THEMA Working Paper n°2022-18
CY Cergy Paris Université, France

The impact of university openings on labor market outcomes

Samia FERHAT

Septembre 2022

The impact of university openings on labor market outcomes*

Samia FERHAT[†]

September 28, 2022

Abstract

In this paper, I study the impact of university openings on labor market outcomes. I focus on university openings that occurred in France in the 90's, and exploit five waves from representative samples of young individuals who left the French education system, starting from wave 1992. I use difference-in-differences estimation techniques, and find that the impact of university openings on labor market outcomes is heterogeneous according to the characteristics of the region where the opening occurs. I find that opening a new university increases the probability of being employed by about 8% points and increases wages by 5% in regions characterized by a lower level of education and a more disadvantaged socio-economical background. In contrast, no impact is found in regions where the unemployment rate is low and where the population is highly educated.

Keywords: Human capital, university openings, labor market outcomes

JEL Codes: I23, I26, J21, J23, J24

*I am grateful to Thomas Brodaty, Laurence Jacquet, Vincent Vandenberghe, Fabian Gouret and Maelys de la Rupelle for their advice and comments. I acknowledge the financial support of Labex MMEDII (ANR11-LBX-0023-01). I also thank CASD for data access.

[†]Thema, CY Cergy Paris University, 33 boulevard du Port, 95000 Cergy. Email: samia.ferhat@cyu.fr.

1 Introduction

In contemporary developed societies, individuals and governments alike are striving to improve labor market outcomes. It is generally considered that a more educated workforce would be less prone to unemployment and would have access to better-paying jobs. In economics, this idea is grounded in the human capital theory (HCT), first coined by Becker (1962), and according to which education may lead to wages increase through an increase in individual productivity. Since universities are producers of human capital, the opening of a new university in a given region may lead to improved labor market conditions in the area. The idea of this paper is to build on existing evidence of university openings on human capital investments in order to measure the causal effect of university openings on labor market outcomes using microeconomic data.

Siegler (2012) is the first to show a causal effect of opening a new university on human capital. For Germany, he finds that the probability to graduate for a young rises by 8 to 10 percentage points when a new university is established in her county between 1960 and 1979. Brodaty et al. (2019) also find a causal impact between university openings and human capital investments. They use individual data from CEREQ *Généneration* surveys on the case of university openings that occurred in France between 1991-1993. They find that opening a new university increases the probability of attaining at least two years of higher education by about 10 percentage points, and the probability of attaining at least four years of higher education by about 5 percentage points. These findings support my argument for the human capital channel through which university openings have an impact on labor market outcomes.

There is a strand of literature that focuses on the economic impacts of universities at the macroeconomic level. Beeson and Montgomery (1990) advocate that universities raise productivity growth through increasing earnings and chances of employability. Cantoni and Yuchtman (2014) argue that universities had an important role in the commercial revolution through the development of legal institutions. Gagnol and Héraud (2001) study the case of Strasbourg university in France and find direct repercussions on companies through human capital improvement. More recently, Valero and Van Reenen (2019) use data of university openings between 1950-2010 for

1500 regions across 78 countries and find that a 10% increase in a region's number of universities per capita is associated with 0.4% higher future GDP per capita in that region. This important role of universities on the economy raises the question on the role they have from a microeconomic perspective. In this paper, I provide answers by focusing on the causal effects of opening a new university on individuals' wages, employability and employment stability in the labor market (labor market outcomes hereafter).The treatment that I consider consists of a series of university openings that occurred in France between 1991-1993. My outcome data come from Céreq Génération Survey and the French version of the EU-Labour Force Survey. The identification rests on a difference- in-difference setup.

Even though there is an extensive literature on returns to education, starting with Becker (1993) and Mincer (1974), [1] these studies focus on the causality between years of schooling and returns to schooling. An interesting aspect in the literature that interests us is its use of College/University proximity as source of exogenous variation of schooling. For instance, Card (1995) uses the variation of the College/University proximity as an instrument for schooling. He exploits the National Longitudinal Survey (NLS) for men and shows that men who grew up in regions with a 4-year college nearby have significantly higher schooling and significantly higher earnings. Other examples of studies that use college proximity as an instrument for schooling are: Kling (2001) and Moretti (2004). In this paper, I also explore this potentially causal link between College/University and schooling and labour market outcomes. What differentiates my approach is twofold. First, I consider the impact of an explicit change (i.e. increase) of proximity that corresponds to university opening in some French counties in the early 1990s. Methodologically, this implies that, instead of using college proximity as IV, I adopt a treatment effect perspective as I am interested in the impact of binary variable (presence or not of a university) on outcome variable of policy interest (i.e. individuals labour market outcomes). More specifically, I implement a difference in differences treatment analysis. Second, I posit (and allow in my specifications) for the impact on the treatment to operate via two channels: i) the traditional "human capital" chan-

[1] See Psacharopoulos and Patrinos (2018) for details on returns to schooling estimation

nel that underpins most of the empirical work done since Mincer (1974), where labour market outcomes gains strictly correspond to educational/human capital gains achieved by those youth who benefited from the easier access to university, and ii) the "spillover" channels resulting from higher companies productivity related to the presence of a higher share of college graduates, and from local externalities related to university research via collaborations with firms (see Anselin et al. (1997), Kantor and Whalley (2014) and Monjon and Waelbroeck (2003)). These knowledge spillovers affect local labor market structure and hence individuals employment and wages. Other potential explanation of the spillover channel is the Altonji and Pierret (2001) Employer Learning-Statistical Discrimination (EL-SD) assumption. The idea is that employers hire workers on the basis of observable characteristics such as educational attainment since it is correlated with their unobserved real productivity. Over time the real productivity gets revealed to employers through work performance, hence, if the statistical discrimination is performed in the early stage of hiring based on observed characteristics, those characteristics become less important for wages as workers reveal their type with time. In this case of university openings, employers do not know the quality of education compared to well established pre-existing universities (université de la Sorbonne, Lille...), employers could discriminate based on the name of the university delivering the degree. Bordón and Braga (2020) show that when university prestige is used to signal workers unobserved productivity, the wage premium is at 13% for college graduates in the first year of labor market. This wage premium drops to 4% after 6 or more years of work experience.

To the best of my knowledge, my paper is the first to focus on university opening per se as an exogenous treatment, synonymous with enhanced proximity for local residents and quantify it's impact on their labor market outcomes. A notable exception is the paper by Berlingieri et al. (2017) who attempt to measure the impact of opening colleges on local labor markets, using the case of college openings in west Germany in 1968. They find that high skilled employment increases by about 12% eight to nine years after the college opening. Another interesting finding is that the wage of these highly educated individuals did not suffer from the rise of the number of tertiary educated individuals entering the labour market . In other words, the labour demand was able to absorb the additional human capital.

The main finding of my paper is that the impact of university openings on labor market outcomes is positive in areas characterised by low supply of higher education, a low initial level of education and a high level of unemployment rate compared to the rest of the country. Opening a new university in such areas increases employment by about 8 percentage points and wages by 5%. I show that the total effect of university openings on labor market outcomes can be disentangled into an impact that goes through human capital formation and an impact that results from spillover effects. I show that the contribution of human capital in the total effect is always positive and significant. However, the effect that is related to spillovers is positive, negative or null depending on the counties where universities opened. The latter finding calls for further investigation on the nature of the potential spillovers discussed in this paper and their magnitude (local firms externalities and EL-SD).

The rest of the paper is organized as follows. Section 2 provides the background of this study and presents the data. Section 3 presents the empirical framework. Section 4 gives the main results. Section 5 provides some robustness checks and I finally conclude in section6.

2 Context and Data

2.1 University openings

I study the case of university openings that occurred in France between 1991-1993, as a part of *Université 2000* plan. The plan suggested by the minister of education "Lionel Jospin" at the time and was officially decided upon on May 23, 1990 by the French Council of Ministers. The main objectives of *Université 2000* were to meet the increasing demand on higher education due to an increase in population growth, and to rebalance university facilities across regions in France and thus contributing to better regional planning. [2].

[2]source: Délégation à l'aménagement du territoire et à l'action régionale (DATAR) (1998)

Four universities were created by decree law of 22nd of July, 1991, in Ile-de-France region (i.e., the Parisian region).These four universities are called *Universités des villes nouvelles* (UVN) because they were implemented in the four new cities *"Villes Nouvelles"* (Évry, Cergy-Pontoise, Saint-Quentin-en-Yvelines and Marne-La-Vallée) that were created by the French government in 1965 to decongest Paris and permit a multipolar urban development. See the Appendix for details on each university creation.

I consider these 7 openings to be my source of exogenous shock to the supply of high skilled labor. One might think that opening a new university is not random and that the government decides to establish universities in a place where it wants to develop economic activity. To tackle this issue, I control for the regional economic activity for both treated and control counties.

2.2 Data

I use information from two different datasets: CEREQ *Génération* surveys and *Les Enquêtes Emploi*. I describe in this subsection both data sources and define the main variables.

2.2.1 CEREQ *Génération* surveys

CEREQ *Génération* surveys are large scale surveys realized by the CEREQ[3] on waves starting from 1992. Each wave is representative of individuals who left the education system between January 1st and December 31st the year of the wave. I exploit 5 waves: 1992, 1998, 2001, 2004, and 2007. A total number of 154,225 individuals were interviewed in the five cross sections considered. For each wave and for every individual I use four sets of information: (i) First, on the last degree attained at the date the individual left the education system, (ii) individuals' situation on the labor market three years after leaving the education system (i.e., whether they are employed or not, whether they have a permanent contract or work as officials and their wage in case they

[3]CEREQ is a research agency working under the aegis of the Ministry of Education, link to website: `http://www.cereq.fr`

are employed), (iii) a set of individual characteristics (e.g, gender, if a grade was repeated before entering 6th grade,[4] the individual's father situation on the labor market at the date the individual left the education system, and whether her parents were born in France or abroad).

(iv) the place of residence at the date the individual entered in middle school. I use this piece of information to assign individuals into treated and control groups in the difference-in-differences set up (see Figure 1). An individual is considered to be treated if she lived at the date she entered in middle school in a county where a university opened between 1991-1993. An individual is considered to be in a control group if there were no university openings between 1991-1993 in the county where she lived at the date she entered middle school. The place of residence at the age of entry in middle school is used in contrast to current place of residence or the place of residence at the date the individual graduates high school, because parents' decision to live in a given county at the age their child is 11 or 12 years old is unlikely to be influenced by the treatment (university openings).

I consider wave 1992 to be the wave before university openings. Most universities were opened at the end of 1991 or 1992 (except for university of la Rochelle that was opened in 1993) and since students enrolled in 1991 would not obtain their higher education degree until at least 1993, as the minimum number of years required to obtain a university degree in France is 2 years. Year 1992 is legitimate to be considered as the year before the treatment in my difference in differences estimation strategy. Indeed, waves 1998, 2001, 2004 and 2007 are the waves after the treatment.

2.2.2 *Les Enquêtes Emploi*

Les Enquêtes Emploi realized by INSEE [5] are the equivalent of Labor Force Surveys on the European Union level. [6] I use 5 waves from *Les Enquêtes Emploi* (1995, 2001,and surveys from the 1st

[4] 6th grade corresponds to the first year of middle school in the French education system, if the pupil doesn't repeat any grade her age should be 11 to 12 on this year

[5] INSEE is the general directorate of the ministry of Economy and Finance. Its mission is to collect, analyze and disseminate information on the French economy and society throughout its territory.
Link to website `https://www.insee.fr/fr/accueil`

[6] Before 2003, these surveys were realized on a yearly basis, and then to have a standardization across the European Union, starting from the 1st of January 2003 the surveys were henceforth realized quarterly. The aim of these surveys is to have detailed measures of labor force activity and unemployment in a way that the population census

quarters of years 2004, 2007 and 2010).[7] By using these waves I control for labor market structure in treated and control counties, 3 years after individuals have left the education system. I define labor market structure in each county by the share of unemployed individuals (for those aged 35 and above to avoid endogeneity) and the share of executives and intermediate professions.

3 Empirical Approach

As a first step, I describe the following function in which labor market outcomes (Y) are explained by schooling (S) and university openings (UO):

$$Y = f(S, UO) = \rho + \beta \times S + \gamma_1 \times UO \tag{1}$$

Second, in line with the "human capital" channel idea, I assume that schooling can be directly impacted by the university openings.

$$S = g(UO) = \lambda + \gamma_2 \times UO \tag{2}$$

Replacing (2) in (1), I obtain:

$$Y = f(g(UO), UO) = \eta + \alpha \times UO \tag{3}$$

where $\alpha = \gamma_2 \beta + \gamma_1$

In other words, the total impact of university openings (α) consists of their impact on schooling (the human capital channel) ($\gamma_2 \times \beta$) plus what I will refer to hereafter the "spillover" channel (γ_1).

is unable to provide (Goux (2003)).

[7] 1st quarters are used for comparability reasons, as the yearly surveys till 2002 were conducted in March of each year (i.e., in the 1st quarter).

3.1 Total effect of university openings

I propose identifying the total impact (α) by estimating the following difference-in-difference model:

$$\begin{aligned} Y_{i,c,t} &= \kappa + \varphi \times Treat_c + \delta \times Post_t + \alpha \times (Treat_c \times Post_t) + \\ &\quad \zeta \times X_{i,c,t} + \mu \times Z_{c,t} + \phi \times D_{2004} + \vartheta \times D_{2007} + \xi_{i,c,t} \end{aligned} \tag{4}$$

$Y_{i,c,t}$ is the outcome 3 years after leaving the education system for individual i who lived in county c at the date she entered middle school. $Treat_c$ is a dummy for individuals in treated counties, i.e., the counties where a university opening occurred, $Post_t$ is a time dummy that switches on for individuals in waves (1998, 2001, 2004 or 2007). $Z_{c,t}$ are county time varying variables: unemployment rate for those who are aged 35 and above, [8]the share of executives and intermediate professions and the number of births 18 years earlier. D_{2004} and D_{2007} are time dummies equal to 1 for individuals who left the education system in 2004 and 2007 respectively and 0 otherwise, in order to account for the reform in the French education system.[9] $X_{i,c,t}$ is a set of individual characteristics (individual's gender, if she repeated at least one grade in primary school, whether the individual's father was active on the labor market at the date the individual left the education system, the last socio-professional category of the individual's father at the date the individual left the education system, and if her both parents were born in France, abroad, or one of them in France and the other abroad). $\xi_{i,c,t}$ is a random error term.[10]

I consider 3 main left-hand side variables to measure labor market outcomes. First, a dummy that takes 1 if the individual is employed and 0 otherwise. Second, the log of monthly wages for full time contractors, wages are deflated using the consumer price index with 2010 as the base

[8]I consider 35 years and above to make sure that the unemployment rate is exogenous and not explained by university openings.

[9]Between 2004-2007, the French universities suppressed progressively the DEUG (i.e., the two year university degree), the Maîtrise (i.e., the four-year university degree), the DESS and the DEA, and implemented the Licence Master Doctorat system (i.e., Bachelor's Master's Doctorate system of the Bologna process)

[10]Contrary to Mincer (1958, 1974) I do not include experience since the potential experience for individuals leaving the education system in a given year is the same (3 years).

year. Finally, a dummy that takes 1 if the individual has a stable employment, i.e. if she has a permanent contract or is a civil servant.

Note that, I do not restrict the data to those who have a university degree so that I can measure the overall effect that university openings have, including on the labour market outcomes of those who have not attended university.

3.2 Total effect decomposition

Estimating Equation (4) gives the total impact of university openings on labor market outcomes: on one side, it consists of the impact that is mediated through human capital formation ($\gamma_2 \times \beta$) and on the other side of the additional impact (γ_1), to which I refer as a spillover effect (i.e. additional jobs for those who did not attend university, consumption increase translating into higher wages...). One may want to decompose the two, a first step is to estimate an augmented DID model that includes, EDUC a measure of educational attainment.

$$\begin{aligned} Y_{i,c,t} &= \kappa' + \varphi' \times Treat_c + \delta' \times Post_t + \gamma_1 \times (Treat_c \times Post_t) + \beta \times Educ \\ &\quad + \zeta' \times X_{i,c,t} + \mu' \times Z_{c,t} + \phi' \times D_{2004} + \vartheta' \times D_{2007} + \xi'_{i,c,t} \end{aligned} \tag{5}$$

The “new” (purged) coefficient of the DID interaction term γ_1 provides a good indication of how much the overall impact of the reform is reduced when conditioning out changes in educational attainment. The estimated coefficient γ_1 tells us about the magnitude of the spillover effect. And β is the Mincerian coefficient of the effect of schooling on labor market outcomes. Gelbach (2016) shows that the difference between α (from Equation (4)) and γ_1 informs on the magnitude of the contribution of educational attainment. The key idea here is the well-known formula of the omitted variable bias. When estimating the first model (without educational attainment EDUC: eq (4)) the value of the estimated $\hat{\alpha}$ deviates from the "true" $\hat{\gamma}_1$ (i.e. which is here simply the DID coefficient purged from the contribution of EDUC in eq (5)) according to:

$$\hat{\alpha} = \hat{\gamma}_1 + \hat{\gamma}_2 \times \hat{\beta} \tag{6}$$

The second term to the right of the Equation represents the bias due to the omission of EDUC. Note also that the term premultiplying β is just the OLS-estimated partial correlation γ_2 obtained when regressing EDUC on the other regressors present in eq (4):

$$\begin{aligned} Educ_{i,c,t} &= \kappa'' + \varphi'' \times Treat_c + \delta \times Post_t + \gamma_2 \times (Treat_c \times Post_t) + \\ &\quad \zeta'' \times X_{i,c,t} + \mu'' \times Z_{c,t} + \phi'' \times D_{2004} + \vartheta'' \times D_{2007} + \xi''_{i,c,t} \end{aligned} \tag{7}$$

It is thus the DID estimate of the impact of university openings on educational attainment. In other words, the difference between the two estimated coefficients $(\hat{\alpha} - \hat{\gamma_1})$ is equal to $\hat{\beta}$ (the relationship between labour market outcome Y and educational attainment), weighted by the DiD coefficient $\hat{\gamma_2}$ (the one that captures the impact of university openings on educational attainment).

Note that, this decomposition is valid under the hypothesis of educational attainment exogeneity, that I relax later on using instrumental variables approach as a robustness check.

4 Results

In Table 1 are reported the main results for the estimates of α in Equation (4). The four columns depict results for separate estimations where treated individuals are those who lived at the age of entry in middle school in Charente-Maritime (first column), Nord and Pas-de-Calais (second column), in Pas-de-Calais only (third column) and in Villes-Nouvelles (last column). Control individuals are those who lived at the age of entry in middle school in neighbouring counties to those where a university opening occurred (see Figure 1 for details). I show separate results for Pas-de-Calais since both universities, Artois university and the university of the Littoral Opal Coast, had their main sites in Pas-de-Calais. The Nord county had only one site (Dunkerque site of the Littoral Opal Coast university) and in this county the university of Lille already existed with a high supply of higher education. Therefore the treatment is less relevant in the Nord county so I provide results when the Nord county is and is not included.

Table 1 shows that the opening of a new university in Charente-Maritime increases the probability to be employed by about 8% points and decreases this probability by about 4% points in Pas-de-Calais. Employment stability increases by 6% points as a result of university openings in Charente-Maritime and decreases by 5% points in Villes-Nouvelles. Finally, wages increase by about 5% in Charente-Maritime.

Before explaining the previous results, Table 2 provides some descriptive statistics to illustrate the particularity of the counties surrounding Paris in comparison to the rest of France. We can see from Table 2 that even before university openings occurred, Villes-Nouvelles had a high share of individuals whose father is or was executive. The population in Villes-Nouvelles is highly educated compared to the other counties where university openings occurred and is also above the average of the rest of France. The unemployment rate is also quite low in the region. Brodaty et al. (2019) show that opening a new university increases the probability to attain at least 2 years of higher education by about 10% points, nonetheless results in Table 1 tell us that this increase in human capital accumulation did not translate into higher employment in Nord and Pas-de-Calais. Over education and skill downgrading hypotheses could be plausible. For instance, Sloane et al. (2020) show that over education lead to lower wages. Valletta (2018) argues that skill downgrading is part of the explanation for the decrease in higher education wage premiums starting year 2000. The idea is that the high skilled with advanced tertiary education degrees replace the lower skilled workers due to a lower demand for advanced cognitive tasks in the workplace (one of the consequences of information technology revolution). Other potential explanation of the negative effect of university openings on employment is the signalling hypothesis. The new universities in Nord-Pas-de-Calais and Villes Nouvelles opened in areas not far from existing and already well known universities by employers (exp: Université de la Sorbonne in Paris and Université de Lille in Nord). When employers are faced with two candidates with the same level of education, they might statistically discriminate based on the name of university that is expected to be correlated with the unobserved productivity.

Moreover, another potential explanation for the negative impact on employment in Pas-de-Calais and Villes-Nouvelles is the fields of studies. Individuals might have studied a field that led them to deteriorate their labor market outcomes. To test this hypothesis, I estimate Equation (4), to which I add controls for the field of the last degree prepared or obtained at the date individuals have left the education system. [11] The fields or majors are grouped into the following categories: Business and Administration, Literature-Languages-Social Sciences, Science & technology, Environmental Sciences, Social & Health, Humanities & Art, Law & Political Sciences and finally Services (for vocational degrees mainly). Table 3 shows results. The negative impact on unemployment decreases and becomes almost non significant when controlling for the field of studies. Therefore, the choice of major could be responsible for decreasing the chances for those who lived in Pas-de-Calais at the date of entry in middle school to find a job.

Furthermore, an interesting question to ask would be whether university openings have an impact on the type of employment. Do these openings give better chances to get high skilled jobs? For this, I look at the probability for an individual who found a job to be working as an executive, to work as an employee and finally as a blue collar (the least qualified job). Table 4 shows the results. I find that opening a new university increases the probability for those who lived in Charente-Maritime to find a job as an executive by about 6% points. This previous impact is negative in Nord and Pas-de-Calais at the level of 2% points. Moreover, university openings seem to decrease the probability to work as a blue collar in Nord and Pas-de-Calais by about 3% points. Results are consistent when I control for the field of the last degree attained (Table 5). Interestingly, the Nord and pas-de-Calais region was initially known for its coal and mining industries. The share of blue collars was therefore particularly high in this region, Table 2 illustrates this fact, where we can clearly see a higher share of individuals whose father is or was blue collar in Nord and Pas-de-Calais compared to the rest of France. As a result, it makes sense to find that opening a new university in this region reduces the probability to work as blue collar worker, this is because universities are supposed to create skilled workers.

To obtain the results for the decomposition of the total impact of university openings on labor market outcomes, between the part that is due to human capital from the part that is due to job

[11] Note that, one could argue that the field of the last degree attained is not entirely exogenous, therefore I do not include it as a control variable in the main specification.

creation in the area, I estimate Equations (4) and (5) separately. I then compute the difference in DiD estimated coefficients from the two regressions (i.e. $\hat{\alpha}$-$\hat{\gamma}_1$) to obtain the magnitude of the contribution of human capital in the effect of university openings on labor market outcomes. Results are given in Tables 6, 7 and 8. They show that the impact of university openings on labor market outcomes that goes through human capital formation is always positive and significant (column(difference)), however the impact related to universities spillover is not as straightforward (the coefficient on the interaction term Treat x Post in column (2)). If we consider results for the probability to be employed shown in Table 6, we see that in the case of Charente Maritime both effects are positive, the total effect of university openings on labor market outcomes is hence positive (column (1)). However, in the case of Nord-Pas-de-Calais the impact of university openings on employment that goes through human capital formation is positive, but the effect that is related to university spillovers is negative. The negative impact seems to be more pronounced, therefore the total effect on the probability to be employed is negative for Nord-Pas-de-Calais. This negative effect of university openings on employment, that goes through universities spillover can be due to statistical discrimination based on the observed characteristic (new university which the quality is unkown by employers). The negative spillover can also be due to an over education and skill downgrading hypothesis. A combination of the two hypotheses (over education-skill downgrading and negative signal due to unknown university quality) is plausible as well to explain the negative spillover effect.

Table 9 gives the estimates of γ_2 , the effect of university opening on human capital in Equation 7. The results from Tables 6 and 7 , when combined with Panel A results of Table 9 allow to illustrate the Gelbach decomposition formula described in Equation 6. In the case of Table 8 where the outcome variable is log monthly wages, results are to be combined with Panel B of Table 9 in order to illustrate the Gelbach decomposition formula. For example, from Table 6 we read in the second column for Charente Maritime County $\hat{\beta}$= 0.026 (reflecting the positive link between labour-market outcome Y and education) and $\hat{\gamma}_1$=0.064 (estimate of the magnitude of the spillover effect of UO). From Table 9, we have for Charente Maritime the DID-estimated positive impact of

UO on educational attainment $\hat{\gamma_2}$=0.474. According to the formula in Equation 6, the total effect of university openings on labor market outcomes is :

$$\hat{\alpha} \quad = \quad \hat{\gamma_1} + \hat{\gamma_2} \times \hat{\beta} = 0.064 + 0.474 * 0.026 = 0.076$$

And this is exactly what we obtain from estimating Equation 4, i.e., the Equation that gives the total effect of university openings, i.e., the effect that goes through human capital channel plus the spillover effect. (Estimated coefficients α are reported in Tables 1 and 6).

Table 1: DID estimates ($\hat{\alpha}$)

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes-Nouvelles
Employment	0.076***	-0.041**	-0.039**	-0.029*
	(0.015)	(0.016)	(0.016)	(0.015)
# obs	20433	23608	16368	29429
Stable employment	0.062***	-0.035**	-0.039**	-0.051**
	(0.012)	(0.014)	(0.014)	(0.021)
# obs	20433	23608	16368	29429
log (wages)	0.048***	0.017*	0.010	-0.006
	(0.012)	(0.007)	(0.006)	(0.016)
# obs	10978	12854	9172	17545
Df	10	6	5	14

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C - L$ degrees of freedom, with C the number of clusters in the regression of interest and L the number of regressors which are invariant within cluster, i.e. the intercept and the dummy $Treat_c$, so $L = 2$.
(iii.) The set of county-specific time-varying variables are: Unemployment rate for those who are aged 35 and above, the share of executives and intermediate professions and the number of newborns in $t - 18$
The set of individual variables are: a dummy variable that is equal to one if she is 12 or more years old when she entered in middle school, a dummy variable equal to one if the individual is a female, a dummy variable equal to one if the individual's father had an economic activity or not when she left the education system, a set of dummies reflecting the last socio-professional category of the individual's father when the individual left the education system –farmer, storekeeper, executive, technician, employee, other, or missing answer (the category of reference is blue-collar)–, a set of dummies reflecting the origins of her parents –two parents born in France, two parents born abroad (the category of reference is only one of her parents was born in France or if he has only one parent).

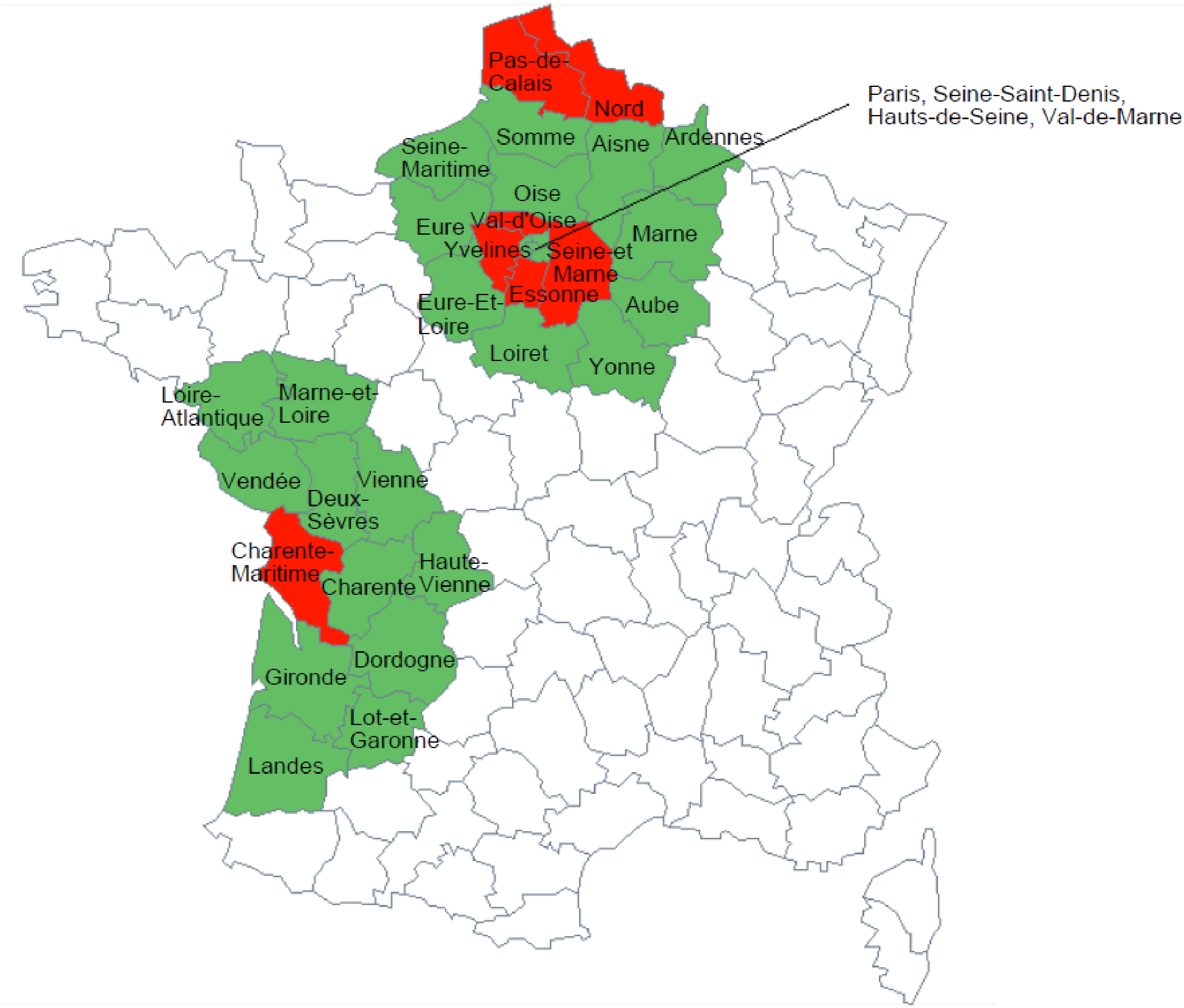

Figure 1: Individuals asssignment into treated and control groups

Notes:(i) individuals are assigned into treated and non treated(control) based on their county of residence at the age of entry in middle school. The red colour on the map corresponds to counties where a new university opened between 1991-1993 and therefore individuals who lived in one of these counties at their age of entry in middle school are considered to be treated no matter the year in which they entered middle school. The green colour corresponds to where individuals in the control group belong.

(ii) The control group for those who lived in Charente-Maritime at their age of entry in middle school are those who lived at their age of entry in middle school in one of the following neighbouring counties: Vendée, Deux-Sèvres, Charente, Gironde, Loire-Atlantique, Marne-et-Loire, Veinne, Haute-Vienne, Dordogne, Lot-et-Garonne and Landes.

(iii) The control group for those who lived in Nord and Pas-de-Calais at their age of entry in middle school are those who lived at their age of entry in middle school in the following neighbouring counties: Somme, Aisne, Ardennes, Seine-Maritime, Marne and Oise.

(iv) The control group for those who lived in Villes-Nouvelles (Val d'Oise, Essonne, Seine-et-Marne and Yvelines) at their age of entry in middle school are those who lived at their age of entry in middle school in the following neighbouring counties: Paris, Seine-Saint-Denis, Hauts-de-Seine, Val-de-Marne, Oise, Eure, Erue-Et-Loire, Loiret, Yonne, Aube, Marne and Aisne.

Table 2: Situation of the counties in 1992

	Mean (Standard-deviation)				
	Charente-Maritime ($c = CM$)	Pas-de-Calais ($c = PC$)	Nord ($c = N$)	Villes Nouvelles ($c = VN$)	Remaining counties of France
Attaining a four-year post-secondary degree or higher diploma					
	0.059	0.062	0.088	0.169	0.113
	(0.236)	(0.241)	(0.283)	(0.375)	(0.316)
Last socio-professional category of the individual's father when he or she left the education system					
Farmer	0.092	0.039	0.023	0.005	0.065
	(0.289)	(0.194)	(0.150)	(0.068)	(0.247)
Storekeeper	0.158	0.059	0.077	0.095	0.121
	(0.366)	(0.236)	(0.267)	(0.293)	(0.326)
Executive	0.070	0.093	0.125	0.254	0.128
	(0.256)	(0.291)	(0.331)	(0.435)	(0.334)
Technician	0.099	0.089	0.111	0.120	0.092
	(0.299)	(0.285)	(0.315)	(0.325)	(0.288)
Employee	0.195	0.170	0.191	0.197	0.193
	(0.397)	(0.376)	(0.393)	(0.398)	(0.394)
Blue collar	0.294	0.454	0.363	0.245	0.313
	(0.456)	(0.498)	(0.481)	(0.430)	(0.463)
County variables					
Unemployment rate	11.2	11.9	11.6	6.004	8.512
(in %)				(0.694)	(1.901)
Observations	281	1002	1412	1697	16315

Notes: Descriptive statistics computed from CEREQ Génération survey data, wave 1992.

Table 3: DID estimates ($\hat{\alpha}$) (adding the field of the last degree attained to the set of control variables)

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes-Nouvelles
Employment	0.070***	-0.022	-0.0215	0.027*
	(0.016)	(0.012)	(0.012)	(0.013)
# obs	20433	23608	16368	29429
Stable employment	0.055***	-0.025	-0.027*	-0.047**
	(0.014)	(0.013)	(0.013)	(0.016)
# obs	20433	23608	16368	29429
log (wages)	0.060***	0.017**	0.016**	-0.006
	(0.013)	(0.005)	(0.006)	(0.015)
# obs	10978	12854	9172	17545
Df	10	6	5	14

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) $*$, $**$ and $***$ refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1, as well as a set of dummy variables reflecting the fields of study at the date the individual leaves the education system: Business and Administration, Literature-Languages-Social Sciences, Science & technology, Environmental Sciences, Social & Health, Humanities & Art, Law & Political Sciences, Services and the reference category is humanities and art.

Table 4: DID estimates ($\hat{\alpha}$) by type of job

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes-Nouvelles
Executives	0.057***	-0.007	-0.022**	0.020*
	(0.012)	(0.011)	(0.008)	(0.011)
Employee	0.008	-0.015	-0.010	-0.051**
	(0.010)	(0.009)	(0.009)	(0.022)
Blue collar	0.001	-0.021*	-0.030**	0.008
	(0.011)	(0.011)	(0.009)	(0.014)
# obs	20433	23608	16368	29429
Df	10	6	5	14

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) $*$, $**$ and $***$ refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table (1)

Table 5: DID estimates ($\hat{\alpha}$) by type of job (adding the field of the last degree attained to the set of control variables)

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes-Nouvelles
Executives	0.051***	-0.001	-0.025*	0.018
	(0.012)	(0.015)	(0.012)	(0.012)
Employee	0.003	-0.023*	-0.011	-0.044**
	(0.010)	(0.011)	(0.010)	(0.018)
Blue collar	-0.002	-0.019**	-0.037***	0.007
	(0.010)	(0.007)	(0.008)	(0.014)
# obs	20433	23608	16368	29429
Df	10	6	5	14

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table (1)

5 Robustness

5.1 Sector of activity

I also control for the sector of activity at which the individual ends up being employed 3 years after having left the French education system. The idea behind this is that one might think the observed effects of UO on wages are simply the result of higher/lower paying industries, so by including the sctor of industry in the set of control variables I rule out this hypothesis. Results are depicted in Table 11. Note that I only provide results for the outcome log wages because the sector of activity is defined for individuals who are employed only. Results are consistent with main findings for Charente Maritime, the effect of UO on employment is positive and statistically significant. I acknowledge that including the sector of activity is somewhat endogenous as university openings influence both labor market outcomes and companies establishments. There are also some omitted factors that impact both the sector of activity and labor market outcomes.

Table 6: Decomposition of the total effect of university openings on employment

Dependent variable=1 if individual is working 3 years after having left the French education system

Charente Maritime		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	0.076***	$\hat{\gamma}_1$ =	0.064***	0.012***
		(0.015)		(0.016)	(0.003)
Educational attainment			$\hat{\beta}$ =	0.026***	
				(0.004)	
# obs				20433	
Df				10	

Nord-Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	-0.041**	$\hat{\gamma}_1$ =	-0.058***	0.017**
		(0.016)		(0.003)	(0.007)
Educational attainment			$\hat{\beta}$ =	0.057***	
				(0.002)	
# obs				23608	
Df				6	

Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	-0.039*	$\hat{\gamma}_1$ =	-0.060***	0.021***
		(0.016)		(0.015)	(0.006)
Educational attainment			$\hat{\beta}$ =	0.057***	
				(0.003)	
# obs				16368	
Df				5	

Villes Nouvelles		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	-0.029*	$\hat{\gamma}_1$ =	-0.041***	0.012*
		(0.015)		(0.014)	(0.006)
Educational attainment			$\hat{\beta}$ =	0.039***	
				(0.002)	
# obs				29429	
Df				14	

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) In the first column, estimated coefficients on Treat x Post ($\hat{\alpha}$) from Equation 4 are reported. In the second column both estimated coefficients from Equation (5) : ($\hat{\gamma}_1$) on Treat x Post and ($\hat{\beta}$) on educational attainment are reported. Finally, the third column provides the difference between the two estimated coefficients ($\hat{\alpha}$) and ($\hat{\gamma}_1$) from Equations 4 and (5).
(iii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1

Table 7: Decomposition of the total effect of university openings on stable employment

Dependent variable=1 if individual has a stable employment (permanent contract or government official) 3 years after having left the French education system

Charente Maritime		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	0.062***	$\hat{\gamma}_1$ =	0.044***	0.018***
		(0.012)		(0.013)	(0.002)
Educational attainment			$\hat{\beta}$ =	0.038***	
				(0.004)	
# obs				20433	
Df				10	
Nord-Pas-de-Calais		**From eq(4)**		**From eq(5)**	**Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)**
Treat x Post	$\hat{\alpha}$ =	-0.035***	$\hat{\gamma}_1$ =	-0.054**	0.019**
		(0.014)		(0.016)	(0.008)
Educational attainment			$\hat{\beta}$ =	0.063***	
				(0.003)	
# obs				23608	
Df				6	
Pas-de-Calais		**From eq(4)**		**From eq(5)**	**Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)**
Treat x Post	$\hat{\alpha}$ =	-0.039**	$\hat{\gamma}_1$ =	-0.063***	0.023***
		(0.014)		(0.014)	(0.007)
Educational attainment			$\hat{\beta}$ =	0.064***	
				(0.004)	
# obs				16368	
Df				5	
Villes Nouvelles		**From eq(4)**		**From eq(5)**	**Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)**
Treat x Post	$\hat{\alpha}$ =	-0.051**	$\hat{\gamma}_1$ =	-0.066***	0.014*
		(0.021)		(0.016)	(0.008)
Educational attainment			$\hat{\beta}$ =	0.045***	
				(0.002)	
# obs				29429	
Df				14	

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) In the first column, estimated coefficients on Treat x Post ($\hat{\alpha}$) from Equation 4 are reported. In the second column both estimated coefficients from Equation (5) : ($\hat{\gamma}_1$) on Treat x Post and ($\hat{\beta}$) on educational attainment are reported. Finally, the third column provides the difference between the two estimated coefficients ($\hat{\alpha}$) and ($\hat{\gamma}_1$) from Equations 4 and (5).
(iii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C - 2$ degrees of freedom. The set of control variables is the same as in Table 1

Table 8: Decomposition of the total effect of university openings on log wages

Dependent variable=log monthly wage of the individual 3 years after having left French education system

Charente Maritime		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha} =$	0.048***	$\hat{\gamma}_1 =$	-0.002	0.050***
		(0.012)		(0.008)	(0.007)
Educational attainment			$\hat{\alpha} =$	0.066***	
				(0.002)	
# obs				10978	
Df				10	

Nord-Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha} =$	0.017*	$\hat{\gamma}_1 =$	-0.013	0.031***
		(0.007)		(0.010)	(0.005)
Educational attainment			$\hat{\beta} =$	0.066***	
				(0.001)	
# obs				12854	
Df				6	

Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha} =$	0.010	$\hat{\gamma}_1 =$	-0.022**	0.032***
		(0.006)		(0.006)	(0.004)
Educational attainment			$\hat{\beta} =$	0.065***	
				(0.001)	
# obs				9172	
Df				5	

Villes Nouvelles		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha} =$	-0.006	$\hat{\gamma}_1 =$	-0.031**	0.025**
		(0.016)		(0.011)	(0.107)
Educational attainment			$\hat{\beta} =$	0.074***	
				(0.003)	
# obs				17545	
Df				14	

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) In the first column, estimated coefficients on Treat x Post ($\hat{\alpha}$) from Equation 4 are reported. In the second column both estimated coefficients from Equation (5) : ($\hat{\gamma}_1$) on Treat x Post and ($\hat{\beta}$) on educational attainment are reported. Finally, the third column provides the difference between the two estimated coefficients ($\hat{\alpha}$) and ($\hat{\gamma}_1$) from Equations 4 and (5).
(iii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C - 2$ degrees of freedom. The set of control variables is the same as in Table 1

Table 9: DID estimates ($\hat{\gamma}_2$) for educational attainment measured by years of schooling

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes-Nouvelles
Panel A: The entire sample				
$\hat{\gamma}_2$	0.474***	0.300*	0.367**	0.310*
	(0.067)	(0.122)	(0.092)	(0.167)
# obs	20433	23608	16368	29429
Panel B: Restricting the sample to individuals who are employed				
$\hat{\gamma}_2$	0.759***	0.488***	0.504***	0.361**
	(0.110)	(0.0656)	(0.057)	(0.132)
# obs	10978	12854	9172	17545
Df	10	6	5	14

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1.

5.2 Alternative definitions of control group

Another potential issue one could think of is the presence of displacement effects in the control group. This means that even if no university opening occurred in the neighbouring counties on the studied period 1992 - 2007, there could still exist some effect that is comparable to a university opening, for instance university expansion in one of the neighbouring counties. For this I choose a control group in which the supply of higher education was constant throughout the period. Paris region is a suitable control group as it matches the previous definition. Tertiary education supply is the highest in France in this region and it has been stable throughout the period. I define Paris region as Paris and 3 of its neighbouring counties where university openings did not occur (Seine-Saint-Denis, Val-De-Marne and Hauts-De-Seine). Results are depicted in Table 10. The positive impact on employment and wages in Charente-Maritime is robust to this alternative definition of control group. I then check another alternative definition of the control group. I consider the control group to be all the counties of France except for the treated ones. Using this definition both the positive impacts on employment and wages in Charente-Maritime and the negative impact on employment in Pas-de-Calais and Villes-Nouvelles persist.

5.3 Instrumental Variables

In equation 5 educational attainment is assumed to be unaffected by UO. What is more, education is assumed to be exogenous, i.e., uncorrelated to the error term in equation 4. However, there are unobserved determinants of Labor market outcomes that are correlated to schooling, hence schooling is not exogenous. To address this potential bias, I use grade repetition before 6th grade as an instrument for educational attainment. The idea is that for grade repetition before 6th grade to be a valid instrument, it needs to determine educational attainment and affects Labor Market Outcomes only through educational attainment. Grenet (2010) shows using French data that month of birth has an important effect on grade repetition in primary school, the effect decreases till end of high school. He argues that month of birth doesn't have a significant effect on labor market outcomes, however it has an effect on education and especially at an early age. This gives confidence in the validity of this instrument.

Results of IV two stage least squares are presented in tables 14, 15 and 16. First stage results show a negative and statistically significant effect of grade repetition in primary school on educational attainment. Second stage results and most importantly the coefficient on $Treat \times Post$ are consistent with main findings (estimates of α in equation 4). The effect of UO on labor market outcomes is overall positive and significant for Charente Maritime and significant but negative in the North of France (Nord-Pas-de-Calais and when excluding Nord county). What is more, multiplying the coefficient on $Treat \times Post$ from first stage by the coefficient on educational attainment from second stage yields the same result as multiplying $\hat{\gamma}_2$ by $\hat{\beta}$) in Gelbach decomposition formula where results are reported in tables 6, 7 and 8.

5.4 Propensity score weighting and Doubly Robust DID

Regarding the DID strategy, one of the main hypotheses that needs to hold is the parallel trend assumption. The conventional DID estimator requires that, in the absence of the treatment, the average outcomes for the treated and control groups would have followed parallel paths over time. This assumption may be implausible if characteristics that are thought to be associated with the dynamics of the outcome variable are unbalanced between the treated and the untreated and also possible before and after the treatment. Since I only have one pre-treatment wave of data, the common trend assumption cannot be tested, and I also cannot consider implementing DID methods accounting for pre-treatment non parallelism. Still we can implement what is know in the literature as conditional DiD (cDID hereafter). The idea is a generalization of the unverifiable (in the case of this paper) parallel trends assumption, to parallel trends conditional on observables. In other words I assume that individuals forming a treated groups with a particular set of characteristics (X) would have, on average, experienced the same changes in their outcomes had they not received the program as individuals with their same observable characteristics forming the control group. In a sense, I have already implemented cDiD by adding control variables (X, Z) to my main DID model (equation 4). What I propose here is to assess the robustness of my results by implementing more elaborate cDID methods. The first one is the probability weighting DID method proposed by Stuart et al. (2014). The second is the Doubly Robust DID (DR DID) proposed by Sant'Anna and Zhao (2020) that combines probability weighting with regression adjustment. See Appendix for details on how these methods are implemented in this paper.

Results of first method (Probability weighting DID à la Stuart et al. (2014)) are presented in Table 12. Estimates are similar to main results estimates in Table 1, which suggest that change in group composition is not an issue in this paper and that the conditional parallel trends assumption holds.

Results of the second method (DR DID) method are presented in table 13. Results are overall consistent with main findings. The cDID estimate of the impact of UO on labor market outcomes is positive and its mean is significantly different from zero for Charente Maritime. When considering Pas de Calais county, results are in line with what is observed once I control for fields of last degree attained (table 3), i.e., the negative effect of UO on employment is not statistically different from zero.

Table 10: DID estimates ($\hat{\alpha}$) for main variables with different control groups

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes-Nouvelles
Paris region as a control group				
Employment	0.152***	0.046**	0.029	0.004
	(0.018)	(0.016)	(0.021)	(0.020)
# obs	10038	20181	13031	18332
Stable employment	0.025	-0.047*	-0.084**	0.011
	(0.022)	(0.021)	(0.022)	(0.020)
# obs	10038	20181	13031	18332
log (wages)	0.051**	0.036*	0.027	0.004
	(0.014)	(0.015)	(0.016)	(0.016)
# obs	5738	11178	7464	11095
Df	3	4	3	6
Remaining counties of France as a control group				
Employment	0.088***	-0.010	-0.016**	-0.057***
	(0.007)	(0.008)	(0.007)	(0.013)
# obs	124812	134955	127805	133106
Stable employment	0.023***	-0.047***	-0.050***	-0.002
	(0.005)	(0.006)	(0.005)	(0.010)
# obs	124812	134955	127805	133106
log (wages)	0.045***	0.031**	0.014***	-0.020*
	(0.004)	(0.013)	(0.005)	(0.011)
# obs	71686	77126	73412	77043
Df	88	89	88	91

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1

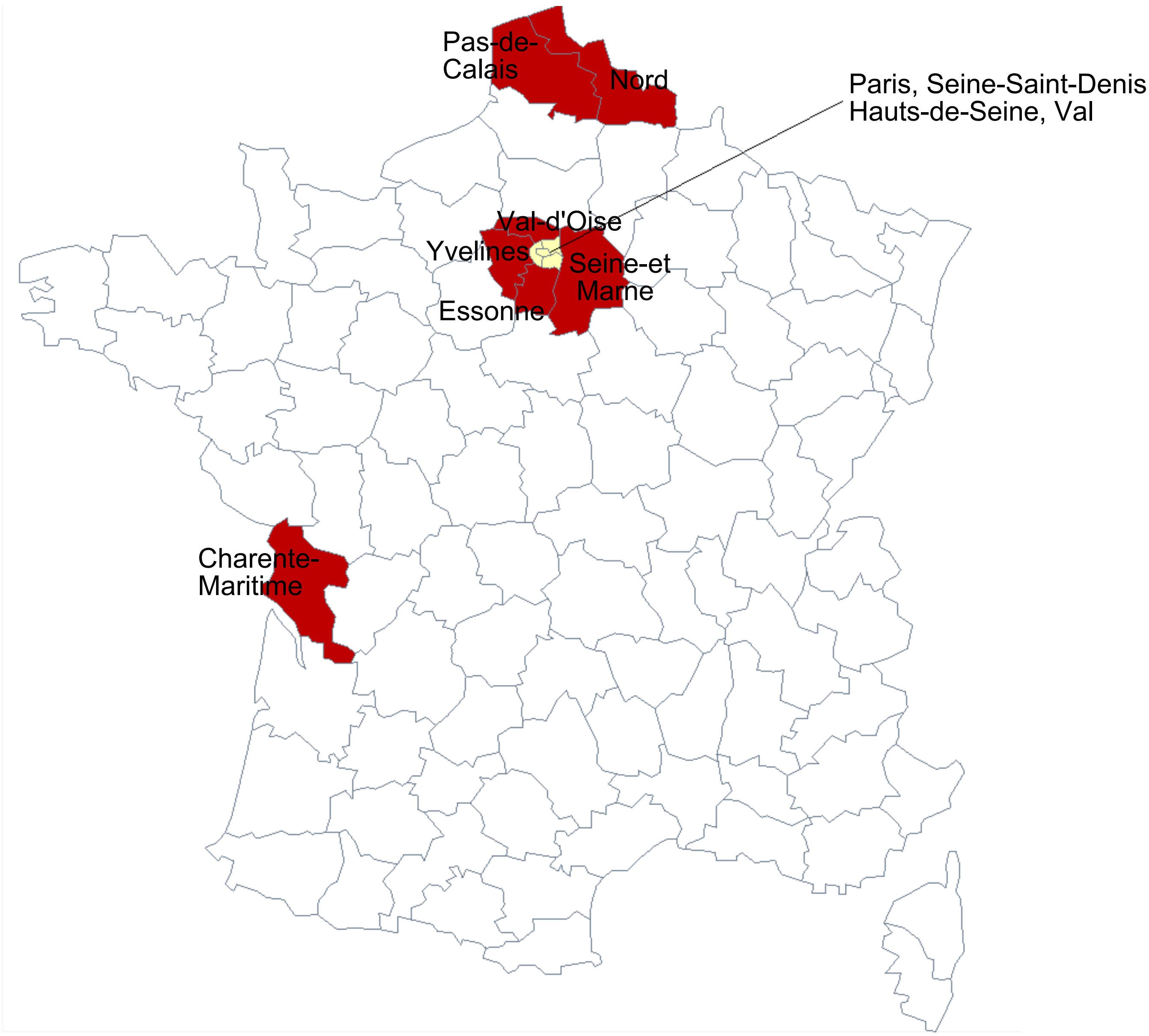

Figure 2: Individuals asssignment into treated and control groups (alternative control group1)

Notes:(i) individuals are assigned into treated and non treated(control) based on their county of residence at the age of entry in middle school. The red colour on the map corresponds to counties where a new university opened between 1991-1993 and therefore individuals who lived in one of these counties at their age of entry in middle school are considered to be treated no matter the year in which they entered middle school. The yellow colour corresponds to where individuals in the control group belong.

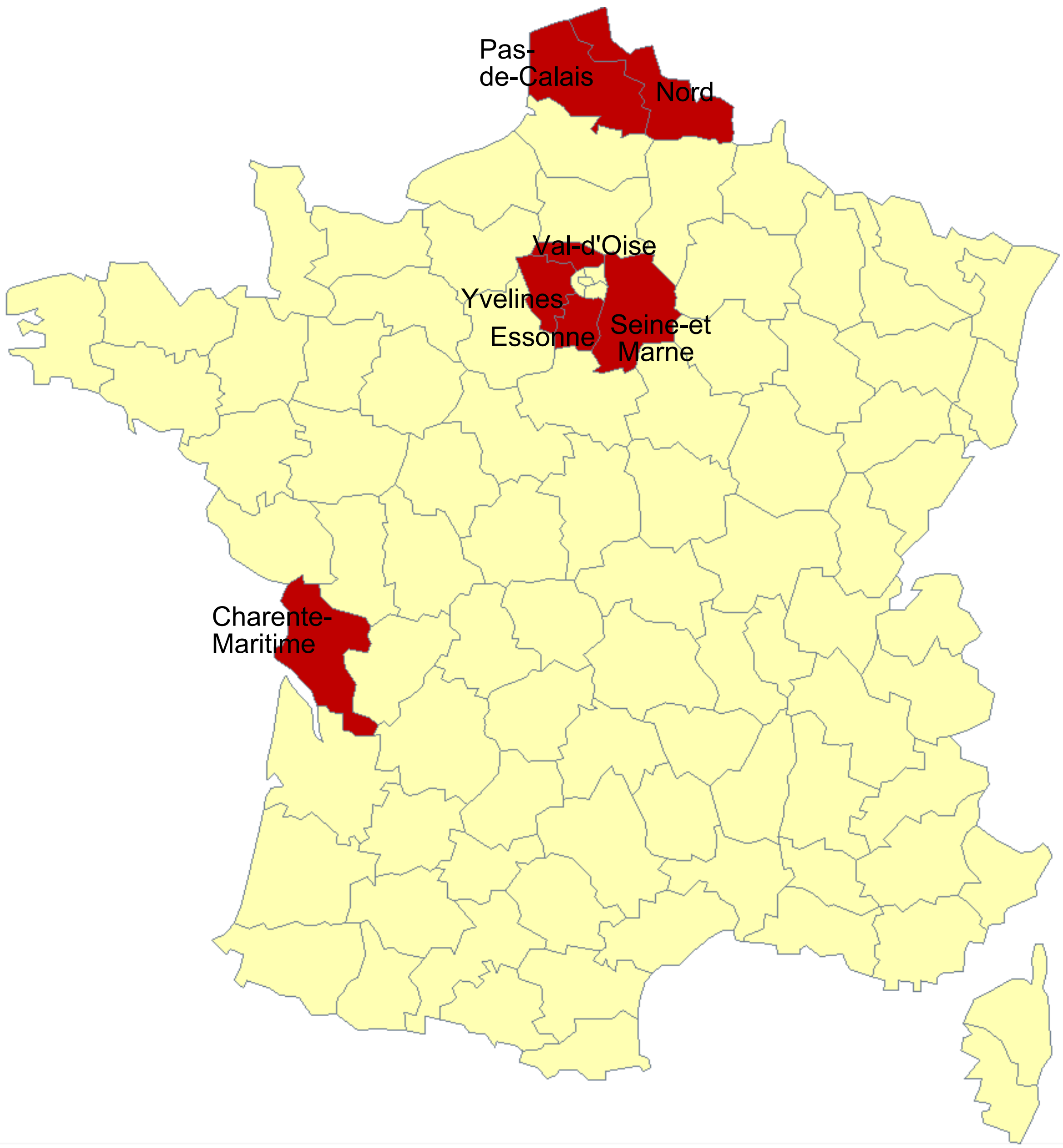

Figure 3: Individuals assignment into treated and control groups (alternative control group2)

Notes:(i) individuals are assigned into treated and non treated(control) based on their county of residence at the age of entry in middle school. The red colour on the map corresponds to counties where a new university opened between 1991-1993 and therefore individuals who lived in one of these counties at their age of entry in middle school are considered to be treated no matter the year in which they entered middle school. The yellow colour corresponds to where individuals in the control group belong (all of France except for treated counties)

Table 11: DID estimates ($\hat{\alpha}$) when controlling for the sector of activity

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes-Nouvelles
log (wages)	0.047***	0.022**	0.0132*	-0.005
	(0.013)	(0.008)	(0.006)	(0.016)
# obs	10978	12854	9172	17545
Df	10	6	5	14

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1, as well as a set of dummy variables reflecting the sector of activity of the company in which the individual is employed: agricultural, industrial, services, construction and other. The reference category is agricultural sector.

Table 12: DID estimates ($\hat{\alpha}$) after integrating propensity score in DID model

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes-Nouvelles
Employment	0.083***	-0.034	-0.046*	-0.032**
	(0.015)	(0.020)	(0.020)	(0.014)
# obs	20433	23608	16368	29429
Stable employment	0.082***	-0.026	-0.034	-0.051**
	(0.011)	(0.016)	(0.018)	(0.021)
# obs	20433	23608	16368	29429
log (wages)	0.048***	0.021**	0.016**	-0.002
	(0.011)	(0.007)	(0.006)	(0.016)
# obs	10978	12854	9172	17545
Df	10	6	5	14

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1.
iii) weights used are propensity score weights wi multiplied by the sample weights à la Ridgeway et al. (2015).

Table 13: cDID estimates ($\hat{\alpha}$) according to Doubly Robust DID method

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes Nouvelles
Employment	0.029***	0.036***	-0.005	-0.074***
	[0.000]	[0.001]	[0.622]	[0.000]
# obs	20433	23608	16368	29429
Stable employment	0.031***	0.062***	0.019*	-0.084***
	[0.000]	[0.000]	[0.091]	[0.000]
# obs	20433	23608	16368	29429
Log(wages)	0.045***	0.034***	-0.010	-0.035***
	[0.000]	[0.001]	[0.256]	[0.000]
# obs	10978	12854	9172	17545
# counties	10	6	5	14

Notes: (i.) Between brackets are p-values from t-test.
(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels.
(iii.) Control variables in the multinomial logistic regression are the following individual characteristics: a dummy for female, a dummy equals to 1 if the individual's father had an economic activity and equals to 0 otherwise, a dummy equal to 1 if individual repeated 6th grade and 0 otherwise, dummies for the origins of parents (whether both born abroad, both born in france or one born in France and the other abroad), and a set of dummies for last socio professional category for the individual father: farmer, business man, exectuive, technician, employee or blue collar. Control variables in the OlS regression are the following county-specific time-varying variables are: Unemployment rate for those who are aged 35 and above, the share of executives and intermediate professions and the number of newborns in $t-18$.

Table 14: IV results: dependent variable=1 if the individual is working 3 years after having left the French education system

	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes Nouvelles
First stage				
Grade repetition before 6th grade	-1.961***	-1,722***	-1.666***	-0.041***
	(0.061)	(0.066)	(0.068)	(0.013)
Treat x Post	0.474***	0.300**	0.367***	0.301*
	(0.067)	(0.122)	(0.092)	(0.167)
Second stage				
Treat x Post	0.058***	-0.057***	-0.059***	-1.828***
	(0.016)	(0.015)	(0.014)	(0.065)
educational attainment	0.039***	0.056***	0.054***	0.039***
	(0.005)	(0.007)	(0.012)	(0.006)
#obs	20,433	23,608	16,368	29,429
Df	10	6	5	14

Notes:(i) Grade repitition before 6th grade dummy variable is used as an instrument for educational attainment measured by years of schooling
(ii.) Cluster-robust standard errors in parentheses (clustering on county).
(iii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1.

Table 15: IV results: dependent variable=1 if the individual has a stable employment 3 years after having left the French education system

	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes Nouvelles
First stage				
Grade repetition before 6th grade	-1.961***	-1.722***	-1.666***	-1.828***
	(0.061)	(0.066)	(0.068)	(0.065)
Treat x Post	0.474***	0.300**	0.367***	0.310*
	(0.067)	(0.122)	(0.092)	(0.167)
Second stage				
Treat x Post	0.035**	-0.054***	-0.063***	-0.065***
	(0.014)	(0.015)	(0.013)	(0.014)
educational attainment	0.057***	0.062***	0.063***	0.050***
	(0.006)	(0.004)	(0.007)	(0.006)
#obs	20,433	23,608	16,368	29,429
Df	10	6	5	14

Notes:(i) Grade repitition before 6th grade dummy variable is used as an instrument for educational attainment measured by years of schooling
(ii.) Cluster-robust standard errors in parentheses (clustering on county).
(iii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1.

Table 16: IV results: dependent variable is log monthly wage 3 years after having left the French education system

	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes Nouvelles
First stage				
Grade repetition before 6th grade	-2.022***	-1.831***	-1.807***	-1.882***
	(0.075)	(0.069)	(0.093)	(0.094)
Treat x Post	0.759***	0.488***	0.504***	0.361**
	(0.110)	(0.066)	(0.057)	(0.132)
Second stage				
Treat x Post	0.006	-0.017*	-0.026***	-0.031**
	(0.011)	(0.008)	(0.006)	(0.011)
educational attainment	0,072***	0.071***	0.071***	0.077***
	(0.006)	(0.004)	(0.006)	(0.004)
#obs	10,978	12,854	9,172	17,545
Df	10	6	5	14

Notes:(i) Grade repitition before 6th grade dummy variable is used as an instrument for educational attainment measured by years of schooling
(ii.) Cluster-robust standard errors in parentheses (clustering on county).
(iii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1.

6 Conclusion

In this paper, I highlight the impact of university openings on labor market outcomes. The main contribution of this paper is that increasing human capital supply in a region does not necessarily translate into better labor market outcomes. Also, I show that there are two channels through which universities impact labor market outcomes: (i) The human capital channel, (ii) spillover channel.

I study the case of university openings that occurred in France in the 90's in a difference-in-differences framework. I exploit five waves from Cereq generation surveys, starting from wave 1992, which corresponds to my before treatment year in the DID strategy, till wave 2007. My main findings are that the probability of being employed increases by about 8% points in the west of France where individuals come from disadvantaged backgrounds, and wages increase by about 5%. On the other hand, in regions where the unemployment rate is low and where individuals' socio-economical background (reflected by parents' occupation) is more advantageous, opening a new university does not have an impact on labor market outcomes. I find that the effect of university openings on labor market outcomes that goes through the human capital channel is always positive. However, the spillover channel can be positive, negative or null.

Finally, this paper constitutes the building block for future research towards understanding the repercussions of increasing the supply of higher education. I provide evidence for a more complex effect of university openings on labor market outcomes than one would intuitively expect. I show that university openings can generate positive or negative externalities "spillovers". One can then think of exploring this spillover channel as a next step in research.

7 Appendix

7.1 Appendix A

- **The probability weighting DID method**

  This method suggested by Stuart et al. (2014) combines propensity score matching with difference-in-differences. It allows to relax the unconditional parallel trends assumption in favour of a conditional parallel trends assumption, and it reduces the selection bias related to the change in the composition of treatment and control groups through time. The propensity score which is the probability of receiving the treatment is used to weigh outcome control and treatment groups based on observed characteristics. The contribution of Stuart et al. (2014) is to weigh 4 groups: treatment group before and after treatment, comparison group before and after treatment. Each individual will have 4 propensity scores reflecting probabilities of being in each of the 4 groups. Unlike the standard inverse probability weight methods where propensity scores for treated and non treated individuals are computed based on before observations only. The propensity scores are obtained by estimating a multinomial logistic regression model where each group is predicted using the set of individual characteristics X. The weights are hence computed as follows:

$$w_i = \frac{e_1(Xi)}{e_g(Xi)} \tag{8}$$

  $e_1(Xi)$ is the probability for individual i to be in group 1, and $e_g(Xi)$ is the probability for individual i to be in group g (s.t. g=1 to 4). Individuals in group 1 have a weight w_1 equals to 1. The set of individuals characteristics I use in the multinomial logistic regression is: a dummy for female, a dummy equals to 1 if the individual's father had an economic activity and equals to 0 otherwise, a dummy equal to 1 if individual repeated 6th grade and 0 otherwise, dummies for the origins of parents (whether both born abroad, both born in france or one born in France and the other abroad), and a set of dummies for last socio professional category for the individual father: farmer, business man, exectuive, technician, employee or blue collar. Once I obtain these propensity scores weights w_i, I multiply them by the sample weight *pondef* for each individual à la Ridgeway et al. (2015). Estimating Equation 4 but using the weights W_i (s.t. $W_i = w_i * pondef$) yields the treatment effects estimates regardless of the change in composition of treatment group across time and/or group.

- **Doubly Robust difference in differences (DR DID)**

  DR DID is based on the conditional parallel trends assumption rather than the parallel trends. This DR DID method was developed by Sant'Anna and Zhao (2020), based on the work of Vermeulen and Vansteelandt (2015) and Graham et al. (2012). It combines Heckman et al. (1997) outcome regression OR approach and Abadie (2005) propensity score weighting approach (Inverse probability weighting IPW). Combining the two approaches allows to control for X vector of individual characteristics in IPW, and for Z vector of county level characteristics in OR. Hence, it gives better chances to get a more accurate estimation of the average treatment on treated (ATT). The estimation of ATT is more precise as only one of the two models (OR / IPW) needs to be correctly specified.

  Recall that DID main objective is to estimate ATT, defined as:

$$ATT = E[Y1(1)|Treat = 1] - E[Y1(0)|Treat = 1] \tag{9}$$

  Which is the treatment effect on the treated group after the treatment period $t = 0$. $Y1(1)$ is the outcome for treated individuals after the treatment. $Y1(0)$, is the unobserved outcome for treated individuals hadn't they received the treatment.

  In the case of repeated cross section, Sant'Anna and Zhao (2020) propose computing the ATT as follows:

$$ATT^{DR} = E\left[(w_1(Treat, Post) - w_0(Treat, Post, X; \pi))(Y - \mu_{0,Y}(Post, Z))\right] \tag{10}$$

Where the idea of propensity weighting (but also the idea of a difference-in-differences estimation) rests on:

$$\begin{aligned} w_1(Treat, Post) &= w_{1,1}(Treat, Post) - w_{1,0}(Treat, Post) \\ w_0(Treat, Post, X) &= w_{0,1}(Treat, Post, X) - w_{0,0}(Treat, Post, X) \end{aligned} \tag{11}$$

With

$$\begin{aligned} w_{1,1}(Treat, Post) &= Post \times Treat \\ w_{1,0}(Treat, Post) &= Treat \times \frac{e_1(X_i)}{e_2(X_i)} \\ w_{0,1}(Treat, Post, X) &= Post \times (1 - Treat) \times \frac{e_1(X_i)}{e_3(X_i)} \\ w_{0,0}(Treat, Post, X) &= (1 - Treat) \times \frac{e_1(X_i)}{e_4(X_i)} \end{aligned}$$

Note that, the use of multinomial logit-estimated propensity scores e(X)'s is a deviation from what Sant'Anna and Zhao (2020) propose.

The implementation of the OR idea corresponds to the presence of:

$$\mu_{0,Y}(Post, Z) \quad = Post \times resid_{0,1}(Z) + (1 - Post) \times resid_{0,0}(Z)$$

Where $resid_{0,1}(Z)$ and $resid_{0,0}(Z)$ are residuals from OLS regressions of the outcome variable on county-specific time varying variables in treatment group after the treatment ($Treat = 0$ and $Post = 1$) and individuals in treatment group before the treatment ($Treat = 0$ and $post = 0$) respectively.

7.2 Appendix B

Details on university creations in France between 1991-1993:

- Université d'Évry Val-d'Essonne:

  The university is located at Essonne county, characterised by important demographic and economic growth perspectives, with 13 different sites, most of which are in Evry district except for few sites in Brétigny and Athis-Mons. Since the creation of the university, there were significant delays in the construction of buildings that penalized the development of the university. The delays forced the university to re-use existing buildings (such as social security building, post office...) also building rentals that constitute a fairly heavy financial burden as operating costs are high. Overall, new constructions constitute only 30% of surface areas excluding IUT.[12] The opening of the university's library in 2001 improved the situation by providing a place for students to communicate and interact. The university's objective upon opening was mainly focused on science and technology (mathematics and computer sciences, physics of materials and science for engineers) and on vocational training in the technological and tertiary sectors. (CNE Évry university Report (2006d)).

- Université de Cergy-Pontoise

  Opened in the Val d'Oise county, it first consisted of a DEUG in Physics, with 90 students, created by a team of professors sent by the ministry of higher education from Orsay (part of Paris-Sud university). sent from the ministry of higher education. In October 1992, the 23,000 m2 building "*Les Chênes*" was opened and dedicated to literature and human sciences. Then, in 1994, the local authorities delivered, in two successive phases, the building of "*Saint-Martin*", in the the municipality of Pontoise, but close to "*Les Chênes*". The building is devoted to sciences (Maths, computer sciences...). In December 1995, the building of "*Neuville*" was opened, in a village where an industrial zone is meant to be developed. The latter hosted the following disciplines: civil engineering, electrical engineering, experimen-

[12] *Institut Universitaire de Technologie* (IUT) is where the two-year vocational degrees called Diplôme Universitaire Technologique (DUT) are prepared

tal physics and heavy chemistry. Several extension operations occurred in the following years. The university also has sites in Argenteuil and Sarcelles, both municipalities are located in Val d'Oise county. Starting january 2020, this Cergy-Pontoise university is called CY Cergy Paris Université. (CNE UCP Report (2006a))

- Université de Versailles Saint-Quentin-en-Yvelines (UVSQ)

  Located in Yvelines county, the university first consisted of two branches of pre-existing universities. The branch of the University of Paris 10-Nanterre, established in 1985 at Yvelines as a result of the increasing number of students especially those coming from Yvelines. The branch covered the disciplines of law, economics and management, human and social sciences, literature and languages. The second pre-existing branch of the university was established in 1987 by the Pierre and Marie Curie University - Paris 6, that also faced an increase in the number of students and hence decided to open a Branch at Versailles (a city in Yvelines county). The latter brings together two DEUG[13](a two year university degree), a DEA[14] (the equivalent of a master of research degree) and a laboratory in the field of fundamental sciences. In November 1991, the IUT, created in Vélizy (at Yvelines county), was added to these two components, then, the following year, the university opened the Yvelines Institute of Sciences and Techniques, a computer engineering school located in Versailles. The development of the university continued in the following years. In 2001, the Paris West Medicine faculty, from the René Descartes - Paris 5 University, was attached to the UVSQ. (CNE UVSQ report (2006c))

- Université de Marne la Vallée (UMLV)

  Created at Seine-et-Marne county, the university's objective was to improve the access to higher education for high school graduates from the east of Ile de France region. The university's sites are located in Champs-sur-Marne and val d'Europe parts of the Marne

[13] *Diplôme d'Etudes Universitaires Générales* was a two year university degree before 2007. This degree no longer exists after the establishment of the LMD system by Bologna process.

[14] *Diplôme d'Etudes Universitaires Générales* no longer provided in French universities after 2007.

la Vallée city and is considered to be a multidisciplinary university (excluding law and health) with a predominance of vocational fields. The university was established starting from the pre-existing branch of the University of Paris 12 - Denis Diderot de-localized in 1989. UMLV suffered from important financial difficulties due to the low commitment of the Seine-et-Marne county compared to its three counterparts in Ile de France (Essonne, Val d'Oise and Yvelines), leading to a lack of programming by the university's administrators and presidents for a real campus creating a better inclusive environment for the community of students and professors/researchers.Hence, the university had to buy buildings with low capacity to host a large number of students and with high maintenance costs which continue to burden university's operating budgets. (CNE UMLV report (2006b))

Two universities created by decree law of november 7, 1991, in Nord-Pas-de-Calais region to mainly relieve the pressure on universities in Lille in terms of the increasing number of students:

- Université d'Artois

  The university is located in different areas of the Pas-de-Calais county and each location is specialized in a different discipline:

  - Human sciences, literature and languages at Arras.
  - Economics, Management and Technology at Béthune.
  - Law at Douai.
  - Sciences at Lens.
  - STAPS (Sciences and techniuqes of physical and sports activities) at Lévin.

  This multipolarity, makes it difficult to communicate across disciplines and reinforces the local specialisation of each city based on the university's component attached to it. When Artois university first opened its doors, it included established branches from the universities of Lille 1 and Lille 3 universities: for example, a DEUG in modern literature of Lille 3 university on the Arras site (existing since 1988) , a DEUG in Economics at Béthume established by Lille 1 university in 1990... However, starting from its first year of operation, it also opened many other programs and degrees of its own. (CNE Artois university report (1996a))

- Université du Littoral

  Created by the same decree law as Artois university and for the same reason (relieving the pressure from Lille), the Littoral university was officially functional starting September 1992. It is the result of the grouping in an autonomous entity, of previous programs from the 3 universities in Lille (Lille 1, Lille 1 and Lille 3). The Calais sites of the University of Lille I were the oldest: there had been a scientific training since 1963, that became a DEUG in 1976. Littoral university is located on twenty sites in Boulogne-sur-Mer, Calais and Saint-Omer (municipalities in the Pas-de-Calais) and Dunkerque (city in Nord) which hosts the university's headquarter. Overall, the number of programs and students is more or less balanced across the different sites, with a slight advantage for Dunkerque. The particularity of this university is in its strategy in covering local needs at the undergraduate level mainly, by providing a panel of discipline choices on each site. This duplication of programs across sites, avoids local specialisations, for example the DEUG of Applied Languages and Modern Literature was taught both in Dunkerque and Boulogne. (CNE Littoral university report (1996b))

The last university opened is the university of la Rochelle, created by decree law of January 20, 1993 in Charente-Maritime county, the university operates autonomously from October of the same year.

- Université de la Rochelle

  The university of la Rochelle is distinguished from the other universities by it not replacing pre-existing branches and doesn't have the declared objective of relieving Poitiers (a district in the bordering county of Vienne). In fact, it functions as a local university, where more than 2/3 of its students coming from the Charentes. In 1993, la Rochelle university was defined as a multidisciplinary higher education establishment, and not as a university with a prevalent thematic. The future growth of this local university is essentially linked to the local demographic prospects of a limited recruitment pool. The university deserves the name of a campus as it is installed in the city and occupies a remarkable site near the historic port of La Rochelle and Minimes Port. Modern and quality constructions are distributed over this vast space, and a large number of studios and private residences offer their services to both students and tourists depending on the season.(CNE La Rochelle university report (1997))

7.3 Appendix C

I illustrate the Gelbach decomposition formula described in Equation 6 using an alternative definition of educational attainment. I define educational attainment as the probability to attain at least two years of higher education. This allows to capture the effect on those who most likely benefited from university openings. Results are shown in the following tables (Tables 17, 18 and 19) and are consistent with results using larger definition of educational attainment measured by years of schooling (Tables 6, 7 and 8). The effect of UO on labor market outcomes that goes through human capital channel is always positive and statistically significant.

Table 17: Decomposition of the total effect of university openings on employment with an alternative definition of educational attainment

Dependent variable=1 if individual is working 3 years after having left the French education system					
Charente Maritime		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha} =$	0.076***	$\hat{\gamma}_1 =$	0.067***	0.009***
		(0.015)		(0.014)	(0.002)
Attaining at least 2 years of HE			$\hat{\beta} =$	0.105***	
				(0.017)	
# obs				20433	
Df				10	
Nord-Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha} =$	-0.041**	$\hat{\gamma}_1 =$	-0.060***	0.020**
		(0.016)		(0.019)	(0.009)
Attaining at least 2 years of HE			$\hat{\beta} =$	0.225***	
				(0.012)	
# obs				23608	
Df				6	
Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha} =$	-0.039*	$\hat{\gamma}_1 =$	-0.065**	0.026***
		(0.015)		(0.018)	(0.008)
Attaining at least 2 years of HE			$\hat{\beta} =$	0.221***	
				(0.017)	
# obs				16368	
Df				5	
Villes Nouvelles		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha} =$	-0.029*	$\hat{\gamma}_1 =$	-0.037***	0.008**
		(0.015)		(0.014)	(0.004)
Attaining at least 2 years of HE			$\hat{\beta} =$	0.158***	
				(0.011)	
# obs				29429	
Df				14	

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) In the first column, estimated coefficients on Treat x Post ($\hat{\alpha}$) from Equation 4 are reported. In the second column both estimated coefficients from Equation (5) : ($\hat{\gamma}_1$) on Treat x Post and ($\hat{\beta}$) on the probability to attain at least two years of higher education are reported. Finally, the third column provides the difference between the two estimated coefficients ($\hat{\alpha}$) and ($\hat{\gamma}_1$) from Equations 4 and (5).
(iii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1

Table 18: Decomposition of the total effect of university openings on stable employment with an alternative definition of educational attainment

Dependent variable=1 if individual has a stable employment (permanent contract or government official) 3 years after having left the French education system

Charente Maritime		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	0.062***	$\hat{\gamma}_1$ =	0.045***	0.012***
		(0.012)		(0.012)	(0.002)
Attaining at least 2 years of HE			$\hat{\beta}$ =	0.148***	
				(0.014)	
# obs				20433	
Df				10	

Nord-Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	-0.035**	$\hat{\gamma}_1$ =	-0.058**	0.023**
		(0.014)		(0.018)	(0.010)
Attaining at least 2 years of HE			$\hat{\beta}$ =	0.260***	
				(0.011)	
# obs				23608	
Df				6	

Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	-0.039**	$\hat{\gamma}_1$ =	-0.069***	0.030***
		(0.014)		(0.016)	(0.009)
Attaining at least 2 years of HE			$\hat{\beta}$ =	0.260***	
				(0.017)	
# obs				16368	
Df				5	

Villes Nouvelles		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	-0.051**	$\hat{\gamma}_1$ =	-0.060***	0.010**
		(0.021)		(0.018)	(0.005)
Attaining at least 2 years of HE			$\hat{\beta}$ =	0.205***	
				(0.013)	
# obs				29429	
Df				14	

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) In the first column, estimated coefficients on Treat x Post ($\hat{\alpha}$) from Equation 4 are reported. In the second column both estimated coefficients from Equation (5) : ($\hat{\gamma}_1$) on Treat x Post and ($\hat{\beta}$) on the probability to attain at least two years of higher education are reported. Finally, the third column provides the difference between the two estimated coefficients ($\hat{\alpha}$) and ($\hat{\gamma}_1$) from Equations 4 and (5).
(iii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C - 2$ degrees of freedom. The set of control variables is the same as in Table 1

Table 19: Decomposition of the total effect of university openings on log wages with an alternative definition of educational attainment

Dependent variable=log monthly wage of the individual 3 years after having left French education system

Charente Maritime		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	0.048***	$\hat{\gamma}_1$ =	0.012	0.036***
		(0.012)		(0.011)	(0.006)
Attaining at least 2 years of HE			$\hat{\alpha}$ =	0.241***	
				(0.008)	
# obs				10978	
Df				10	

Nord-Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	0.017*	$\hat{\gamma}_1$ =	-0.007	0.024***
		(0.007)		(0.010)	(0.004)
Attaining at least 2 years of HE			$\hat{\beta}$ =	0.221***	
				(0.007)	
# obs				12854	
Df				6	

Pas-de-Calais		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	0.010	$\hat{\gamma}_1$ =	-0.016	0.026***
		(0.006)		(0.009)	(0.004)
Attaining at least 2 years of HE			$\hat{\beta}$ =	0.221***	
				(0.009)	
# obs				9172	
Df				5	

Villes Nouvelles		From eq(4)		From eq(5)	Difference ($\hat{\alpha}$- $\hat{\gamma}_1$ = $\hat{\gamma}_2 \times \hat{\beta}$)
Treat x Post	$\hat{\alpha}$ =	-0.006	$\hat{\gamma}_1$ =	-0.031**	0.025**
		(0.016)		(0.011)	(0.107)
Educational attainment			$\hat{\beta}$ =	0.074***	
				(0.003)	
# obs				17545	
Df				14	

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).
(ii.) In the first column, estimated coefficients on Treat x Post ($\hat{\alpha}$) from Equation 4 are reported. In the second column both estimated coefficients from Equation (5) : ($\hat{\gamma}_1$) on Treat x Post and ($\hat{\beta}$) on the probability to attain at least two years of higher education are reported. Finally, the third column provides the difference between the two estimated coefficients ($\hat{\alpha}$) and ($\hat{\gamma}_1$) from Equations 4 and (5).
(iii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. For the level of significance, I consider $C-2$ degrees of freedom. The set of control variables is the same as in Table 1

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. The Review of Economic Studies, 72(1):1–19.

Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. The quarterly journal of economics, 116(1):313–350.

Anselin, L., Varga, A., and Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. Journal of urban economics, 42(3):422–448.

Becker, G. (1993). Human capital: a theoretical and empirical analysis, whit special reference to education. New York: NationalfeureauofEconomicResearch, Columbia.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. Journal of political economy, 70(5, Part 2):9–49.

Beeson, P. and Montgomery, E. B. (1990). The effects of colleges and universities on local labor markets. Technical report, National Bureau of Economic Research.

Berlingieri, F., Gathmann, C., and Lauber, V. (2017). College openings and local economic development. working paper.

Bordón, P. and Braga, B. (2020). Employer learning, statistical discrimination and university prestige. Economics of Education Review, 77:101995.

Brodaty, T., Ferhat, S., and Gouret, F. (2019). The impact of university openings on human capital formation. Mimeo Thema.

Cantoni, D. and Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. The Quarterly Journal of Economics, 129(2):823–887.

Card, D. (1995). Using geographical variation in college proximity to estimate the return to schooling. Aspects of Labor Market Behavior: Essays in Honor of John Vanderkamp, University of Toronto Press, Toronto.

Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel (CNE) (1996a). Rapport d'évaluation de l'université d'artois. Rapport du CNE.

Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel (CNE) (1996b). Rapport d'évaluation de l'université du littoral. Rapport du CNE.

Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel (CNE) (1997). Rapport d'évaluation de l'université de la rochelle. Rapport du CNE.

Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel (CNE) (2006a). Rapport d'évaluation de l'université de cergy-pontoise. Rapport du CNE.

Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel (CNE) (2006b). Rapport d'évaluation de l'université de marne-la-vallée. Rapport du CNE.

Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel (CNE) (2006c). Rapport d'évaluation de l'université de varsailles-saint-quentin-en-yvelines. Rapport du CNE.

Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel (CNE) (2006d). Rapport d'évaluation de l'université d'evry-val d'essonne. Rapport du CNE.

Délégation à l'aménagement du territoire et à l'action régionale (DATAR) (1998). Développement Universitaire et Développement Territorial: L'Impact du Plan U2000 (1990-1995). La documentation Française.

Gagnol, L. and Héraud, J.-A. (2001). Impact économique régional d'un pôle universitaire: application au cas strasbourgeois. Revue d'Economie Régionale & Urbaine, (4):581–604.

Gelbach, J. B. (2016). When do covariates matter? and which ones, and how much? Journal of Labor Economics, 34(2):509–543.

Goux, D. (2003). Une histoire de l'enquête emploi. Économie et statistique, 362(1):41–57.

Graham, B. S., de Xavier Pinto, C. C., and Egel, D. (2012). Inverse probability tilting for moment condition models with missing data. The Review of Economic Studies, 79(3):1053–1079.

Grenet, J. (2010). La date de naissance influence-t-elle les trajectoires scolaires et professionnelles? Revue économique, 61(3):589–598.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. The review of economic studies, 64(4):605–654.

Kantor, S. and Whalley, A. (2014). Knowledge spillovers from research universities: evidence from endowment value shocks. Review of Economics and Statistics, 96(1):171–188.

Kling, J. R. (2001). Interpreting instrumental variables estimates of the returns to schooling. Journal of Business & Economic Statistics, 19(3):358–364.

Mincer, J. (1958). Investment in human capital and personal income distribution. Journal of political economy, 66(4):281–302.

Mincer, J. (1974). Schooling, experience, and earnings. human behavior & social institutions no. 2.

Monjon, S. and Waelbroeck, P. (2003). Assessing spillovers from universities to firms: evidence from french firm-level data. International Journal of Industrial Organization, 21(9):1255–1270.

Moretti, E. (2004). Workers' education, spillovers, and productivity: evidence from plant-level production functions. American Economic Review, 94(3):656–690.

Psacharopoulos, G. and Patrinos, H. A. (2018). Returns to investment in education: a decennial review of the global literature. Education Economics, pages 1–14.

Ridgeway, G., Kovalchik, S. A., Griffin, B. A., and Kabeto, M. U. (2015). Propensity score analysis with survey weighted data. Journal of causal inference, 3(2):237–249.

Sant'Anna, P. H. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. Journal of Econometrics, 219(1):101–122.

Siegler, B. (2012). The effect of university openings on local human capital formation: Difference-in-differences evidence from germany. Technical report, BGPE Discussion Paper.

Sloane, P. J. M. et al. (2020). Overeducation, skill mismatches, and labor market outcomes for college graduates. IZA World of Labor.

Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M. E., and Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. Health Services and Outcomes Research Methodology, 14(4):166–182.

Valero, A. and Van Reenen, J. (2019). The economic impact of universities: Evidence from across the globe. Economics of Education Review, 68:53–67.

Valletta, R. G. (2018). Recent flattening in the higher education wage premium: Polarization, skill downgrading, or both? In Education, skills, and technical change: Implications for future US GDP growth, pages 313–342. University of Chicago Press.

Vermeulen, K. and Vansteelandt, S. (2015). Bias-reduced doubly robust estimation. Journal of the American Statistical Association, 110(511):1024–1036.