

# Causal Inference for Policy Evaluation

## Assignment 1

Marco Gortan, Felix Schulz, Benjamin Weggelaar

April 4, 2025

### Question 1

Table 1: Share of Age Groups by Treatment Status

Treatment Status	<i>Share</i> $\leq 40$
D=0	0.617
D=1	0.629

The differences in the distribution of age across treatment and control groups are relatively small, suggesting that the treatment and control groups are fairly balanced with respect to age. However, the treatment group appears to have a slightly younger composition on average.

Table 2: Balance Metrics

Variable	Standardized Bias	Difference in Means	P-value
Age	2.601	0.013	0.073

Here, we test the  $H_0$  that the difference in means between the treatment and control group is zero. The p-value of 0.07 is above the conventional 5% significance level. This suggests that the difference in means is not statistically significant. This indicates that the treatment and control group are balanced with respect to the age variable.

### Question 2

We leave the interpretation of the results to Q3.

### Question 3

Figure 1 shows the evolution of the Average Treatment Effect (ATE) over time for the two groups of people, below and above 40 years old. Confidence bands are at the 95% confidence level.

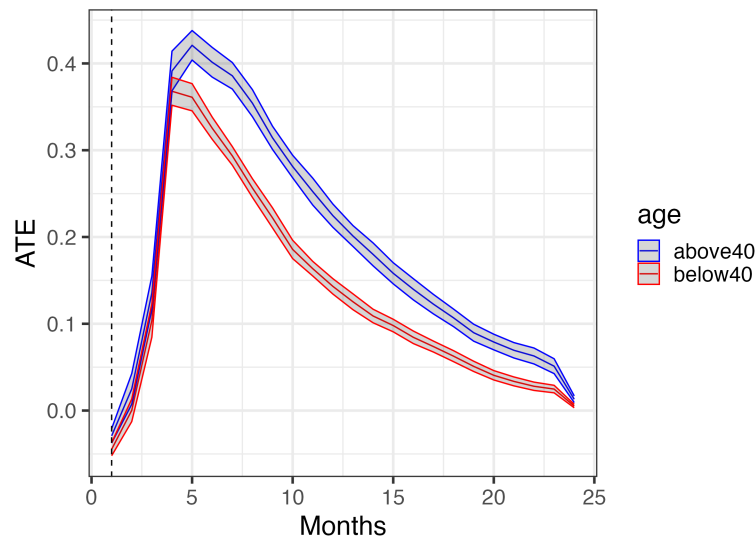


Figure 1: Program and Employment: ATEs over 24 months by age group

We notice that ATE is negative and significant one month after the implementation of the program, but it quickly turns positive, reaching a maximum level after four (five) months for the group of people below (above) 40 (50) years old. At the peak, treated people are between 35% and 45% more likely to be employment after four to five months.

The effect is, however, temporary, and decreases to zero 24 months after the implementation of the program, for both the groups. In general, the program seems to have been more successful for older people above 40.

ATET refers to the average treatment effect for those that end up treated. If the treatment group differs from the non-treated in their potential outcome, ATE and ATET are not the same. Note that the idea of the inverse probability weighting is to reweigh the groups of the treated and non-treated according to the observed characteristics, but there might be other factors that have not been taken into account, are related to the probability of ending up treated, and potentially explain the employment situation.

## Question 4

### (a) Discuss potential bias from labor market competition

The control group would have lower employment than the true counterfactual. The estimated effect would be larger than it actually is.

### (b) Identify assumption being violated

The bias arises from a violation of A1, the stable unit treatment value assumption. SUTVA defines our outcome as a function of our treatment status and potential outcomes

alone. The treatment received by the treated group should not impact the outcome of the non-treated group.

### **(c) Think of reverse bias scenario**

If treatment impacted the outcomes of the control group in the other direction, it can decrease the bias coming from the SUTVA violation explained in (a). An example would be if participant success raised overall demand for labor. Then non-participants' outcomes might improve when others are treated. If the size of this bias is the same as the initial SUTVA violation explained in (a), it would fully counter this bias.

## **Question 5**

The treatment group includes all locations that gain access to fast Internet in a given quarter or year. Specifically, these are locations situated within 500 meters of the national backbone network at the time when that backbone becomes connected to at least one submarine Internet cable.

The control group, by contrast, consists of locations that are farther than 500 meters from the backbone network and therefore are not directly exposed to fast Internet when the submarine cable arrives.

The post-treatment period starts at time period  $t$  at which country  $c$ 's backbone network is connected to at least one submarine cable. The pre-treatment period is thus the time period before the network has been connected.

It is important to note that the comparison is within countries, meaning that the control groups only include non-connected locations in the same country where some locations did receive the treatment.

## **Question 6**

### **(a) Why not individual fixed effects?**

The data is in repeated cross-sections, so the survey that are distributed in the same location in different time periods might have a different composition. Therefore, the authors control for any fixed differences between connected and unconnected parts of each grid-cell, which helps address potential selection bias in the placement of the backbone infrastructure.

### **(b) Why not location-specific time fixed effects?**

The treatment variable varies at the location level, so including location-specific time period fixed effects would absorb all the time-varying shocks at the location level, and thus absorb the variation that allows them to estimate the treatment effect.