Econometrics II - Workgroup 3

Due date: Sunday, January 22, 11.59 pm. Hand in your solutions as a **single .pdf file** including your code via Canvas. Include your R (or any other language) code by using R Markdown (preferred) or by using the package "minted" in your .tex file (see a template on Canvas). Each team has to come up with a unique name (without names or student numbers). Both teammates have to submit solutions via Canvas. Each subquestion has an equal number of points.

Problem 1: Treatment effects

In a population, each individual is uniquely characterized by its color. There are three types of individuals, purple, blue and green individuals. A researcher wants to know how a treatment affects outcomes. The table below shows for each color how many individuals are in the treatment and control group and what the average outcomes are for individuals in the treatment and control group.

color	number of individual		average outcome	
	treated	control	treated	control
purple	100	100	9	7
blue	75	25	13	8
green	25	75	10	9

(*i*) Compute for each of the colors the treatment effect.

(*ii*) Compute the average treatment effect in the full population.

(*iii*) Compute the average treatment effect on the treated.

(*iv*) Give an example where the average treatment effect on the treated would be more useful to consider than the average treatment effect, and explain why.

Problem 2: Financial rewards to students

Bachelor students often have difficulties completing all first-year courses. This is costly. When failing too many courses the study gets delayed, because students cannot start second-year courses. This might imply that students enter the labor

market later or even drop out of university. There are also costs to the university. Students retake exams, which need to be graded or sit in classes a second time. More crowded classrooms have negative consequences for fellow students. Therefore, universities use different interventions for stimulating students to improve their study achievements.

Economists usually first think about financial incentives when intending to change behavior. In this assignment we consider a policy intervention in which students were promised financial rewards for good study achievements. The data are collected from a randomized experiment performed among first-year economics students. At the start of the academic year all students were randomized into three groups: a control group, a low-reward group and a high-reward group. The students in the low-reward group could earn 500 guilders (about 227 euro) when completing all first-year courses before the start of the next academic year. For the high-reward group, this was 1500 guilders (about 681 euro).

The data set contains the following variables: bonus500 and bonus1500 indicate if a student has been assigned to the low-reward and the high rewards group, respectively; pass indicates if the student completed all first-year courses before the start of the second academic year; stp2001 gives the total number of credit points collected in the first year (maximum 60) and stp2004 describes the total number of credit points after three academic years; myeduc and fyeduc describe the years of education of the mother and father, respectively; p0 is the subjective assessment of the student about his/her success probability at the start of the study; math contains the high-school math score; job indicates if the student had a job while studying; dropout indicates if the student dropped out of the economics study; and effort describes the average number of hours the students devoted to studying.

(i) Compute the fraction of students in all three groups (control, low-reward and high-reward) that complete all first-year courses before the start of the second academic year. Show within a table that background characteristics are balanced over the treatment groups.

(ii) Use the linear probability model to regress the dummy variable for completing all courses on the assignment of the three treatment groups. Interpret the treatment effects. Next include as additional regressors father's education, high-school math score and the subjective assessment about the pass probability.

(*iii*) Next also include as regressors in your model whether a student has a job and the amount of study effort. Comment on this approach. Do you consider this an improvement over (*ii*)?

(*iv*) Use your preferred model specification to estimate the effects of the financial incentives on some other outcomes: dropping out and credit points collected (in the first year and after three years).

Prior to conducting the experiment, the researcher performed some power calculations.

(*v*) Given the sample size and the estimates you have obtained above, what would be the minimum detectable effect size of this experiment?

(*vi*) Initially, the researchers were aiming at an increases in the pass rate of 10%-points. How large should the sample size of the experiment have been in that case?

(*vii*) Students assigned to the control group provide counterfactuals to both the high-reward and the low-reward treatment. Show that by assigning more students to the control group than to each treatment group the same MDE can be achieved with fewer students in the experiment.