Causal Data Analysis

Study questions

1. The government wants to raise the employment probability of long-term unemployed workers. If a firm hires such a worker, it will pay his/her wage for the first one year. You have data on the population of adult individuals for 5 years, including observations both before and after the policy introduction. With the panel data, how would you estimate the effect of the subsidy on the employment of long-term unemployed to decrease any selection bias? Who would you choose as a control group? Write down the estimation equation and interpret the coefficient. What condition you need to have a causal estimate? Can you think of any spillovers which may contaminate the estimation?
2. Demonstrate mathematically, that a fixed effects regression estimate is not biased by variables that do not vary at the subject level, even if they are correlated with both the outcome and the treatment variables.
3. You want to analyze the effects of industrial robots on firm efficiency. You have panel data.
    a. Write down the estimation equation for fixed effects estimation. Interpret the coefficient of interest (as precisely as possible).
    b. Why is it important to control for the aggregate trend Yt?
    c. Under what condition would the coefficient be the ATET?
    d. Explain why is the coefficient estimated by fixed-effects probably closer to the causal effect than a coefficient estimated by simple OLS.
    e. Provide an example when the coefficient will be biased (your example should be specific to the problem discussed).
4. Instrumental variables
    a. Write down the conditions for a variable z being an instrument for variable x (dependent variable is y)?
    b. 2. Give the intuition for why an instrumental variable can provide causal estimates. Give an example.
    c. 3. How can you estimate a relationship with an instrumental variable?
5. You want to estimate the effects of firm subsidies on firm productivity with panel data, but you realize that many unsubsidized firms are very different from the subsidized firms. Describe how matching and panel data methods can be combined to obtain a good estimate of the treatment.