

## Econometrics II - Workgroup 5

### Problem 1

A university introduces a housing program which offers student housing to second-year students who in their first year have a grade point average (GPA) exceeding 8.0. A researcher applies a difference-in-differences model to estimate the effect of this housing program on academic achievement. The treatment group contains students with a first-year GPA between 8.0 and 9.0, the control group includes students with a first-year GPA between 7.0 and 8.0. The researcher uses as outcomes the first-year GPA (before treatment) and the second-year GPA (after treatment) of the students. Does this approach give a consistent estimator for the causal effect of the housing program? Briefly explain your answer.

### Problem 2

Even though not everyone likes to think about it in that way, the marriage market is a market like all markets, where demand and supply can have an influence on market outcomes. What is often seen as an important determinant of marriage market outcomes is the sex ratio, the ratio of men to women in the population. The prediction would be that the position of men in the marriage market improves with a reduction in the sex ratio (and vice versa). In this exercise we look at the impact of male scarcity on out-of-wedlock births. Such births could reflect men having more partners or men being able to shift the cost of child rearing to single mothers.

The dataset `marriagemarket.dta` contains data on the percentage of out-of-wedlock births in France in 91 departments, both before and after World War I (the variable `illeg`). The variable `post` indicates whether this percentage is measured before or after the war, `mortality` is the military mortality rate during the war and `sr` is the sex ratio ( $\text{\#males 18-59} / \text{\#females 15-49}$ ). Finally, `depc` is the department code for the 91 departments.

- (i) Regress the number of out-of-wedlock births on the sex ratio, using only the observations from the pre-war period. Discuss your result. How can a difference-in-differences approach using the military mortality rate during WWI improve on this estimation strategy?

- (ii) Generate a dummy variable that indicates whether the military mortality in a region is above the median military mortality or not. Make a table with the mean percentage of out-of-wedlock births for the high and low mortality regions, both before and after the war. Use the numbers from the table to calculate the difference-in-differences estimator.

The variable mortality is actually a continuous variable. In a regression we will therefore use it as a continuous treatment variable.

- (iii) Estimate the following model, which estimates the difference-in-differences estimator in a regression equation. What is the interpretation of the coefficients  $\beta_1$  and  $\beta_2$ ? What do you conclude about the effect of male scarcity on the number of out-of-wedlock births?

$$\text{illeg} = \beta_0 + \beta_1 \text{post} \times \text{mortality} + \beta_2 \text{post} + u$$

*In order to run this regression you will first have to create a new variable that is equal to  $\text{post} \times \text{mortality}$*

- (iv) Run the same regression but now include dummies for all the departments. Discuss your results. Do you prefer this estimation over the estimation of question iii)? Why?
- (v) What is the key assumption when you apply difference-in-differences? What would be a way to investigate the plausibility of this assumption? Why is that not possible with this dataset?