# 58920-01 Causal Inference for Policy Evaluation

## Assignment

Please hand in one single PDF per group via email to conny.wunsch@unibas.ch. The submission deadline is **Tuesday, 15 July 2025, 7 pm**. Please be precise and concise in your answers and explain all notation, equations, figures and tables you include. There is no minimum or maximum number of pages. Submit whatever you think is necessary. Contact me by email if you have any questions regarding the assignment.

## Background

We consider a training program for unemployed workers in Switzerland. The program is offered by the regional placement offices (RPOs) and targets workers' basic computer and technical skills to improve their employment prospects, in light of the digitization of the economy. The program lasts 3 months, and workers can only enroll once within any month of their unemployment spell. The program was implemented in 2013 and is only offered to workers older than 45 entering unemployment after this date (i.e. workers turning 45 during the spell, or workers older than 45 but already unemployed at the time of the reform are not eligible). Prior discussions about this policy change have remained confidential, so that neither caseworkers nor the unemployed had knowledge about it before this date. The program is not mandatory, and workers can choose to participate at any time after its implementation, given that they satisfy the eligibility criteria. From discussions with caseworkers, we know that certain characteristics may affect program participation, such as the worker's demographic characteristics, education, characteristics of last job, as well as other factors that may affect employment. We are interested in evaluating the effect of the program on individual labour market outcomes.

## Data

You have representative data on unemployment spells for workers in Switzerland who entered unemployment during the years 2011 and 2015. The data include the following variables.

| | |
|---|---|
| Id | Personal identifier (the same individual can have multiple spells) |
| Date_start | Date of start of unemployment spell |
| Date_end | Date of end of unemployment spell |
| Sex | Sex -- 0 male, 1 female |
| Age | Age at the beginning of the unemployment spell |
| Marits | Marital status – 0 not married, 1 married |
| Canton | Canton of residence – 26 cantons |
| Region | Region of residence – 7 great regions, as defined by the Swiss Federal Office of Statistics |
| Nationality | Nationality – 1 Swiss, 2 EU/EFTA, 3 Other |
| Mother_tongue | Mother tongue – 0 German, 1 French, 2 Italian, 3 English, 4 other |
| Educ | Education – 0 school, 1 apprenticeship or matura, 2 specialized or university |
| Insured_earn | Earnings insured by unemployment insurance |
| Lastj_occpt | Occupation of the last pre-unemployment job (23 categories) |
| Lastj_fct | Function in the last pre-unemployment job – 0 self-employed, 1 managerial role, 2 specialist role, 3 learning or other |
| Lastj_rate | % of full-time job (activity rate) in the last pre-unemployment job |
| Child_subsidies | 1 receives subsidies for dependent children, 0 otherwise |
| Contr_2y | Months of contribution (employment) in the last 2 years prior to unemployment |
| Unempl_r | Unemployment rate (in %, canton-level) |
| GDP_gr | GDP growth (in %, canton-level) |
| Treat | Month in unemployment spell when treatment started (0 when not treated) |

**IMPORTANT:** You are not allowed to use the data for any other purpose than the assignment and you are not allowed to share the data with anyone outside your group.

## Difference-in-differences (DiD)

You want to exploit the introduction of the program to estimate its effects based on DiD. As outcome variables you want to use unemployment duration and the probability of being employed 12 months after entering unemployment.

1. What would be the treatment group, and what would be the control group? Explain. [4 points]
2. Generate the two outcome variables of interest and provide descriptive statistics for all four relevant groups (treated pre/post, controls pre/post). Discuss the results. [4 points]
3. Compare mean observed characteristics for all four relevant groups (treated pre/post, controls pre/post). Discuss the results. [6 points]
4. Which assumptions do you need to identify the effect of interest based on DiD in this setup? [5 points]
5. Plot average outcomes for treatment and control group for each month in the observation period. Discuss the results. [6 points]
6. Discuss the validity of the identifying assumptions in this specific case. Provide and discuss supporting evidence if possible, incl. event study estimates. [8 points]
7. Implement the DiD using OLS. Describe in detail what you do and why. Discuss whether or not you need additional control variables and if so, why. Discuss the results. [8 points]
8. Implement the DiD using a semi-parametric estimator based on the propensity score. Describe in detail what you do and why. Discuss the results and compare them to the OLS estimates. [10 points]
9. Discuss different ways of exploiting the panel structure of the data and implement one of them. Discuss to which extent the panel approach relaxes assumptions underlying the estimation approaches you used before. Describe in detail what you do and why. Discuss the results and compare them the other estimates. [10 points]
10. Check for heterogeneous program effects with respect to observed characteristics for which you suspect different effects. Describe in detail what you do and why. Discuss the results. [5 points]

**GOOD LUCK!**