Y|Z=z'] &= E[(D_z - D_{z'})Y_1 + (1 - D_z - 1 + D_{z'})Y_0] \\ &= E[(D_z - D_{z'})(Y_1 - Y_0)] \\ &= E[Y_1 - Y_0|D_z = 1, D_{z'} = 0]P(D_z = 1, D_{z'} = 0), \\ &\quad \text{(Law of total probability)} \end{aligned}$$

Similar for D, using the same logic,

$$E[D|Z=z] - E[D|Z=z'] = E[D_z - D_{z'}] = P(D_z = 1, D_{z'} = 0)$$

Taken Together

Now take the ratio, we have:

$$\frac{E[Y|Z=1]-E[Y|Z=0]}{E[D|Z=1]-E[D|Z=0]} = E[Y_1 - Y_0|D_z = 1, D_{z'} = 0]$$

Taken together, we know

$$\beta_{IV} \equiv \frac{Cov(Y,Z)}{Cov(D,Z)} = \frac{E[Y|Z=1]-E[Y|Z=0]}{E[D|Z=1]-E[D|Z=0]} = E[Y_1 - Y_0|D_1 = 1, D_0 = 0] = \text{LATE}$$

Remarks

- What do we learn after all these derivations?
- $\beta_{IV} = \text{LATE} = E[Y_1 - Y_0 | D_1 = 1, D_0 = 0]$
- The IV gives some specific treatment effects
- The treatment effects for the compliers
- Importantly, these treatment effects depend on the IV (Z) you use
- Very often these compliers are not the policy-relevant ones (e.g., earthquake IV)

Relation to Roy Model

- People find that the Roy model and LATE framework are equivalent
- Model 1: Imbens and Angrist (1994)
  1. $Y = DY_1 + (1 - D)Y_0$
  2. $D = ZD_1 + (1 - Z)D_0$
  3. (Y_1, Y_0, D_1, D_0)
- Model 2: Non-parametric Roy
  1. $Y = DY_1 + (1 - D)Y_0$
  2. $D = \mathbf{1}\{U < \nu(Z)\}$ for ν an unknown function
  3. (Y_1, Y_0, U)

Vytlacil (2002)

- These two are equivalent.
- The key is that $U < \nu(Z)$, it's a separate index.
- For instance, $\nu(U, Z)$ doesn't work.
- This can be used to define MTE:

$$E[Y_1 - Y_0 | U = u]$$

- Those with small u (close to 0) often choose $D = 1$.

Weighted MTE

- Everything is a weighted MTE (I write without proofs):
- ATE:

$$E[E[Y_1 - Y_0|U]] = \int_0^1 \text{MTE}(u)du$$

  (when we normalize U to uniform)
- ATT:

$$\int_0^1 \text{MTE}(u)\frac{P(p(Z) \geqslant u)}{P(D = 1)}du$$

- Those with low values of u are more highly weighted.