Causal Inference for Policy Evaluation
Assignment 2

Marco Gortan, Felix Schulz, Benjamin Weggelaar

April 18, 2025

Question 1

Table 1: Internet access and employment: Non-parametric bootstrap estimates of ATET

ATET	SD	95% CI
0.0083	0.0042	[0, 0.016]

(b) **i.** The parametrically estimated effects are identical until at least the fourth decimal position. This is to be expected since both are calculating the same two differences: In the parametric model, *Submarines* captures the trend before and after treatment across all groups, while *Connected* captures the difference between treatment and control group shared across periods. In the nonparametric approach we are explicitly calculating these two differences (first $\tilde{Y}_i = \bar{Y}_{i,1} - \bar{Y}_{i,0}$ and then $\tilde{Y}_1 - \tilde{Y}_0$). **ii.** Standard errors differ by a small amount. Asymptotically and if the OLS assumptions on error variance hold, both should be identical.

(c) These standard errors are substantially larger than the previous estimates, suggesting the presence of heteroskedasticity and/or serial correlation in the error terms.

(d) The point estimate is larger than the non-parametric estimate. The fixed effect approach used here uses the within estimator, focusing on within-location variation, additionally accounting for temporal fixed effects. The remaining variation after demeaning significantly differs and thereby also the coefficients. We expect to obtain different coefficients as, in the original specification, we were not taking into account unobservable differences among locations (that differ in their initial employment situation) and the common effects that certain year-quarters have on the employment level. The low, insignificant, pooled estimate of *Treatment* might reflect that many treated locations were starting from low employment levels. When we eliminate this initial heterogeneity, we can properly estimate the effect.

Question 2

(a) Yes, it is a good idea, because we expect skill to be correlated with both treatment, as skilled people might decide to go live close to the grid, and the outcome, as they artificially increase the employment level. Note that if *Skilled* was a time-invariant characteristic of each location, than it would not be needed to be included, as its effect would be absorbed by location fixed effect. On the other hand, if *Treatment* increases exogenously Skill and Employment, it might be a 'bad' control, since it will cause a downward bias in our response.

(b) The estimate is lower compared to point 1d, suggesting that skilled people go to live close to newly connected grids. Indeed, we see that the coefficient of *Skilled* is positive and significant, absorbing part of the effect that was present in point 1d.

Table 2: Parametric ATET estimates

	(1b)	(1c)	(1d)	(2)
Intercept	0.7192^{***}	0.7192^{***}		
	(0.0012)	(0.0038)		
Connected	0.0478^{***}	0.0478^{***}		
	(0.0031)	(0.0100)		
Submarines	-0.0400^{***}	-0.0400^{***}		
	(0.0019)	(0.0039)		
Treatment	0.0083	0.0083	0.0217^{**}	0.0149^{*}
	(0.0051)	(0.0101)	(0.0079)	(0.0059)
Skilled				0.5695^{***}
				(0.0038)
Fixed Effects			✓	✓
Clustered SEs		✓	✓	✓
Num. obs.	280641	280641	280641	280641
Num. groups: time			10	10
Num. groups: location			3169	3169

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Question 3

(a) The first thing we observe is that *Connected* locations have a higher share of educated people. One potential explanation is that living closer to the grid already allowed faster internet, attracting higher educated people. Moreover, we see that not connected locations are witnessing downward trend in the share of highly educated, possibly due to the fact educated people might migrate more to other areas / countries with job opportunities more similar to their education. Finally, we observe an increase of highly educated people in connected areas after fast internet arrived, possibly due to an increase in job opportunities.

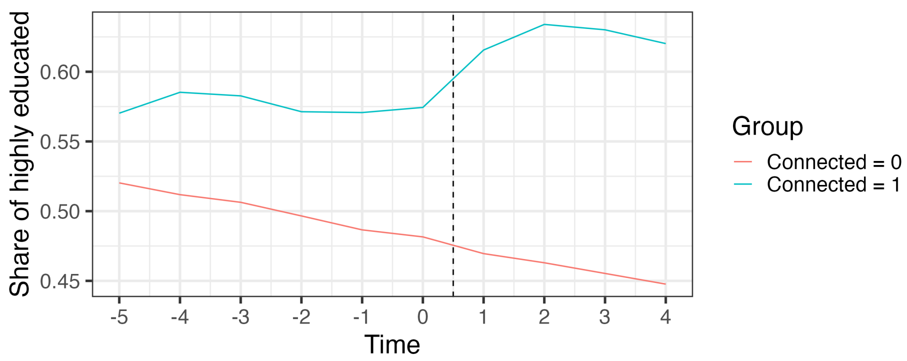

Figure 1: Common trends for education

(b) No. Already prior to treatment, education in those locations further away from the internet network showed a vastly different trend than those that would be treated later. Therefore we cannot confidently assume parallel trends after treatment.

(c) Our estimate on employment would be upward biased because we observe a (i) downward trend in not connected locations and a (ii) a flat trend in connected locations before fast internet arrived.

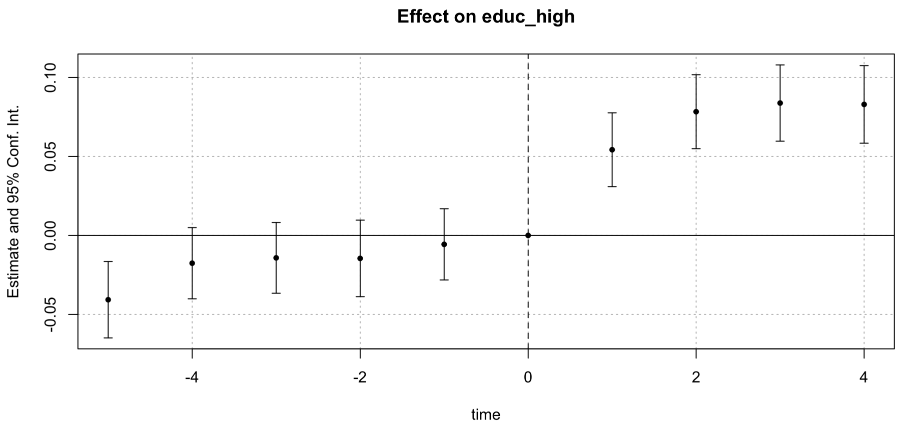

Figure 2: Event study: education

Question 4

(a) Yes, they are similar to the analytical standard errors with unit-clustering from 1(c).

Table 3: Internet access and employment: Non-parametric bootstrap estimates of ATET, Cluster-robust SEs

ATET	SD	95% CI
0.0083	0.009	[-0.009, 0.026]

Question 5

(a) In this particular example, regressing female labor supply on the number of children would lead to endogeneity. This is because deciding on the number of children may be influenced by unobserved family factors such as their preferences for leisure over work, access to child care, or family support. This violates the exogeneity or zero conditional mean error assumption that OLS relies on for unbiased and consistent estimates.

(b) The estimated treatment effect is a local average treatment effect (LATE), also known as the complier average causal effect (CACE), because the authors identify the effect of having a third child on labor supply, for those parents whose decision of having a third child is affected by the sex composition of the first two children (same sex). Therefore, the estimated effect is specific to that group.

(c) OLS over-estimates the causal effect because it is unable to isolate the effect of an additional child on labor supply, since it doesn't take into account unobserved factors such as preferences for leisure over work, access to child care, or family support. The OLS estimates are more negative compared to the IV estimates, possibly due to the fact it is more common for women with less child care support to have more than two children (for instance, women with big career prospectives, who probably have access to child care support beyond family, might usually have less children).