Labor Economics

Assignment: IV, Panel, and Roy Model

1 Normalization of the Selection Equation

In class, we said that ATT= $\int_0^1 \text{MTE}(u)\frac{P(p(Z)\geq u)}{P(D=1)}du$. This result relies on the following normalization.

$$D = \mathbf{1}\{U \leq \nu(Z)\} = \mathbf{1}\{\tilde{U} \leq p(Z)\}$$

, where p is the propensity score.

1. We write $D = \mathbf{1}\{F_U(U) \leq F_U(\nu(Z))\}$, where F_U is the CDF of the continuously distributed U. What is the distribution of $F_U(U)$? Let $\tilde{U} := F_U(U)$ from now on.

2. Define $p(z) := P(D = 1|Z = z)$, show that $p(z) = F_U(\nu(z))$.

3. Show that $D = \mathbf{1}\{\tilde{U} \leq p(Z)\}$.

2 Derivation of the Weights for LATE

In this exercise, we try to show that LATE of instrument z and z' can be written as:

$$\text{LATE}_{z'}^{z} = \frac{E[Y|Z=z] - E[Y|z=z']}{E[D|Z=z] - E[D|z=z']} = \int_0^1 \text{MTE}(u) \times \frac{\mathbf{1}\{u \in [p(z'), p(z)]\}}{p(z) - p(z')} du$$

1. Which part of the most right-hand side is $E[D|Z = z] - E[D|z = z']$ corresponding to?

2. Show that $E[Y|Z = z] = E[Y_1|U \leq p(z)]p(z) + E[Y_0|U > p(z)](1 - p(z))$

3. Show that $E[Y|Z = z] = \int_0^{p(z)} E[Y_1|U = u]du + \int_{p(z)}^1 E[Y_0|U = u]du$

4. Show that $E[Y|Z = z] - E[Y|Z = z] = \int_{p(z')}^{p(z)} \text{MTE}(u)du$

3 Policy Relevance Treatment Effect

We introduced LATE in class, but the "ideal" treatment effect depends on the research question. Let's take attending college for example.

1. Let $D \in \{0,1\}$ be attending college or not and let the outcome Y being the future average earning. What is the ATE measuring?

2. What is the ATT measuring?

3. Let Z be the tuition, Z^* be the tuition under the new policy, p^* be the propensity score under new policy, and $D^* = \mathbf{1}\{U \leq p^*(Z^*)\}$. Therefore $Y^* = D^* Y_1 + (1 - D^*) Y_0$. Define $\beta_{\text{PRTE}} = \frac{E(Y^*) - E(Y)}{E(D^*) - E(D)}$. Write down an argument why β_{PRTE} is more interesting than ATE or ATT.

4. When will β_{PRTE} equal to LATE?

4 Arellano-Bond

Consider the following model

$$Y_{it} = \rho Y_{it-1} + \delta_i + \epsilon_{it}$$

Assume $Cov(\epsilon_{it}, Y_{is}) = 0 \, \forall s \leq t - 1$ (sequential exogeneity).

1. Show that the fixed effect estimator cannot recover ρ consistently.

2. Take the first difference to difference out the individual fixed effect.

3. With the first difference, can OLS recover ρ? If not, can you propose an instrument?

5 Show that TWFE is Biased

1. Draw the figures in class to explain why TWFE cannot recover a postively weighted average of cohort specific ATT.

2. Simulate a data to illustrate your point above.

3. Look up the paper "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." Learn to use one of the packages listed in the paper's appendix. Show that you can recover one of the treatment effect from your simulation.