

Training for the exam

Exercise 1 : Measuring the efficiency of policies encouraging older workers' training

A series of measures of the "Plan Concerté pour l'emploi des seniors" enacted in 2009 in France aimed at stimulating employment among workers aged 50 and over. One measure consisted in encouraging employers to train their older workers. The question is : what is the effect on older workers' training on their probability of being employed in the same firm next year ?

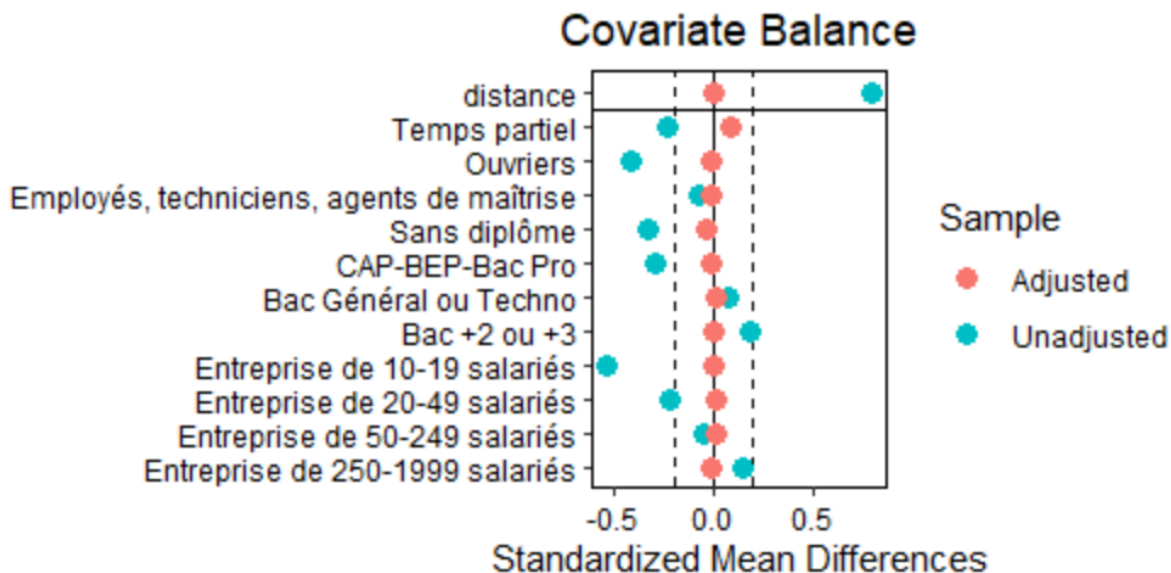
For this study we have French data that contain information on training rates, occupational and educational level of workers aged 50-54 and on the size of the firm in which they are employed in 2015. The outcome is a binary variable Y_i equal to one if a worker i initially employed in 2015 is still employed in the same firm in 2016. In addition, D_i is a dummy variable indicating whether a worker i participated in a training session in 2015.

- a. Using the notations of the potential outcomes framework, write the expression of Average Treatment Effect on the Treated (ATT). In this setting, what is the treatment ? What is the counterfactual ? What is the main issue to estimate the ATT ?
- b. Write a naive estimator to estimate the ATT. Is it an unbiased estimator ? Why ?
- c. Write the equation of a OLS model regressing the dependent variable Y_i on the treatment D_i adjusting for a vector X_i of individual characteristics.
- d. Instead of a OLS regression, we estimate this effect using a propensity-score based nearest-neighbour matching method. What does the propensity score mean in this setting ?
- e. To estimate this model with R we indicate the following options :
 - `replace=TRUE`
 - `caliper=0.25`
 - `ratio=2`

What do these options mean ? For each line, explain why is it recommended to use these options.

- f. Figure 1 shows the standardized mean differences for each individual characteristic between the treated and the non-treated group of individuals before and after matching. We also have plotted two vertical lines at -0.2 and 0.2. Is the balancing property verified in that case ? Give an example of one variable that strongly differs across both groups.

FIGURE 1 – Testing the balancing property



- g. At last we find an Average effect of the Treatment on the Treated of +4 percentage points. What does it mean ?
- h. What is the key assumption in our model ? Is this assumption credible in our case ? Which method should we use to address this issue ?

Exercise 2 : Estimating the effect of childbearing on labour supply

You try to estimate the effect of childbearing on mothers' labour supply using an instrumental variable strategy. More precisely, your explanatory variable of interest

D_i is a binary variable indicating whether the mother i has had a third child. The dependent variable Y_i is a dummy variable equal to one if the mother i is employed. your strategy consists in instrumenting D_i by two variables : a dummy $Z_{i1} = 1$ is the mother has already had two boys and a dummy $Z_{i2} = 1$ is the mother has already had two girls.

- a. Why OLS estimate from a regression of Y_i on D_i and a set of exogenous covariates X_i would be biased ?
- b. What are the two properties for the instruments to be valid ?
- c. After summarizing the corresponding 2SLS estimate with R, the diagnosis tests are the following

TABLE 1 – Diagnosis tests after 2sls estimate of the effect of childbearing on mothers' labour supply

	df1	df2	statistic	p-value
Weak instruments	2	329 495	228.74	0.00001
Sargan	1	NA	0.333	0.5641

What is the meaning of each test ? What can you conclude from the results obtained in the first line ? And from those obtained in the second line ?

- d. You find that having already two children of the same sex increases by 0.06 (or 6 percentage points) the probability of having a third child. In addition you find that employment rate among women who had their two first children of the same sex is equal to 0.52 (52%) while the employment rate among women who had their two first children of different sex is 0.528 (52.8%). Compute the Wald estimate of the causal effect of having a third child on mothers's employment rate.

Exercise 3 : Evaluation of a job training program through a Randomized Control Trial

You want to measure the impact of a job training program that focused on low-income youths with less than a secondary education. The main outcome of your study is the employment rate. A random sample of eligible youths was selected to undergo training. However, among those assigned to the program, 20% do not effectively participate in training sessions. Among those not assigned to the program, no one has participated in this training program.

- a. What is the main advantage of a Randomized Control Trial to evaluate this kind of program ?
- b. After explaining these terms and under the assumption of monotonicity, give the proportion of always-takers, never-takers and compliers in that case. Justify your answers.
- c. You observe that one year after the end of the training program, 50% of youths in the control group and 60% of youths that participated in training are employed. In addition, 53% of youths that have been assigned to the program are employed one year after the end of the program. What is the Average effect of this Treatment on the Treated (ATT) here ?
- d. This causal effect is measured under the Stable Unit Treatment Value Assumption (SUTVA). Explain the meaning of this assumption. Why would it mean if this assumption was not verified ?

Exercise 4 : Evaluation of minimum wage policy on employment in US

Card and Krueger (1994) investigated the effect of a rise in the minimum wage in US on employment at fast-food restaurants, since these restaurant are big minimum-wage employers. They exploit the fact on April 1 1992, New Jersey raised the state minimum wage from 4.25\$ to 5.05\$ while this minimum wage has remained stable (4.25\$) in Pennsylvania. They collected data on average full time employment in fast-food restaurants in these two states in February 1992 and in Novembre 1992. They show the followig descriptive statistics :

TABLE 2 – Evolution of average full time employment in fast-food restaurants in New Jersey and in Pennsylvania

	Pennsylvania	New Jersey
February 1992	23.33	20.44
November 1992	21.17	21.03

- a. Evaluate the causal effect of a rise in the minimum wage on fast-food restaurants' employment using a difference-in-differences method.
- b. What is the key assumption for this measure to be a valid estimator of the causal effect of a rise in minimum wage on employment ?
- c. Suggest a test to convince the reader that this assumption is credible in this setting.