Notes for derivation of LATE and PRTE

LATE

Consider the Wald estimand (assume conditioning on X is implicit):

$$LATE = \frac{\mathbb{E}[Y|Z=z] - \mathbb{E}[Y|Z=z']}{\mathbb{E}[D|Z=z] - \mathbb{E}[D|Z=z']}$$

Lets take a closer look at the numerator:

$$\begin{aligned}
\mathbb{E}[Y|Z=z] &= \mathbb{E}[Y_1 D + Y_0(1-D)|Z=z] = \mathbb{E}[Y_1|Z=z, D_z=1]\Pr[D_z=1] + \mathbb{E}[Y_0|Z=z, D_z=0]\Pr[D_z=0] \\
&\underbrace{=}_{Z \perp\!\!\!\perp (Y_1, Y_0)} \mathbb{E}[Y_1|D_z=1]\Pr[D_z=1] + \mathbb{E}[Y_0|D_z=0]\Pr[D_z=0] \\
&\underbrace{=}_{D_z=1 \iff U \le p(z) \text{ and } U \sim \mathbb{U}[0,1]} \mathbb{E}[Y_1|U \le p(z)] \underbrace{\Pr[U \le p(z)]}_{=p(z)} + \mathbb{E}[Y_0|U > p(z)] \underbrace{\Pr[U > p(z)]}_{=1-p(z)} \\
&= \int_0^{p(z)} \mathbb{E}[Y_1|U=u] p(z) \underbrace{\frac{1}{p(z)}}_{U|_{U \le p(z)} \sim \mathbb{U}[0, p(z)]} du + \int_{p(z)}^{1} \mathbb{E}[Y_0|U=u](1-p(z)) \underbrace{\frac{1}{1-p(z)}}_{U|_{U > p(z)} \sim \mathbb{U}[p(z), 1]} du \\
&= \int_0^{p(z)} \mathbb{E}[Y_1|U=u] du + \int_{p(z)}^{1} \mathbb{E}[Y_0|U=u] du
\end{aligned}$$

Similarly:

$$\mathbb{E}[Y|Z=z'] = \int_{0}^{p(z')} \mathbb{E}[Y_1|U=u]du + \int_{p(z')}^{1} \mathbb{E}[Y_0|U=u]du$$

As a result:

$$\begin{aligned}\mathbb{E}[Y|Z=z] - \mathbb{E}[Y|Z=z'] &= \int_{p(z')}^{p(z)} \mathbb{E}[Y_1|U=u]du - \int_{p(z)}^{p(z)} \mathbb{E}[Y_0|U=u]du \\ &= \int_{p(z')}^{p(z)} \mathbb{E}[Y_1 - Y_0|U=u]du = \int_{p(z')}^{p(z)} MTE(u)du \\ &= \int_{0}^{1} MTE(u)\mathbb{1}[u \in [p(z'), p(z)]]du\end{aligned}$$

Denominator:

$$\mathbb{E}[D|Z=z] = \mathbb{E}[\mathbb{1}[U \leq p(z)]|Z=z] = \Pr[U \leq p(z)] = p(z)$$

Finally:

$$LATE = \int_{0}^{1} MTE(u)\frac{\mathbb{1}[u \in [p(z'), p(z)]]}{p(z) - p(z')}du$$

PRTE

Define Policy Relevant Treatment Effect as:

$$\text{PRTE} = \frac{\overbrace{\mathbb{E}[Y^*] - \mathbb{E}[Y]}^{\text{the mean effect}}}{\underbrace{\mathbb{E}[D^*] - \mathbb{E}[D]}_{\text{per net person}}}$$

Where:

$$Y = Y_1 D + Y_0(1 - D)$$
$$Y = Y_1 D^* + Y_0(1 - D^*)$$

For instance if the policy operates through the value of the instrument (which is only a very specific example, in general it may also affect $p(.)$ function), then:

$$D^* = \mathbb{1}[U \leq p(Z^*)]$$
$$\Pr[D^* = 1] = \Pr[U \leq p(Z^*)]$$

Meaning that policy changed the probability of selection into treatment. Think of the example when reduced tuition increases the probability of enrolling into college.

Note that we assume *policy invariance* meaning that the policy does not affect the potential outcomes! It is very frequently (but there are other possibilities) meaning that we abstract away from general equilibrium effects. In the example with college education we assume that increased ecrollment into college does not affect wage outcomes on labor market Y_1 and Y_0.

Now lets derive the expression for PRTE:

$$\begin{aligned}
\mathbb{E}[Y] = \mathbb{E}[Y_1 D + Y_0(1 - D)] &\underbrace{=}_{LIE} \mathbb{E}[\mathbb{E}[Y_1 D|U = u]] + \mathbb{E}[\mathbb{E}[Y_0(1 - D)|U = u]] \\
&= \mathbb{E}[\mathbb{E}[Y_1 \mathbb{1}[u \leq p(Z)]|U = u]] + \mathbb{E}[\mathbb{E}[Y_0 \mathbb{1}[u > p(Z)]|U = u]] \\
&= \mathbb{E}[\mathbb{E}[Y_1|u \leq p(Z), U = u]\Pr[u \leq p(Z)]] + \mathbb{E}[\mathbb{E}[Y_0|u > p(Z), U = u]\Pr[u > p(Z)]] \\
&\underbrace{=}_{Z \perp\!\!\!\perp (Y_1, Y_0)} \mathbb{E}[\mathbb{E}[Y_1|U = u]\Pr[u \leq p(Z)]] + \mathbb{E}[\mathbb{E}[Y_0|U = u]\Pr[u > p(Z)]] \\
&\text{recall that outer expectation is over } U \\
&= \int_0^1 (\mathbb{E}[Y_1|U = u]\Pr[u \leq p(Z)] + \mathbb{E}[Y_0|U = u]\Pr[u > p(Z)])\, du
\end{aligned}$$

Now, note that in $\Pr[u \leq p(Z)]$ and $\Pr[u > p(Z)]$ we calculate probability with respect to measure over Z and we never assumed continuity of its distribution, meaning that we did not assume anything about the CDF of $p(Z)$. So $\Pr[u > p(Z)] = F_p^-(u)$ and $\Pr[u \leq p(Z)] = 1 - F_p^-(u)$. Where $F_p^-(u)$ is the left limit at the point u (recall that we require CDF to be right continuous, but not necessary left continuous). As a result we have:

$$\begin{aligned}\mathbb{E}[Y] &= \int_0^1 \big(\mathbb{E}[Y_1|U=u](1-F_p^-(u)) + \mathbb{E}[Y_0|U=u]F_p^-(u)\big)\, du \\ &= \int_0^1 \mathbb{E}[Y_1|U=u] - F_p^-(u)\mathrm{MTE}(u)du\end{aligned}$$

Similarly:

$$\begin{aligned}\mathbb{E}[Y^*] &= \int_0^1 \big(\mathbb{E}[Y_1|U=u](1-F_{p^*}^-(u)) + \mathbb{E}[Y_0|U=u]F_{p^*}^-(u)\big)\, du \\ &= \int_0^1 \mathbb{E}[Y_1|U=u] - F_{p^*}^-(u)\mathrm{MTE}(u)du\end{aligned}$$

As a result:

$$\mathbb{E}[Y^*] - \mathbb{E}[Y] = \int_0^1 (F_p^-(u) - F_{p^*}^-(u))\mathrm{MTE}(u)du$$

Lastly, denominator:

$$\mathbb{E}[D] \underbrace{=}_{\text{LIE with outer expectation over Z}} \mathbb{E}[\mathbb{E}[\mathbb{1}[U \leq p(z)]|Z=z]] = \mathbb{E}[p(Z)]$$

$$\mathbb{E}[D^*] = \mathbb{E}[p^*(Z^*)]$$

Finally:

$$\text{PRTE} = \int_0^1 \frac{(F_p^-(u) - F_{p^*}^-(u))}{\mathbb{E}[p^*(Z^*)] - \mathbb{E}[p(Z)]} \text{MTE}(u) du$$

Note, using integration by parts we can show that weights integrate to 1: $\int F du = uF(u) - \int u f du$.

Figures and discussion from the end of the class

Figure 1 shows MTE in the case of selection on gains, ATE and 2 LATE's for 2 different pairs of instruments:

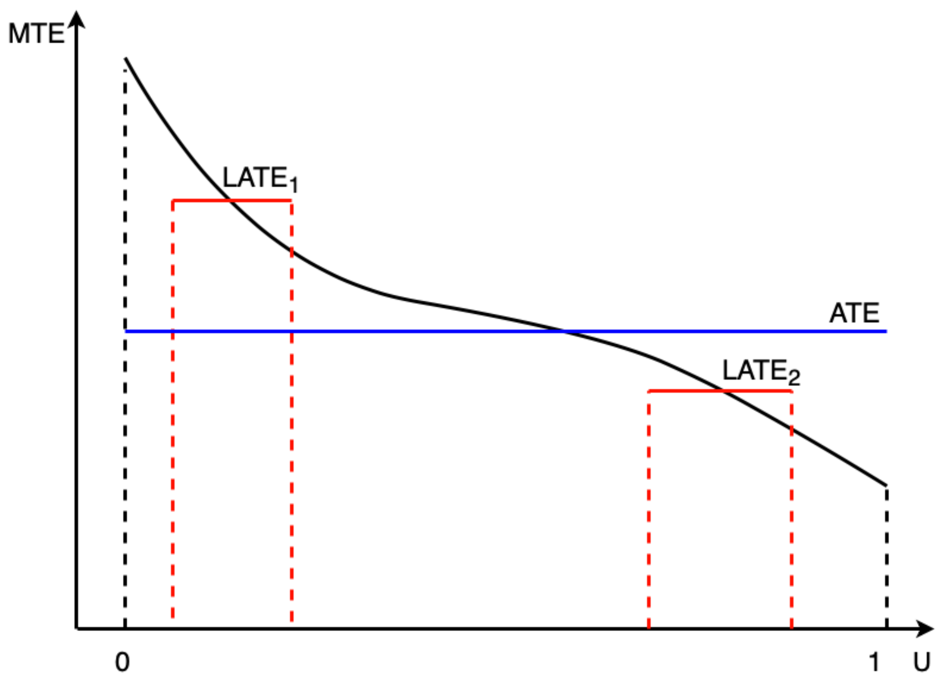

Figure 1: ATE and LATE for 2 pairs of instruments in the case of selection on the gains

Next, Marco asked a good question: why did I decompose ATT in the following way:

$$\text{ATT} = \mathbb{E}[Y_1 - Y_0 | D = 1] = \sum_z \mathbb{E}[Y_1 - Y_0 | D = 1, Z = z] \Pr[Z = z] = \sum_z \mathbb{E}[Y_1 - Y_0 | D_z = 1] \Pr[Z = z]$$

Instead of:

$$\mathbb{E}[Y_1 - Y_0|D = 1] = \mathbb{E}[\mathbb{E}[Y_1 - Y_0|D = 1, U = u]|D = 1] = \dots$$

As you will (most likely) end up doing in the Problem Set 1. The answer is that I wanted to visualize each component of ATT in the way that I think is useful, rather than derive the expression for ATT.

Ana asked an interesting question(if I understood it correctly, it happens to me that in the evening I hear with Russian accent): Why do we have a distribution of gains conditional on $U = u$? Note, that if selection is purely on gains: $D = \mathbb{1}[\underbrace{U_0 - U_1}_{=U} \leq X'(\beta_1 - \beta_0)]$ then clearly, by setting $U = u = U_1 - U_0$ we set $Y_1 - Y_0 = X'(\beta_1 - \beta_0) + u$ and conditional on X there is no variance left - so in this case the answer to the question is we do not have a distribution.

But this is a very particular case, generally, there are other parameters involved into decision to enroll - like costs. Meaning that setting $U = U_C + U_0 - U_1 = u \implies U_0 - U_1 = u - U_C$, so that we have variance in unobservable costs. In this case the answer to the question: we have distribution of gains because of varying unobservable costs.