Part 8: Policy Evaluation- Local Average Treatment Effects

Chris Conlon

March 31, 2020

Applied Econometrics

What about IV

So what does IV do?

- Let's assume a binary instrument $Z_i = 1$
- $Y_i(1), Y_i(0)$ depends on the value of T_i
- But now we endogenize $T_i(1), T_i(0)$ where the argument is the value of Z_i.
- We observe $\{Z_i, T_i = T_i(Z_i), Y_i = Y_i(T_i(Z_i))\}$.

IV Assumptions

So what does IV do?

Independence $Z_i \perp Y_i(1), Y_i(0), T_i(1), T_i(0)$. Instrument is as if randomly assigned and does not directly affect Y_i

This is not implied by random assignment. In that case there would be four potential outcomes $Y_i(z, t)$

Random Assignment $Z_i \perp Y_i(0,0), Y_i(0,1), Y_i(1,0), Y_i(1,1), T_i(1), T_i(0)$.

Exclusion Restriction $Y_i(z, t) = Y_i(z', t)$ for all z, z', t.

Thus we require both RA and ER to guarantee Independence. The second assumption is a substantive one.

We only observe (Z_i, T_i) not the pair $T_i(0), T_i(1)$ so we cannot determine compliance types directly! (See the picture)

IV Assumptions

$T_i(1)$	$T_i(0)$ 0	1
0	never-taker	defier
1	complier	always-taker

IV Assumptions

We are stuck without further assumptions, so we assume:

Monotonicity/No Defiers $T_i(1) \geq T_i(0)$

- Works in many applications (classical drug compliance).
- Implied by many latent index models with constant coefficients
- Works as long as sign of $\pi_{1,i}$ doesn't change

$$T_i(z) = 1[\pi_0 + \pi_1 z + \varepsilon_i > 0]$$

IV Assumptions

Table 2: Compliance Type by Treatment and Instrument

		Z_i = 0	Z_i = 1
W_i	0	complier/never-taker	never-taker/defier
	1	always-taker/defier	complier/always-taker

IV Assumptions

Table 3: Compliance Type by Treatment and Instrument given Monotonicity

		Z_i = 0	Z_i = 1
W_i	0	complier/never-taker	never-taker
	1	always-taker	complier/always-taker

LATE Derivation

- We can derive the expression for β_{IV} as:

$$\beta_{IV} = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[T_i|Z_i = 1] - E[T_i|Z_i = 0]} = E[Y_i(1) - Y_i(0)|complier]$$

- We can derive the expression for π_c (the fraction of compliers):

$$\pi_c = E[T_i|Z_i = 1] - E[T_i|Z_i = 0]$$

- Proof see Angrist and Imbens

How Close to ATE?

Angrist and Imbens give some idea how close to the ATE the LATE is:

- $E[Y_i(0)|\text{never-taker}]$ and $E[Y_i(1)|\text{always-taker}]$ can be estimated from the data
- Compare these to their respective compliers $E[Y_i(0)|\text{complier}]$, $E[Y_i(1)|\text{complier}]$.
- When these are close then possibly $ATE \approx LATE$.

How Close to ATE?

Angrist and Imbens give some idea how close to the ATE the LATE is:

$$\widehat{\beta}_1^{TSLS} \to^p \frac{E[\beta_{1i}\pi_{1i}]}{E[\pi_{i1}]} = LATE$$

$$LATE = ATE + \frac{Cov(\beta_{1i}, \pi_{1i})}{E[\pi_{1i}]}$$

- Weighted average for people with large π_{1i}.
- Late is treatment effect for those whose probability of treatment is most influenced by Z_i.
- If you always (never) get treated you don't show up in LATE.

How Close to ATE?

- With different instruments you get different π_{1i} and TSLS estimators!
- Even with two valid Z_1, Z_2
    - Can be influential for different members of the population.
    - Using Z_1, TSLS will estimate the treatment effect for people whose probability of treatment X is most influenced by Z_1
    - The LATE for Z_1 might differ from the LATE for Z_2
    - A J-statistic might reject even if both Z_1 and Z_2 are exogenous! (Why?).

Example: Cardiac Catheterization

- $Y_i =$ surival time (days) for AMI patients
- $X_i =$ whether patient received cadiac catheterization (or not) (intensive treatment)
- $Z_i =$ differential distance to CC hospital

$$\begin{aligned} SurvivalDays_i &= \beta_0 + \beta_{1i} CardCath_i + u_i \\ CardCath_i &= \pi_0 + \pi_{1i} Distance_i + v_i \end{aligned}$$

- For whom does distance have the great effect on probability of treatment?
- For those patients what is their β_{1i}?

Example: Cardiac Catheterization

- IV estimates causal effect for patients whose value of X_i is most heavily influenced by Z_i
  - Patients with small positive benefit from CC in the expert judgement of EMT will receive CC if trip to CC hospital is short (compliers)
  - Patients that need CC to survive will always get it (always-takers)
  - Patients for which CC would be unnecessarily risky or harmful will not receive it (never-takers)
  - Patients for who would have gotten CC if they lived further from CC hospital (hopefully don't see) (defiers)
- We mostly weight towards the people with small positive benefits.

Local Average Treatment Effect

So how is this useful?

- It shows why IV can be meaningless when effects are heterogeneous.
- It shows that if the monotonicity assumption can be justified, IV estimates the effect for a particular subset of the population.
- In general the estimates are specific to that instrument and are not generalisable to other contexts.
- As an example consider two alternative policies that can increase participation in higher education.
  - Free tuition is randomly allocated to young people to attend college ($Z_1 = 1$ means that the subsidy is available).
  - The possibility of a competitive scholarship is available for free tuition ($Z_1 = 1$ means that the individual is allowed to compete for the scholarship).

Local Average Treatment Effect

- Suppose the aim is to use these two policies to estimate the returns to college education. In this case, the pair $\{Y^1, Y^0\}$ are log earnings, the treatment is going to college, and the instrument is one of the two randomly allocated programs.
- First, we need to assume that no one who intended to go to college will be discouraged from doing so as a result of the policy (monotonicity).
- This could fail as a result of a General Equilibrium response of the policy; for example, if it is perceived that the returns to college decline as a result of the increased supply, those with better outside opportunities may drop out.

Local Average Treatment Effect

- Now compare the two instruments.
- The subsidy is likely to draw poorer liquidity constrained students into college but not necessarily those with the highest returns.
- The scholarship is likely to draw in the best students, who may also have higher returns.
- It is not a priori possible to believe that the two policies will identify the same parameter, or that one experiment will allow us to learn about the returns for a broader/different group of individuals.

Local Average Treatment Effect

Finally, we need to understand what monotonicity means in terms of restrictions on economic theory.

- To quote from Vytlacil (2002) Econometrica:
  "The LATE assumptions are not weaker than the assumptions of a latent index model, but instead impose the same restrictions on the counterfactual data as the classical selection model if one does not impose parametric functional form or distributional assumptions on the latter."
- This is important because it shows that the LATE assumptions are equivalent to whatever economic modeling assumptions are required to justify the standard Heckman selection model and has no claim to greater generality.
- On the other hand there are no magical solutions to identifying effects when endogeneity/selection is present; this problem is exacerbated when the effects are heterogeneous and individuals select into treatment on the basis of the returns.