Causal Inference: Policy Evaluation — Assignment 1

Amit Aryal, Nicholas Putney, Dominik Schawrzkopf

2025-04-04

## Question 1

## [1] "Difference in means by treatment status and standardized bias"

## E(X|D=0) E(X|D=1) Difference s.e. p-value Abs. SB  
## under\_40yo 0.617 0.629 0.013 0.007 0.073 2.601  
## Observations 10336.000 8748.000 NA NA NA NA

The dummy variable under\_40yo seems well balanced between treatment and control groups, with low standardized bias (2.6%) and an insignificant difference in means at the 5% significance level.

We should still account for age differences in our analysis due to:

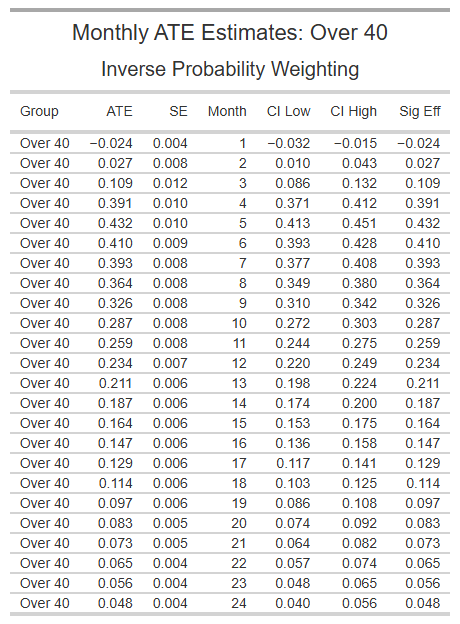
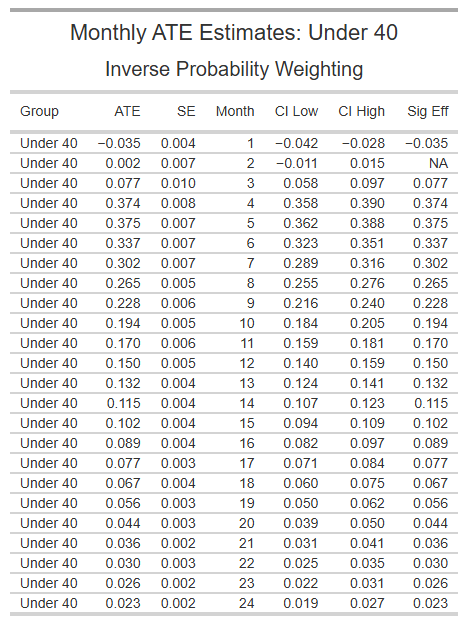
1. Heterogeneity: Age might moderate treatment effects (e.g., younger vs older participants respond differently), so including it enables subgroup or interaction analysis.
2. Robustness: Including covariates that are predictors of the outcome of interest protects against chance imbalances or unobserved heterogeneity in smaller subgroups.
3. Improved precision: if age group is a good predictor of employment status, then including it in the analysis will lead to a decrease mean squared error, leading to smaller standard errors for the ATE estimate.

## Question 2

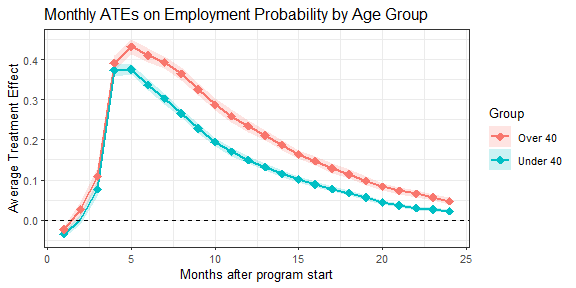
### (a) see R script

### (b) see R script

### (c)



## Question 3



The figure would suggest that for the first month, participants in both age groups have a lower probability of employment than non-participants, perhaps due to their participation in the program in that month. The program had a large impact on both groups in the first five to six months, peaking at a change in employment probability by about 40 percentage points. After month 5, the effect of the program begins to decay. However, this decay is greater in the under 40 age group, reflected in a smaller average treatment effect.

The average treatment effect (ATE) is the effect of the program had everyone been treated, even including those who were not treated. The average treatment effect on the treated (ATET) is the effect of the program only on participants. This means that the ATET estimate is only relevant to participants and would not say something about the effect on the program on non-participants. This two quantities could differ if factors related to employment differed systematically between participants and non-participants. For example, the ATET might be higher if those who actually participated in the program are more responsive to the program compared to non-participants had they participated.

## Question 4

### (a)

In the case where jobseekers who participate in the online application program are more likely to get a job but at the expense of other unemployed workers, we should not compare program participants to non-participants to estimate the ATE(T), because these estimates would be biased. We would in this case overestimate the treatment effect of the online application program.

### (b)

In this case, the Stable Unit Treatment Value Assumption (SUTVA) would be violated due to general equilibrium effects in the labor market.

### (c)

Other violations of SUTVA could be spillover effects from treated to untreated (e.g. when participants would share their gained knowledge from the program with non-participants). Here the bias would be an underestimation of the treatment effect.

## Question 5

The control and treatment groups are defined by the difference in distance between their location and the backbone network. Treatment groups are defined as individuals living less than 500m away from the network (T1), 500-1500m away (T2), 1500-2500m (T3) and 2500m-3000m (T4). The control group contains individuals located more than 3500m from the backbone network. The post-treatment period (t) is defined as at least one submarine cable has arrived in the country at time (t). The pre-treatment period as no submarine cable has arrived at time (t).

## Question 6

### (a)

The authors do not control for individual fixed effects due to the nature of the available data. DHS and Afrobarometer are cross-sectional and not longitudinal. Therefore it is not feasible to incorporate individual fixed effects.

### (b)

It would not be a good idea to control for location-specific time period fixed effects, because the treatment varies across locations over time, which is exactly the variation those fixed effects would absorb.