

Psreg: a Stata routine for estimating the blocking with regression adjustment

Martina Bazzoli (FBK-IRVAPP)

Silvia De Poli (JRC- Seville)

Daniela Piazzalunga (Università di Trento)

XVII Conferenza Italiana degli Utenti di Stata

Motivation

The starting point is the usual linear regression of the outcome Y on the binary treatment status D , controlling for the observable confounders X :

$$Y_i = \alpha_0 + \alpha_1 D_i + \alpha_2 X_i + \varepsilon_i$$

The distribution of X is highly imbalanced across treated and controls



Average causal effect could be badly biased

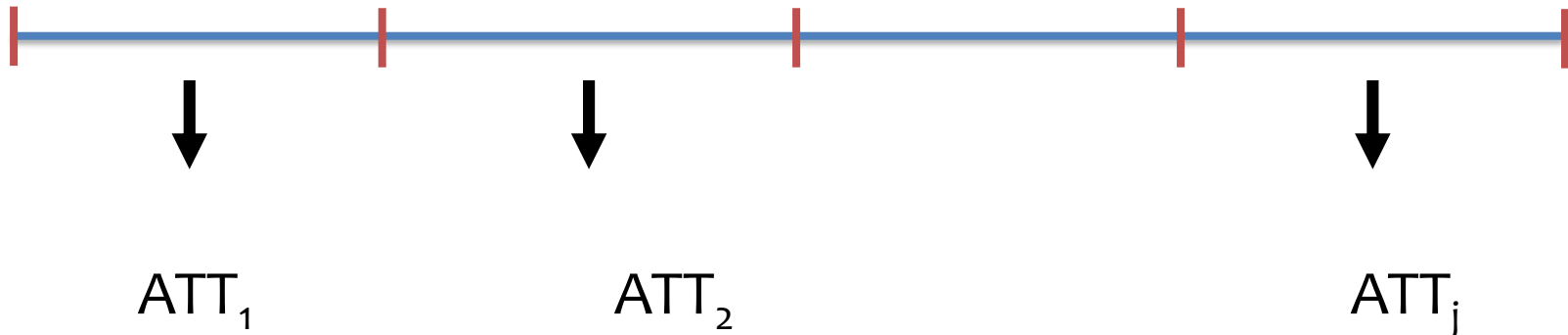
Blocking with regression adjustments estimator

Blocking with regression adjustments estimator, proposed by Imbens (J. Human Resources, 2015).

This estimator relies on the use of the propensity score (Rosenbaum and Rubin, 1983) and computes the average treatment effect on treated (ATT) by performing linear regressions within blocks (i.e. intervals) of the propensity scores, controlling for the observable characteristics.

Blocking with regression adjustments estimator

- Estimate propensity score $p(x)$
- Partitioning the pscore $[0,1]$ into J intervals (blocks)
- Treatment effect estimated by OLS within each block with covariates X and treatment D



ATT

Weighted average of the block specific estimates:

$$ATT = \sum_j w_j ATT_j \quad \text{Var}\{ATT\} = \sum_j w_j^2 \text{var}\{ATT_j\}$$

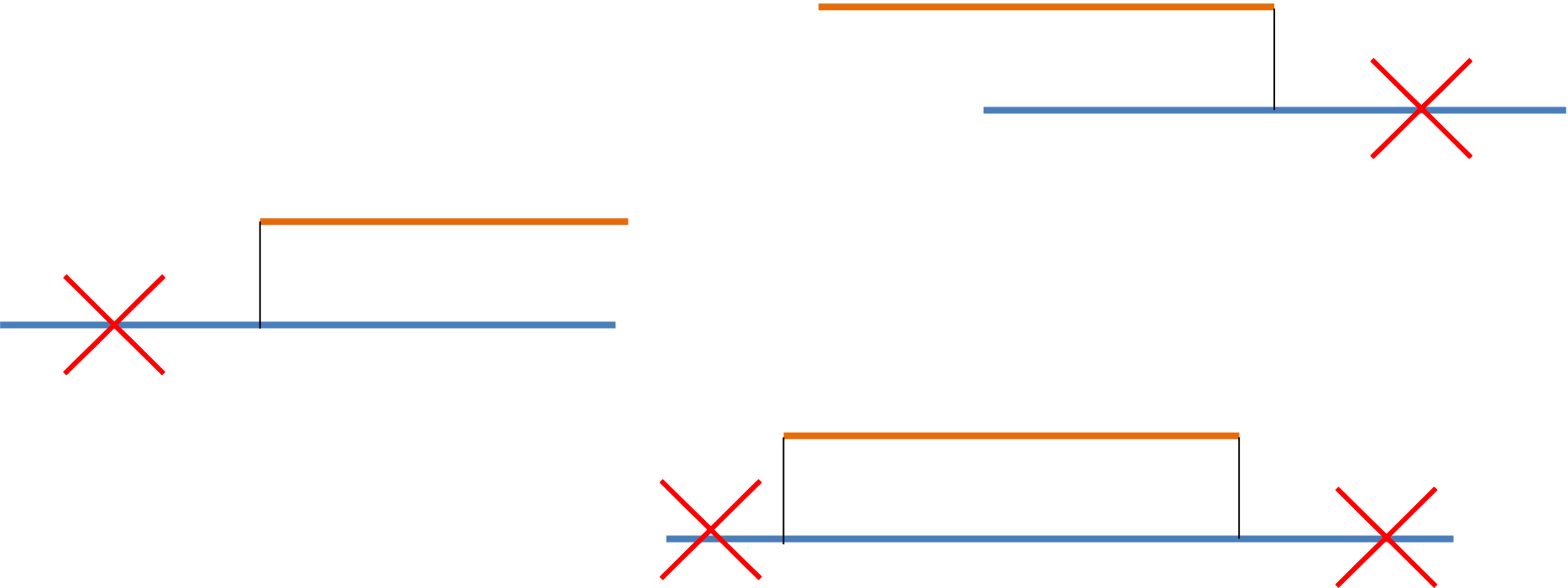
w_j the ratio of the number of treated units in the j -th block to the total number of treated units.

Number of blocks

- Fixed number of blocks with equal number of treated
- Algorithm for selecting the number of blocks (Imbens and Rubin, 2015)

Common Support

Treated Controls



Psreg

psreg depvar treatment indepvars

logit uses logit instead of the default probit to estimate the propensity score.

ate computes the ATE instead of the default ATT.

common restricts analysis to the common support dropping treated observations whose propensity score is higher than the maximum or lower than the minimum propensity score of the untreated observations. In the case of ATE it also drops untreated observations whose propensity score is higher than the maximum or lower than the minimum propensity score of the treated.

groups defines a specific number of blocks, with an equal number of units in each block. The number of units refers to the number of treated in the case of ATT and to the number of observations (treated and untreated) in the case of ATE. The default is the optimal number of blocks according to the algorithm presented by Imbens (2015).

quietly does not print the regressions within each block

The training programmes

In 2010, the *Agenziadel Lavoro*(AL) of Trentino organized 30 training courses for unemployed people.

- 64 courses
- 854 participants
- Average duration: 100 days – 410 hours

Programmes were designed to train their participants to perform these jobs

- construction supervisor, computer network technician, web technician
- electricians, carpenters, bakers
- accountants, office clerks

Research question

Did the training courses increase the participants' chances of employment in the months following attendance?

Data

administrative archives

- information on individual training courses
- registers of the local *Centri per l'Impiego* (CPI, the Italian public employment agencies delivering placement services to the unemployed).
- the *Comunicazioni Obbligatorie* (COB) archive, i.e. the archive of firms' mandatory communications to the CPI.

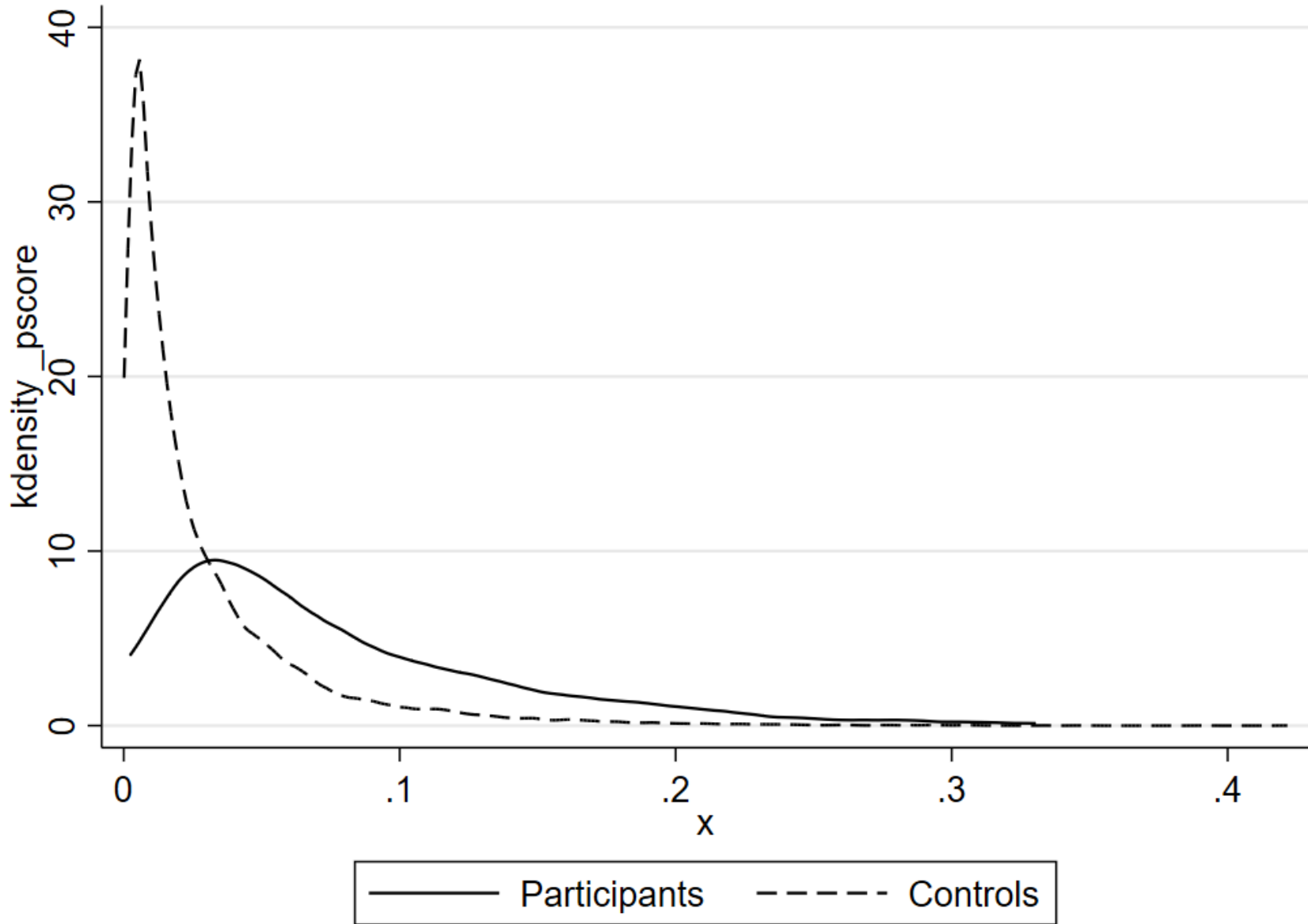
Methodology

- **blocking with regression adjustment**

Comparing

- **Treated:** participants in the training courses
- **Controls:** individuals recorded as unemployed at the starting date of individual training courses in the registers of the local Centri per l'Impiego

Distribution of the propensity score



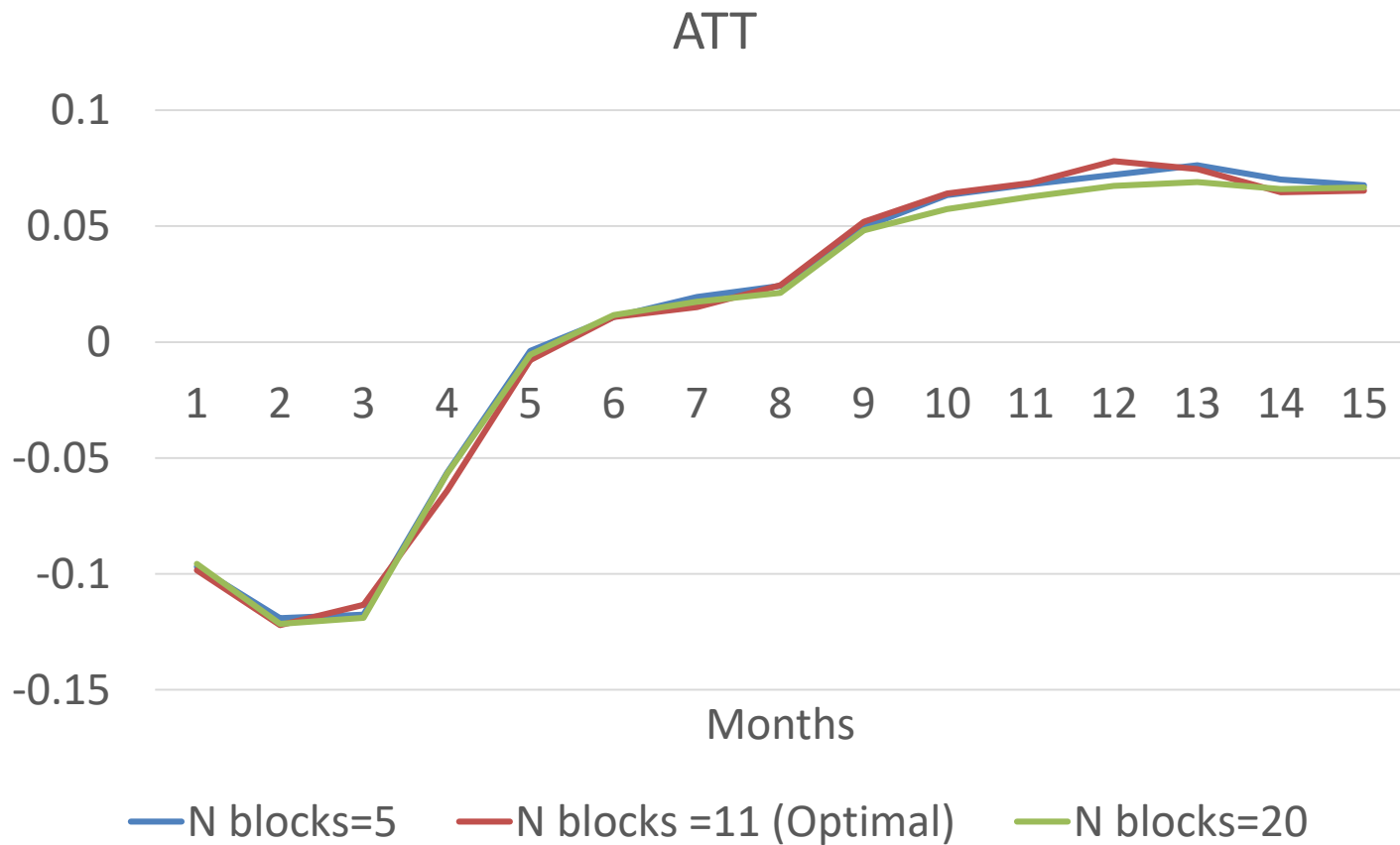
Average propensity score in each block

Block	Number of Treated	Propensity score	
		Treated	Controls
1	13	0.0034	0.0029
2	13	0.0066	0.0064
3	26	0.0099	0.0096
4	53	0.0154	0.0149
5	105	0.0251	0.0247
6	211	0.0429	0.0422
7	105	0.0659	0.0652
8	106	0.0895	0.0893
9	105	0.1252	0.1233
10	53	0.1680	0.1679
11	53	0.2369	0.2365
Total		0.0772	0.0290

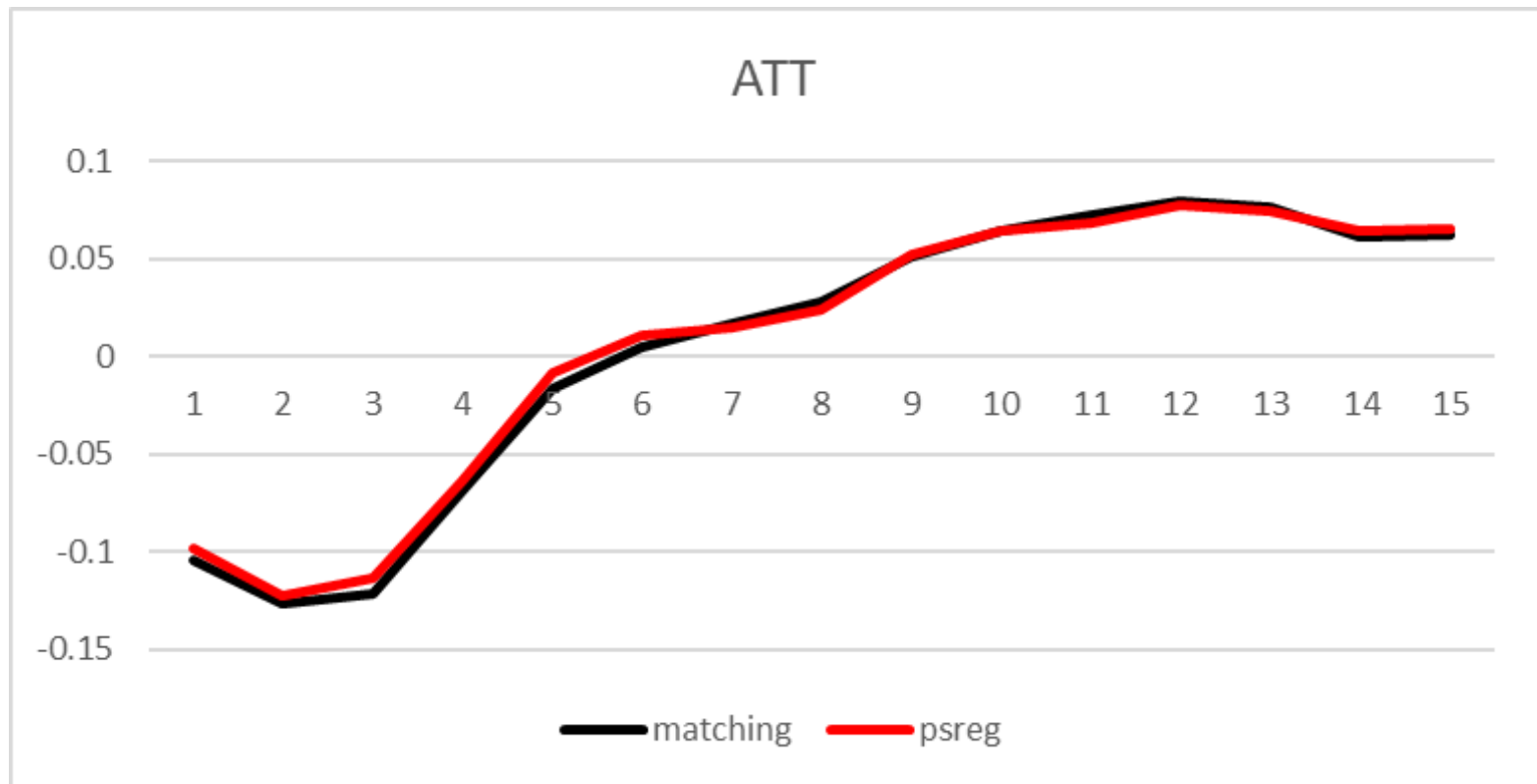
ATT

Month	Att	Se	
1	-0.098	0.011	***
2	-0.122	0.014	***
3	-0.113	0.015	***
4	-0.064	0.016	***
5	-0.008	0.016	
6	0.011	0.016	
7	0.015	0.016	
8	0.025	0.017	
9	0.052	0.017	***
10	0.064	0.017	***
11	0.069	0.017	***
12	0.078	0.017	***
13	0.075	0.017	***
14	0.065	0.017	***
15	0.065	0.017	***

ATT using different number of blocks

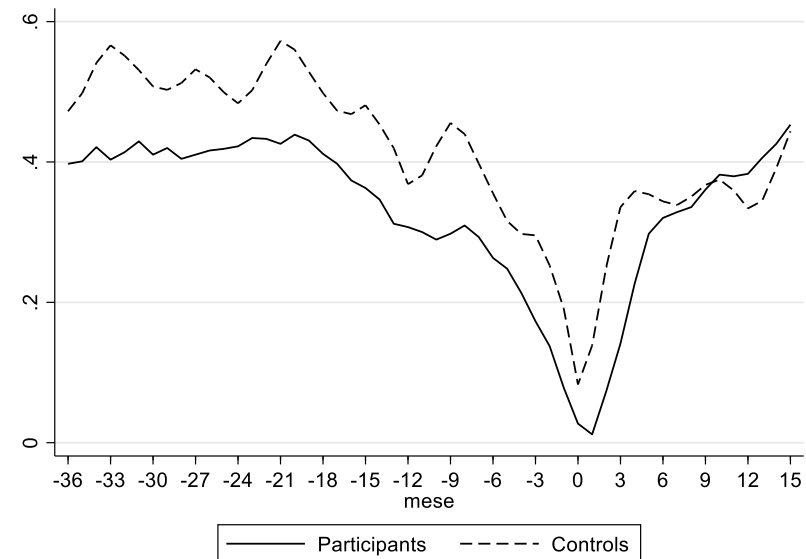


Matching vs Blocking

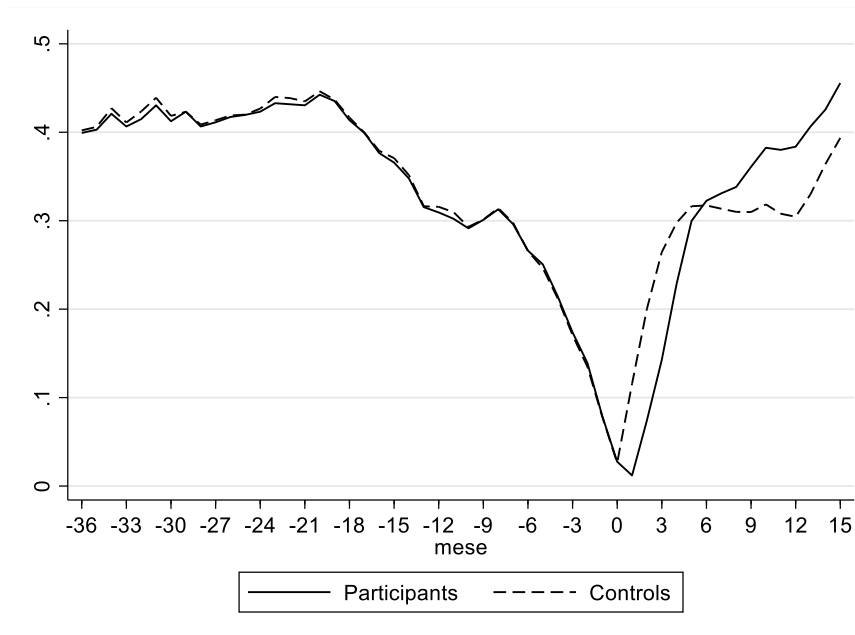


Probability of being employed for the treated and the controls

Before the matching



After the matching



Future work

- Take into account weights
- Show ATT for a subgroup (e.g. gender) as interaction between treatment and the group
- Saving the data of each blocks

Thanks for the attention