Problem Set 1.2

Labour Economics, Winter Term 2025/26

*Submit by Sunday, 16 November, 22:45h **on Moodle!***

Learning objectives

- Conduct instrumental variables (IV) regressions to address OVB.
- Conduct IV regressions to address measurement error.
- Interpret experimental data and compare results to OLS.

Tasks

Download the data `ps1_more_realistic_data.Rda` and open it in R Studio. Reality is often messier than what we saw in *Problem Set 1.1*. We commonly have variables measured with error, such as ln_wage_ME or $hours\_ME$, and lack information on important factors ($motivation$). We will use the randomly assigned wage increases $wage\_premium$ to try to address these issues. As said before, these could be thought of as individuals being drafted into an income support program like the SSP in British Columbia / New Brunswick (Canada) that we discussed in lecture.

a) Repeat part of the descriptive statistics (those you find informative) from questions a)–c) in *Problem Set 1.1*. What differences can you notice in the data?

b) Run again the regression of $hours$ on ln_wages. What is the difference now that you only have measurement-error versions of these variables?

c) Before using the wage premium as a quasi-experiment, what check would you like to do? Does the data pass this check?

d) Do a balancedness check on $motivation$ using the Data from PS1.1. Is this in line with your check from c)?

e) Use the wage premium as an instrumental variable for log wages. What do you see; interpret also the first-stage regression.

f) From b) and e), what are the differences to question d) in *Problem Set 1.1*? Is this what you would expect (*hint:* consider also our lecture notes on IV)?

g) Add *Education* as a control variable to the IV estimation. Do you obtain substantively different results and why (not)?

h) Now compare your IV results to the full regression f) from *Problem Set 1.1*. What is your preferred estimate of the elasticity of labour supply in this data?

Notes: You can work in teams of 1–3 students. Please upload your code as well as a pdf-file with discussions on what you found in the data in response to the tasks above. It should be clear which lines of code and answers in the `.pdf` refer to which question.