Econometrics 3: Problem Set 2

Auplat Valentin, Hamburger Tom, Godechot Solal

07 of February, 2025

Instructions

Questions containing (*) do not necessarily have to be handed in. However, students are expected to think about them and be prepared to discuss in class.

Submit to: metrics2324@gmail.com. Deadline: One minute before the start of the first tutorial.

1 Rubin Model and Roy Model

We consider a labor market training program that may increase a given individual's wage from w_0 to w_1. Participation has a cost c, and we write:

$$w_1 = w_0 + \delta$$

where $\delta > 0$. Assume δ and w_0 are initially independent.

1. What is the treatment impact of a given individual i? What is the average treatment impact in the population?

For a given individual, the treatment impact is :

$$(w_{1i} - c) - w_{0i} = \delta_i - c \iff w_{1i} - w_{0i} = \delta_i$$

Average treatment impact in the population is thus given by :

$$ATE = \mathbb{E}[(w_{1i} - c) - w_{0i}] = \mathbb{E}[\delta] - c \iff ATE = \mathbb{E}(w_{1i} - w_{0i}) = \mathbb{E}(\delta_i)$$

2. Write the decision model of attending the training (Roy model).

Under the Roy model, each individual chooses to participate in the training iff her net benefit is positive: Benefit from training $= w1 - w0 = \delta$
Cost of training $= c$
We can write the decision rule as an indicator T_i :

$$T_i = \begin{cases} 1 & if \;\; \delta_i > c \\ 0 & if \;\; \delta_i \leq c \end{cases}$$

3. What does the average wage of treated and untreated individuals measure under this decision rule?
Once individuals sort themselves according to $\delta > c$ or $\delta \leq c$, we observe the subpopulation averages:
Average wage among the treated : corresponds to the mean (post-training) wage for treated agents (with a $\delta_i > c$).

$$\begin{aligned} &\mathbb{E}\left[w_{1i}|T_i = 1\right] \\ &= \mathbb{E}\left[w_{0i} + \delta_i - c|\delta_i > c\right] \\ &= \mathbb{E}\left(w_{0i}\right) + \mathbb{E}[\delta_i|\delta_i > c] - \mathbb{E}(c) > \mathbb{E}(w_{0i} + \delta_i - c) = \mathbb{E}(w_{1i}) \end{aligned}$$

Therefore, $\mathbb{E}(w_{1i})$ will be overestimated in the population.
Average wage among the untreated : corresponds to the mean wage for non treated agents (with a $\delta \leq c$).

$$\begin{aligned} &\mathbb{E}\left[w_{0i}|T_i = 0\right] \\ &= \mathbb{E}\left[w_{0i}|\delta_i \leq c\right] \\ &= \mathbb{E}(w_{0i}) \end{aligned}$$

4. What parameter do you estimate by comparing average wages of treated and untreated individuals? Interpret that parameter.

$$
\begin{aligned}
&\mathbb{E}\left[w_1 - c|T=1\right] - \mathbb{E}\left[w_0|T=0\right] \\
&= \mathbb{E}\left[w_0 + \delta - c|\delta > c\right] - \mathbb{E}\left[w_0|\delta \leq c\right] \\
&= \mathbb{E}\left[w_0\right] + \mathbb{E}\left[\delta|\delta > c\right] - c - \mathbb{E}\left[w_0\right] \\
&= \mathbb{E}\left[\delta|\delta > c\right] - c \\
&= \mathbb{E}\left[w_1 - c|T=1\right] - \mathbb{E}\left[w_0|T=1\right] = TT
\end{aligned}
$$

This quantity is a "naïve" difference in means. It is the observed difference in outcomes between the self-selected groups (those for whom training pays off vs. those for whom it does not). It is not the average treatment effect for the whole population. It is, particularly in this case, the "treatment on the treated". But we cannot observe it directly with this formula (because that would require comparing w_1 and w_0 for the same individuals with $\delta > c$, rather than comparing to those with $\delta \leq c$).

5. Suppose a randomized controlled trial is run with groups $Z = 0$, $Z = 1$, and $Z = 2$. Compute the value of the average wage in each of these groups.

For the group $Z = 0 : Cost = c,\ T = 1$ iff $(\delta > c)$

$$\mathbb{E}(w_i|Z=0) = \mathbb{P}(\delta \leq c) \times \mathbb{E}(w_0|\delta \leq c) + \mathbb{P}(\delta > c) \times \mathbb{E}(w_1 - c|\delta > c)$$

$$
\begin{aligned}
&= (1 - p_0) \times \mathbb{E}(w_0|\delta \leq c) + p_0 \times \left[\mathbb{E}(w_0) + \mathbb{E}(\delta|\delta > c) - c\right] \\
&= \mathbb{E}(w_0) + p_0 \times \left[\mathbb{E}(\delta|\delta > c) - c\right]
\end{aligned}
$$

For the group $Z = 1 :$ No program offered, $T = 0$

$$\mathbb{E}(w_i|Z=1) = \mathbb{E}(w_0)$$

For the group $Z = 2 : Cost = c - s$, $T = 1$ iff $(\delta > c - s)$

$$\mathbb{E}(w_i|Z=2) = \mathbb{P}(\delta \leq c - s) \times \mathbb{E}(w_0|\delta \leq c - s) + \mathbb{P}(\delta > c - s) \times \mathbb{E}(w_1 - c - s|\delta > c - s)$$

$$
\begin{aligned}
&= (1 - p_2) \times \mathbb{E}(w_0|\delta \leq c) + p_2 \times \left[\mathbb{E}(w_0) + \mathbb{E}(\delta|\delta > c - s) - c + s\right] \\
&= \mathbb{E}(w_0) + p_2 \times \left[\mathbb{E}(\delta|\delta > c - s) - c + s\right]
\end{aligned}
$$

6. Show that comparing $Z = 1$ and $Z = 0$ identifies the same parameter as in the absence of the experiment.

With p_0 the estimate of trainee in group 1 :

$$\mathbb{E}(w_i|Z=0) - \mathbb{E}(w_i|Z=1)$$

$$
\begin{aligned}
&= \mathbb{E}(w_0) + p_0 \times \left[\mathbb{E}(\delta|\delta > c) - c\right] - \mathbb{E}(w_0) \\
&= p_0 \times \left[\mathbb{E}(\delta|\delta > c) - c\right]
\end{aligned}
$$

As in question 4, we are estimating : $\mathbb{E}(\delta|\delta > c)$ the average gain for those who take part in the program.

7. Explain how the independence of δ and w_0 ensures that a naive wage comparison can identify a treatment parameter. Why can't we obtain ATE?

The independence between δ_i and w_i ensures that the average wage of the untreated population is the same as that of the treated population if they hadn't been treated. Thus, we can identify the treatment effect by avoiding the selection bias. $\mathbb{E}(w_{0i}|T_i = 0)$ can be used as a counterfactual for $\mathbb{E}(w_{0i}|T_i = 1)$. We get the local treatment effect for those who self-selected into the program.

But we can't compute ATE, because we can only identify an effect on those who choose to take part to the program. As δ_i is not independent from T_i, we cannot observe $E(w_{1i}|T_i = 0)$ which is the counterfactual of the program on those who will not participate.

8. Identify the impact of the training on the population induced by subsidy s (using groups $Z = 2$ and $Z = 0$). Define the parameter and explain how to compute it.

We are looking for the Locale Average Treatment Effect. To do so, we need to identify a 'compliers' group. That is what we get thanks to the subsidy because it creates a group that take part to the program if and only if they get the subsidy.By still assuming the randomization between groups, we can write the LATE as:

This is the Wald estimator.
This is reduced form/first stage

$$LATE = \frac{E(w_i|Z_i = 2) - E(w_i|Z_i = 0)}{E(T_i|Z_i = 2) - E(T_i|Z_i = 0)}$$

We should have conditional proba, but we don't because of randomness

$$\iff LATE = \frac{E(w_{0i}) + Pr(\delta_i > c - s)E(\delta_i|\delta_i > c - s) - [E(w_{0i}) + Pr(\delta_i > c)E(\delta_i|\delta_i > c)]}{Pr(\delta_i > c - s) - Pr(\delta_i > c)}$$

$$\iff LATE = \frac{Pr(\delta_i > c - s)E(\delta_i|\delta_i > c - s) - Pr(\delta_i > c)E(\delta_i|\delta_i > c)}{Pr(\delta_i > c - s) - Pr(\delta_i > c)}$$

$$\iff LATE = \frac{Pr(c > \delta_i > c - s)E(\delta_i|c > \delta_i > c - s)}{Pr(c > \delta_i > c - s)}$$

$$\iff LATE = E(\delta_i|c > \delta_i > c - s)$$

9. What can you estimate when $s = c$?

If $s = c$, then $\delta_i > 0$ so the cost of the program is covered. Logically, all the agents in the treatment group will take part to the program (so they will all be treated). As the assignment to treatment group is still random, we can write:

$$E(w_{1i}|Z_i = 2) = E(w_{0i}) + E(T_i|Z_i = 2)E(\delta_i|\delta_i > 0)$$

$$\iff E(w_{1i}|Z_i = 2) = E(w_{0i}) + 1 \times E(\delta_i)$$

$$\iff E(w_{1i}|Z_i = 2) = E(w_{1i})$$

Therefore, if $s = c$, we could estimate the average treatment effect:

$$\iff E(w_i|Z_i = 2) - E(w_i|Z_i = 1) = E(w_{1i}) - E(w_i|Z_i = 1)$$

$$\iff E(w_i|Z_i = 2) - E(w_i|Z_i = 1) = E(w_{1i}) - E(w_{i0})$$

And the last line is the Average Treatment Effect. Therefore, under the condition that $c = s$, we can compute the ATE.

10. Suppose now that δ and w_0 are correlated with $w_0 = a+\rho\delta+\epsilon$. Show that the naive wage comparison does not identify a treatment parameter. Discuss the sign of the bias based on ρ.

$$\begin{aligned}
&\mathbb{E}(w_{1i} \mid \mathcal{T}_i = 1) - \mathbb{E}(w_{0i} \mid \mathcal{T}_i = 0) \\
&= \mathbb{E}(w_{0i} \mid \delta_i > c) + \mathbb{E}(\delta_i \mid \delta_i > c) - c - \mathbb{E}(w_{0i} \mid \delta_i \leq c) \\
&= \mathbb{E}(a + \rho\delta_i + \epsilon \mid \delta_i > c) + \mathbb{E}(\delta_i \mid \delta_i > c) - c - \mathbb{E}(a + \rho\delta_i + \epsilon \mid \delta_i \leq c) \\
&= \rho\,(\mathbb{E}(\delta_i \mid \delta_i > c) - \mathbb{E}(\delta_i \mid \delta_i \leq c)) + \mathbb{E}(\delta_i \mid \delta_i > c) - c + a - a + 0 \\
&= \rho\,(\mathbb{E}(\delta_i \mid \delta_i > c) - \mathbb{E}(\delta_i \mid \delta_i \leq c)) + \mathbb{E}(\delta_i \mid \delta_i > c) - c
\end{aligned}$$

Because $(\mathbb{E}(\delta_i \mid \delta_i > c) - \mathbb{E}(\delta_i \mid \delta_i \leq c))$ is different from 0, we can write:

$$\mathbb{E}(w_{1i} \mid \mathcal{T}_i = 1) - \mathbb{E}(w_{0i} \mid \mathcal{T}_i = 0) \neq \mathbb{E}(\delta_i \mid \delta_i > c) - c$$

Which is what we found in question 4. $\rho\,(\mathbb{E}(\delta_i \mid \delta_i > c) - \mathbb{E}(\delta_i \mid \delta_i \leq c))$ is a bias in the naive estimation. We cannot identify the treatment parameter because of this correlation. Since $\mathbb{E}(\delta_i \mid \delta_i > c) > \mathbb{E}(\delta_i \mid \delta_i \leq c)$ so the sign of the bias depends on ρ. If δ and w are positively correlated, then $\rho > 0$, so the bias will be positive and the average wage of the treated without the training is greater than the average wage of the untreated. It is the contrary if δ and w are negatively correlated.

Thus $\mathbb{E}(w_{0i} \mid \mathcal{T}_i = 1) \neq \mathbb{E}(w_{0i} \mid \mathcal{T}_i = 0)$, so the naive comparison does not identify the treatment parameter.

11. Show that comparing $Z = 1$ and $Z = 0$ still identifies the same treatment parameter.

$$\begin{aligned}\mathbb{E}(w_i \mid Z = 0) - \mathbb{E}(w_i \mid Z = 1) &= \mathbb{E}(w_{0i}) + Pr(\delta_i - c > 0)\,[\mathbb{E}(\delta_i \mid \delta_i - c > 0) - c] - \mathbb{E}(w_{0i}) \\ &= Pr(\delta_i - c > 0)\,[\mathbb{E}(\delta_i \mid \delta_i - c > 0) - c]\end{aligned}$$

$\mathbb{E}(\delta_i \mid \delta_i - c > 0)$ gives the same treatment parameter as before. The difference and these averages comes from the training because the first group has the opportunity to take it, but not the second group. That is why we are able to estimate it;

Selection on gains: self-selection based on the potential gains from the program.

Selection on level: selection based on the initial status (level) of the individuals.

SUTVA: Di(Z)=Di(Z') if Zi=Zi'. Same for outcome: Y(Z,D)=Y(Z',D') if Zi=Zi', Di=Di'. So neither outcome or treatment status can vary depending on the "level" of treatment which must be the same, and on the treatment received by others.