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. **Y_{1i} denotes the 2016 employment status if the individual i had participated in training in 2015 and Y_{0i} denotes the same variable if the individual i had not participated in training in 2015. The treatment here is the participation in a training session in 2015. The ATT is the average causal effect of the treatment on the treatment group. The ATT writes as follows :**

$$ATT = E(Y_{1i} - Y_{0i}|D_i = 1)$$

The counterfactual here is the employment status of the treated individuals if they had not participated in training in 2015. This counterfactual has the following expression :

$$E(Y_{0i}|D_i = 1)$$

The main issue to estimate the ATT is that we do not observe at the same time Y_{1i} and Y_{0i} for the same individual.

b. **One naive estimator to estimate the ATT would be the difference between employment rates among treated individuals and employ-**

ment rates among control ones. This naive estimator would write as :

$$E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 0)$$

This estimator suffers from a selection bias : we could guess that individuals that participated in training in 2015 have very different characteristics than the non-participants and that these differences may affect their employment status one year later.

c. **The corresponding OLS model writes as :**

$$Y_i = \alpha + \gamma D_i + \beta X_i + \epsilon_i$$

where α is the intercept, ϵ_i is the error term assuming that $E(\epsilon_i) = 0$. The ATT would be measured by γ under the exogeneity assumption.

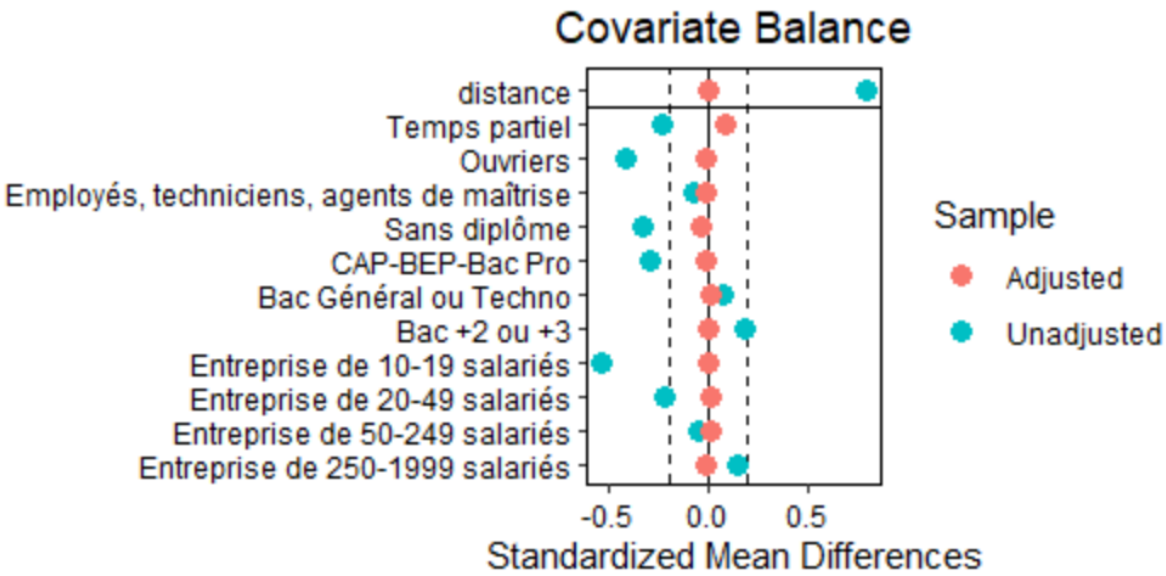
d. **The propensity score is the conditional probability of participating in training in 2015 given a set of covariates X_i.**

e. **The option "replace" means that we apply a matching procedure with replacement. This allows to choose several times the same control individual as nearest neighbour. It is recommended because, without replacement, it could be that some individuals in the control group matched with the treated group may have a too different value from their propensity score.**

The option "caliper=0.25" indicates that we apply a caliper matching : we consider that if the propensity score difference between the nearest neighbour in the control group and the individual in the treatment group exceeds 0.25 standard deviations of the pscore value, we do not match this individual (we discard him from our matched sample). This is recommended because it allows to guarantee good quality matches, i.e. very similar individuals in the control group matched with treated ones.

The option "ratio=2" means that we apply a 2-nearest neighbour matching approach. That means that for each treated individual, we match him with the 2 closest individuals in the control group in terms of propensity score value. This allows to increase the precision of our estimates.

FIGURE 1 – Testing the balancing property



f. Figure 1 shows clearly that the balancing property is not verified in that case. Indeed, standardized mean differences between treated and control group exceed 0.2 in absolute value for many variables. In particular, we observe that the proportion of treated individuals in small firms is strongly smaller that the same proportion computed in the control group (standardized mean difference strongly negative).

g. This means that participation in training increases the probability for a worker aged 50-54 to stay in the same firm the next year by 4 percentage points.

h. The key assumption is the Conditional Independence Assumption : The treatment is independent from the outcome given a set of observable covariates. In that case it is not credible because of a reverse causality bias. We could guess that individuals participate in training BECAUSE they plan to stay in the same firm the next year. In that case, the main variable of interest is endogenous and this issue can not been solved with a matching approach.

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. **OLS estimate will be biased because it could be that fertility decisions are caused by mothers' employment status. In that case, there is a reverse causality bias that raises endogenity issue. The variable indicating whether the woman has three children is not independent from her employment status.**

b. **To be valid, an instrument has to check two conditions. First, it has to be strongly correlated with the endogenous variable. In that case, the fact of having first two children of the same sex has to influence strongly the decision of having a third child. This is the first-stage condition. Second, the instrument has to be exogenous, i.e. independent from employment status (not correlated with the error ter). Here, we could guess that having two boys or two girls do not result from a choice : this is the chance. So, the instrument is randomly assigned and we could say that it is quite independent from employment status.**

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

c. **The first test consists in checking whether the excluded instruments used are strongly correlated with the endogenous variable. This corresponds to a F-test comparing an unrestricted regression where the endogenous variable is regressed on all excluded and included instruments and a restricted one in which the endogenous variable is regressed only on included instruments. This measures**

the joint significance of these tw excluded instruments in the first-stage. The F-test statistic, also referred to as the Cragg-Donald statistic is 228.74 and the p-value equals 0.00001. We conclude that we reject H0 : instruments are not jointly significant at 0.001% level. The instruments are not weakly correlated but strongly correlated with endogenous variable. The first-stage condition is verified.

The second test is possible since there are more excluded instruments than endogenous variables. We are in a over-identified case so we can do a Sargan test : regressing the residual (error term) from the second-stage equation on the set of instruments and multiplying the R^2 of this regression by the number of observations. The null hypothesis H0 is that instruments are valid that is that the residual (error term) are not correlated with excluded and included instruments. The p-value of this test is 0.5641 so we can not reject H0 even at a 10% level. This shows that instruments are valid in that case.

d. **The Wald estimate consists in dividing the reduced-form effect, that is the effect of the instrument on the dependent variable by the first-stage effect, that is the effect of the instrument on the endogenous variable. Here we know that the first-stage effect is 0.06. We observe that the employment rate difference between women who first two children of the same sex and those who had their two first children of different sex is $0.52 - 0.528 = -0.008$. So the Wald estimate is $-0.008/0.06 \approx_{-} 0.1333$. This shows that having a third child reduces by 13.33 percentage points the women'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. Since the program is randomly assigned, we can be sure that it is independent from the outcome variable. There is no selection bias so we can use the naive estimator mentioned in the exercise 1.

b. The always-takers are individuals that always participate in the training program even if they have not been assigned to thus program. In that case, there are no always-takers. The never-takers are individuals who never participate in the training program even if they have been assigned to this program. We know that among those assigned to the program, 20% do not effectively participate, so the proportion of never-takers is 20%. The compliers are those who participate in the program only if they have been assigned to it. In that case, the proportion of compliers equals $1 - 0.2 = 0.8$**.**

c. Since there are never-takers, you have a non-compliance issue that you have to address using an 2SLS method rather than a simple OLS one. The 2SLS estimate is obtained computing the Wald estimator. The reduced-form effect or Intention-To-Treat (ITT) effect, i.e. the effect of treatment assignment on employment rate one year after the end of the program is $53\% - 50\% = 0.03$**. The effect of treatment assignment on effective participation in the program is 0.8 (the proportion of compliers). So the corresponding Wald estimator is** $0.03/0.8 = 0.0375$**. This program has increased by 3.75 percentage points the employment rates of participants.**

d. SUTVA implies that the ATT does not depend on the proportion of treated individuals. This assumes that there are no externalities of the treatment on non-treated individuals outcome. In the case of job training programs, it is not credible : we could guess that increasing the number of participants will be detrimental on the probability for non participants of finding a job.

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. **Before and after the rise in the minimum wage the number of jobs (in full time equivalent) in fast-food restaurants in New Jersey has increased from 20.44 jobs to 21.03 jobs. The increase is** $21.03 - 20.44 = 0.59$. **Over the same period, the number of full-time equivalent jobs in fast-food in Pennsylvania has decreased from 23.33 to 21.17, that is an evolution of -2.16. The difference-in-differences method consists in comparing the evolution of the number of jobs before and after the rise in minimum wage in concerned state (New Jersey) to the same evolution observed in not concerned state (Pennsylvania). This estimator would be :** $0.59 - (-2.16) = 2.75$. **This suggests that the rise in the minimum wage would have increased the number of full time equivalent jobs in fast-foods by 2.75.**
   Note that a before-after estimation is not relevant in this setting. Indeed, some macro shocks could have affected the state of New Jersey before and after the change in the minimum wage.

b. **The key assumption here is the common (parallel) trend assumption : the number of jobs in fast-foods would have evolved in the same way if the New Jersey had not been affected by a rise in the minimum wage.**

c. **To be more convincing, we could do a placebo test. We could compare the evolution in the number of jobs between 1990 and 1991 (or in previous years) in both states, to test whether these trends were parallel before the change.**