



HAL
open science

The Impact of University Openings and Political Outbreaks on Education and Labor Market Outcomes

Samia Ferhat

► **To cite this version:**

Samia Ferhat. The Impact of University Openings and Political Outbreaks on Education and Labor Market Outcomes. Economics and Finance. CY Cergy Paris Université, 2022. English. NNT: 2022CYUN1094 . tel-03736151

HAL Id: tel-03736151

<https://theses.hal.science/tel-03736151v1>

Submitted on 22 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



École doctorale n° 405 : Économie, Management, Mathématiques, Physique et Sciences
Informatiques (EM2PSI)

Théorie Économique, Modélisation et Applications (THEMA)

Thèse de doctorat

pour l'obtention du titre de
Docteur en Sciences Économiques
délivré par

CY Cergy Paris Université

The Impact of University Openings and Political Outbreaks on Education and Labor Market Outcomes

présentée et soutenue publiquement par

Samia Ferhat

le 10 Juin 2022

préparée sous la direction de **Mme Laurence Jacquet** et de **M. Thomas Brodaty**

Jury

M. Thomas Brodaty	Maître de Conférences, CY Cergy Paris Université, Co-directeur de thèse
Mme Hélène Couprie	Maîtresse de Conférences, Aix-Marseille Université, Examinatrice
Mme Laurence Jacquet	Professeur, CY Cergy Paris Université, Directrice de thèse
Mme Eva Moreno Galbis	Professeur, Aix-Marseille Université, Rapporteur
M. Vincent Vandenberghe	Professeur, Université Catholique de Louvain, Rapporteur

Thèse de doctorat

THE IMPACT OF UNIVERSITY OPENINGS AND POLITICAL OUTBREAKS ON EDUCATION AND LABOR MARKET OUTCOMES

présentée et soutenue publiquement par

Samia Ferhat

le 10 Juin 2022

préparée sous la direction de **Mme Laurence Jacquet** et de **M. Thomas Brodaty**

Jury

M. Thomas Brodaty	Maître de Conférences, CY Cergy Paris Université, Co-directeur de thèse
Mme Hélène Couprie	Maîtresse de Conférences, Aix-Marseille Université, Examinatrice
Mme Laurence Jacquet	Professeur, CY Cergy Paris Université, Directrice de thèse
Mme Eva Moreno Galbis	Professeur, Aix-Marseille Université, Rapporteur
M. Vincent Vandenberghe	Professeur, Université Catholique de Louvain, Rapporteur

Remerciements

Tout d'abord, je tiens à remercier ma directrice de thèse Laurence Jacquet pour ses retours sur mon travail, pour ses conseils, son soutien et sa rigueur qui m'ont fait progresser. Je la remercie d'avoir cru en moi et en ma capacité à aller au bout de ce projet de thèse. Je remercie également Thomas Brodaty, qui a suivi de près mon travail, d'abord pendant mon mémoire de master puis pendant ma thèse. Je le remercie pour ses conseils, ses suggestions, et surtout les nombreuses heures que nous avons passées à travailler ensemble. J'ai beaucoup appris de Thomas et je lui suis reconnaissante de m'avoir donné goût à la recherche.

Je tiens également à remercier Vincent Vandenberghe pour ses précieux commentaires, sa relecture et son suivi qui ont permis d'améliorer considérablement mon chapitre 2. Je le remercie aussi de m'avoir accueilli à l'Université Catholique de Louvain pour un séjour de recherche qui m'a d'ailleurs permis d'avoir des échanges enrichissants avec de nombreux chercheurs, notamment : Joseph Gomes, Jean-François Maystadt et Elsa LeRomain. Un grand merci aussi aux doctorants de UCLouvain : Daniele, Léo, Arno, Mathilde et Esmeralda pour avoir rendu mon séjour très agréable.

Merci à Hélène Couprie et Eva Moreno Galbis d'avoir fait partie de mon jury de thèse. Je remercie Hélène pour notre collaboration dans l'enseignement et l'encadrement des projets des étudiants, ce fut un plaisir de travailler ensemble. Je remercie également Eva que j'ai eu l'opportunité de rencontrer à l'AMSE et qui m'a fait découvrir la version homogénéisée des Enquêtes Emploi.

Merci à mes co-auteurs. Je remercie en particulier Rozenn Hotte, avec qui j'ai beaucoup appris sur l'analyse des données et l'économie du genre et de la famille. Je la remercie également pour ses encouragements et tout le temps passé ensemble. Je suis très heureuse de l'avoir rencontrée et j'espère que nos collaborations se poursuivront. Je remercie également Fabian Gouret, pour nos échanges, son engagement et le travail qu'il a effectué sur l'article. Je remercie

Philip Verwimp, pour ses idées et pour avoir cru en notre projet, ce qui m'a motivé à m'accrocher.

Plus largement, je tiens à remercier tous les chercheurs, qui ont assisté à mes séminaires et qui m'ont fait des suggestions. Je remercie plus particulièrement Maelys de la Rupelle, Pierre André, Nathalie Picard, Pauline Morault et Olivier Charlot. Je remercie aussi Michel Poitevin qui m'a invité au CIREQ de l'université de Montréal, ainsi que les chercheurs que j'y ai rencontrés, notamment Raphael Godefroy et Joao Galindo.

Je remercie CY Cergy Paris Université, le laboratoire Thema et l'école doctorale EM2PSI d'avoir financé ma thèse et mes conférences. Je remercie également le Labex MMEDII qui a contribué au financement de mes données et mon séjour à l'étranger. Merci à Gabriel Desgranges, de m'avoir donné l'opportunité d'enseigner et pour avoir facilité mes missions d'enseignements en termes de planning et matières. Merci au personnel administratif, et plus particulièrement aux secrétaires du Thema : Yolande, Lisa, Imen et Cécile.

Mes remerciements vont à ma meilleure amie Dallal, qui a toujours été présente pour moi malgré la distance. Je la remercie de m'avoir écoutée et rassurée pendant toutes ces années de thèse que nous avons traversées ensemble. Je remercie aussi les thésards et docteurs, avec qui j'ai passé différentes phases de la thèse, pour leur soutien et pour tous les échanges intéressants que nous avons eus : Enxhi, José, Jérémie, Fatoumata, Erwan, Arnaud, Aisha et Mehdi. Je remercie José en particulier avec qui j'ai partagé le bureau et de nombreux moments d'introspection. Je remercie aussi mes amies en dehors du milieu académique : Sara, Sarah et Sonia. Merci à Sara pour les moments que nous avons partagés au Canada avec Amna et la famille Ramadan.

Enfin, je tiens à remercier ma famille, en particulier mes parents qui ont beaucoup sacrifié pour moi. Je remercie ma sœur Ilham d'avoir été là pour moi surtout pendant les périodes où j'étais malade. Merci à Nabil, Imène, Besma et Salah. Merci à mon frère Fouad d'être mon exemple par son intérêt pour le savoir et la science.

Abstract

This thesis explores the impact of university openings on education and labor market outcomes on the one hand, and the impact of political outbreaks on women's marital outcomes on the other hand. In the first two chapters, I use difference-in-differences estimation techniques to explore the effects of university openings in France during the 1991-1993 period, when seven new universities were opened in counties where there were none before. In the first chapter, I find that the opening of these universities increases the likelihood of individuals attaining at least two years of higher education by about 10 percentage points. In the second chapter, I examine whether university openings have an impact on employability and wages. The first two chapters highlight the heterogeneity of the effects of university openings and the important role of the initial characteristics of the areas where universities open. Opening a new university is most effective in areas where higher education provision is lower, and unemployment is higher. In the third chapter of this thesis, I use duration models to assess the effect of conflict on women's marital outcomes, taking evidence from the Arab Spring events in Egypt (2011-2014) where thousands of protesters were killed, injured, or arrested. In this chapter, I examine the impact of these violent events on the age of entry into marriage and motherhood, which are major determinants of women's welfare and labor force participation. I find that women in rural areas are marrying earlier and having children earlier as a result of the Arab Spring in Egypt.

Résumé

Dans cette thèse, j'étudie l'impact des ouvertures d'universités sur l'éducation et le marché du travail français, ainsi que les conséquences du printemps arabe égyptien sur le mariage et la fécondité des femmes. Dans les deux premiers chapitres, j'utilise la méthode des doubles différences afin d'étudier les effets des ouvertures d'universités en France pendant la période 1991-1993, lorsque sept nouvelles universités ont été ouvertes dans des départements où il n'y en avait pas auparavant. Dans le premier chapitre, je montre que l'ouverture de ces universités augmente d'environ 10 points de pourcentage la probabilité pour les individus d'atteindre au moins un niveau d'étude BAC+2. Dans le deuxième chapitre, j'examine l'impact de ces ouvertures sur l'employabilité et les salaires. Les deux premiers chapitres mettent en lumière l'hétérogénéité des impacts des ouvertures de nouvelles universités et leur dépendance aux caractéristiques socio-économiques locales. L'ouverture d'une nouvelle université est plus efficace dans les zones où l'offre d'enseignement supérieur est plus faible et le chômage plus élevé. Dans le troisième chapitre, à l'aide des modèles de durée, j'évalue l'impact du printemps Arabe en Egypte (2011-2014) – où des milliers de manifestants ont été tués, blessés ou arrêtés – sur l'âge des femmes au mariage et à la première naissance, qui sont des déterminants majeurs de leur bien-être et de leur participation au marché du travail. Je montre que les femmes des zones rurales se marient plus tôt et ont des enfants plus tôt à la suite du printemps Arabe en Égypte.

Contents

Remerciements	i
Abstract	iii
Résumé	iv
List of Tables	v
List of Figures	viii
General Introduction	1
1 The impact of university openings on human capital formation	8
1.1 Introduction	9
1.2 The French post-secondary education system and the creation of new universities in France (1991-1993)	14
1.2.1 The French higher education system before 2007	14
1.2.2 The creation of new universities in France (1991-1993)	15
1.3 The main data: The “ <i>Génération</i> ” survey	23
1.4 Objective, method and descriptive statistics	27
1.4.1 ATEs and DiDs for Charente-Maritime	28
1.4.2 Solving for multicollinearity, spillover and substitution effects using alternative groups of control for Charente-Maritime	29
1.4.3 ATEs and DiDs for the other treated counties	31
1.5 Main results	34
1.6 Robustness checks and falsification tests	40

1.6.1	Robustness checks	40
1.6.2	Falsification tests	44
1.7	Conclusion	52
2	The impact of university openings on labor market outcomes	55
2.1	Introduction	56
2.2	Context and Data	59
2.2.1	University openings	59
2.2.2	Data	60
2.3	Empirical Approach	63
2.3.1	Total effect of university openings	64
2.3.2	Total effect decomposition	65
2.4	Results	66
2.5	Robustness	77
2.5.1	Sector of activity	77
2.5.2	Alternative definitions of control group	77
2.5.3	Instrumental Variables	78
2.5.4	Propensity score weighting and Doubly Robust DID	78
2.6	Conclusion	88
	Appendices	89
2.A	Appendix A	90
2.B	Appendix B	92
2.C	Appendix C	96
3	Entry into Marriage, Motherhood and the Arab Spring: Evidence from Egypt	100
3.1	Introduction	101
3.2	Literature review	102
3.3	Conflict setting	104
3.4	Data	106
3.4.1	The ELMPS	106
3.4.2	The Egyptian Revolution database	106

3.4.3	Definition of the Treatment	107
3.4.4	Descriptive Statistics	109
3.5	Estimation approach	110
3.5.1	Graphical Evidence	110
3.5.2	Difference-in-Differences	112
3.5.3	Duration model	114
3.6	Regression results	115
3.6.1	Linear Probability Models	115
3.6.2	Duration Model	119
3.7	Robustness checks and Discussion	121
3.7.1	Alternative definition of the treatment	121
3.7.2	At the district level	121
3.7.3	Alternative specification: discretized duration model	122
3.7.4	Threats to the Identification Strategy	122
3.8	Channels and Interpretation	128
3.8.1	By which period of the Revolution are the results driven?	128
3.8.2	Economic shock versus social norms and religiosity	128
3.9	Conclusion	132
	Appendices	133
3.A	Appendix	133
	General Conclusion	143
	Bibliography	147

List of Tables

1.1	Situation of the counties in 1992 –Main variables–	18
1.2	Assessing potential spillover and displacement effects	32
1.3	DiD estimates of new university openings (<i>concluded on next page</i>)	38
1.3	DiD estimates of new university openings (<i>continued from previous page</i>)	39
1.4	DiD estimates of the Universities of the <i>villes nouvelles</i> : Paris region versus Paris county as a control group	40
1.5	DiD estimates of new university openings, controlling for substitution bias	42
1.6	DiD estimates of new university openings, excluding the wave <i>Génération 98</i> (<i>concluded on next page</i>)	43
1.6	DiD estimates of new university openings, excluding the wave <i>Génération 98</i> (<i>continued from previous page</i>)	44
1.7	Falsification tests: Falsified treatment groups	45
1.8	Falsification tests: 2007-2010 (<i>concluded on next page</i>)	47
1.8	Falsification tests: 2007-2010 (<i>continued from previous page</i>)	48
1.9	Falsification tests: 2004-2007 (<i>concluded on next page</i>)	49
1.9	Falsification tests: 2004-2007 (<i>continued from previous page</i>)	50
1.10	Reverse causality tests: Are university openings caused by the demand for education?	52
2.1	DID estimates ($\hat{\alpha}$)	70
2.2	DID estimates ($\hat{\alpha}$) (adding the field of the last degree attained to the set of control variables)	71
2.3	DID estimates ($\hat{\alpha}$) by type of job	71

2.4	DID estimates ($\hat{\alpha}$) by type of job (adding the field of the last degree attained to the set of control variables)	72
2.5	Decomposition of the total effect of university openings on employment	73
2.6	Decomposition of the total effect of university openings on stable employment	74
2.7	Decomposition of the total effect of university openings on log wages	75
2.8	DID estimates ($\hat{\gamma}_2$) for educational attainment measured by years of schooling	76
2.9	DID estimates ($\hat{\alpha}$) when controlling for the sector of activity	82
2.10	DID estimates ($\hat{\alpha}$) for main variables with different control groups	83
2.11	DID estimates ($\hat{\alpha}$) after integrating propensity score in DID model	84
2.12	cDID estimates ($\hat{\alpha}$) according to Doubly Robust DID method	84
2.13	IV results: dependent variable=1 if the individual is working 3 years after having left the French education system	85
2.14	IV results: dependent variable=1 if the individual has a stable employment 3 years after having left the French education system	86
2.15	IV results: dependent variable is log monthly wage 3 years after having left the French education system	87
2.16	Decomposition of the total effect of university openings on employment with an alternative definition of educational attainment	97
2.17	Decomposition of the total effect of university openings on stable employment with an alternative definition of educational attainment	98
2.18	Decomposition of the total effect of university openings on log wages with an alternative definition of educational attainment	99
3.1	Geography of the Revolution	106
3.2	Descriptive statistics	109
3.3	LPM Model: Marriage before 23 years old	117
3.4	LPM Model: Child before 23 years old	118
3.5	Marital outcome at the governorate level, for rural women - Discrete time duration model	120
3.6	Marital outcome at the governorate level, rural women -Treatment 3 and Treatment 4- discrete time duration model	125

3.7	Marital outcome at the district level - discrete time duration model, for rural women	126
3.8	Alternative specification	127
3.9	Duration model: Occurrence of the event. Age at marriage and first birth according to Revolution period in rural areas - discrete time duration model . . .	130
3.10	Impact of the revolution on the probability of marriage according to the degree of conservatism in rural areas (discrete time duration model)	131
3.11	LPM Model: Age at marriage	134
3.12	LPM Model: Age at first child	135
3.13	Marital outcome at the governorate level, for ALL women - Discrete time Duration model	136
3.14	Marital outcome at the governorate level, for URBAN women - Discrete time Duration Model	137
3.15	Marital outcome at the governorate level, for rural women - Cox duration model	138
3.16	Marital outcome at the governorate level, for rural women - Current place of residence, discrete time duration model	139
3.17	Outcomes at the governorate level for non migrant rural women between 2011-2018, discrete time duration model	140
3.18	Logit Model: Marriage before 23 years old	141
3.19	Logit Model: Child before 23 years old	142

List of Figures

1.1	Counties with new university openings (1991-1993)	20
1.2	Growth rate of student enrollment at the university (relative to 1986, according to university location)	21
1.3	Economic and social trends according to location (relative to population size) .	22
2.1	Individuals assignment into treated and control groups	62
2.2	Individuals assignment into treated and control groups (alternative control group1)	80
2.3	Individuals assignment into treated and control groups (alternative control group2)	81
3.1	Number of people killed during the protests for 1000 inhabitants by governorates	108
3.2	Geographic distribution of people killed during the protests for 1000 inhabitants by governorates	108
3.3	Average age at marriage	111
3.4	Probability of marriage before 23	111
3.5	Probability of marriage before 23 for rural women	113

General Introduction

This thesis investigates the consequences of two different events on individuals' outcomes in terms of education, employment, wages and marriage. The first event consists in university openings (UO) in France between 1991-1993, where seven new universities opened in places where there were none before. The second event is the Arab Spring in Egypt where thousands of protesters were killed, injured or arrested between 2011-2014. The first two chapters of the thesis are dedicated to the impact of the first event (University openings) on educational attainment, employment and wages. The third chapter is dedicated to the impact of the second event (Arab Spring) on women entry into marriage and motherhood.

Universities are the driving force behind the production of a highly skilled workforce in any country. In a context where economic prosperity relies on the ability of a country to produce competitive and reliable services (health, information technology, energy, etc.), the role of universities as a place for training and producing academic research has become increasingly important. While it is well established that universities significantly contribute to the accumulation of human capital, much is less known about the marginal impact of opening new universities in a given region/country. This especially leads to questioning how such a contribution can be measured and identified.

Between 2000 and 2018, the global higher education gross enrolment ratio increased from 19% to 38% (UNESCO, 2020). Countries worldwide are experiencing an increasing demand for higher education mainly due to the returns higher education entails. The latter vary across countries and tend to be higher in developing countries. For instance, in a case study of the expansion in public universities in Ethiopia, Somani (2021) shows that higher educational attainment almost doubles the probability of paid employment and almost doubles hourly wages. In the case of developed countries such as the United States, a study from the Georgetown University Center on Education and the Workforce shows that the difference in earnings between

those who attend college and those who do not is continuously growing (Carnevale et al., 2011). The same study shows that the average lifetime earnings in 1999 of someone who holds a bachelor degree was 75% higher compared to someone who holds a high school degree. In 2009 the average lifetime earnings for bachelor degree holders is 84% higher relative to high school degree holders.

In this context of increasing demand for higher education, it is crucial for policymakers to assess whether existing universities are able to meet this growing demand. The first block of this thesis is dedicated to study the impact of opening a new university in a place where none existed before on the local population. The answer to this question as simple as it may sound, is not straightforward. In fact, one might think that opening a new university would have a positive impact on both individual's educational attainment and labor market outcomes, regardless of the location where the university is created. I show that the impact is very sensitive to the choice of location where the new university is implemented.

In particular, I study the case of university openings in France where a total number of seven universities were opened in counties where none existed before between 1991-1993. These universities are: Cergy-Pontoise in the Val-d'Oise county, Marne-la-Vallée in the Seine-et-Marne county, Evry Val-d'Essonne in the Essonne county, and Versailles Saint-Quentin-en-Yvelines in the Yvelines county. These four universities are called *Universités des villes nouvelles* (UVN) because they were implemented in the four new cities that were created by the French government in 1965 to allow a multipolar urban development. Two other universities were opened in Nord and Pas-de-Calais counties: Artois and Littoral Opal Coast. Finally, La Rochelle university opened in Charente Maritime county on January 1993. The motives for focusing on France in this thesis are mainly related to the fact that there are no heavy costs on the entry into universities. Tertiary education in France is mostly public with relatively low tuition fees. Another reason is the perfect framework these UO set for a difference-in-differences analysis: before openings and after. Treatment group – individuals in counties where university openings occurred – and control group – individuals in counties that have not witnessed university openings on the period studied. Also, France features one of the highest youth unemployment rates compared to the average of OECD and European Union (EU) countries combined.¹ Therefore, it is appealing to study how

1. For the year 2018, youth unemployment rate was at 22% in France whereas it was only at 11.7% for the average of OECD countries and 16.5% for the average of EU countries (OECD, 2022).

such a reform, which is specifically designed for youth, would affect their employability and career prospects.

The second block of this thesis (chapter 3) is dedicated to investigate the impact of political outbreaks on entry into marriage and motherhood. Political outbreaks can result in persistent civil conflicts, with devastating economic and societal implications. The MENA region (Middle East and North Africa) has witnessed during the previous half decade more conflicts than any other part of the world (Rother et al., 2016). Notable examples are the Algerian Civil War (1990-1998), the Gulf War (1990-1991), and the 2008 Lebanon conflict. The consequences of the Arab Spring conflict, which began on December 10, 2010 in Tunisia with the self-immolation of a street vendor "Mohamed Bouazizi" whose merchandise was confiscated by the police, have left their mark. The spread of the damage reached the western world as well. The Islamic State of Iraq and the Levant (ISIL) was formed causing insecurity in both MENA and Western world. Refugees fled the region towards mainly Europe and Turkey, leaving behind their homes and loved ones.

In traditional societies such as those prevailing in MENA region, women are financially dependent on men and gender inequalities are the highest in the world. According to the Global Gender Gap Index (GGGI) of Sharma et al. (2021) – measured in four dimensions: economic participation and opportunity, educational attainment, health and survival, and political empowerment – the MENA countries have the worst performance in closing the gender gap. In terms of female labor force participation (FLFP), Ghazalian (2022) shows a moderate increase in FLFP as a result of Arab Spring events. In fact, two opposite effects on FLFP can be distinguished in a context of conflict. On the one hand, the uncertainty caused by the conflict may incentivise women to participate in the labor force to support their families; this is the "added worker effect hypothesis" suggested by Lee and Cho (2005). The latter hypothesis is dependent on the level of discrimination against women in the country of interest in the labor market, given that women are perceived to have lower productivity compared to males. On the other hand, political and economic instabilities can also discourage women to participate in the labor market because of the resulting rise of conservatism and religious groups in society. Chaney (2013) argues that in times of economic recession, religious authorities gain power, which sort of describes what happened in Egypt as the Muslim Brotherhood took power in the heat of the conflict. Nonetheless, the effect of political instability on women's welfare is not only limited to FLFP as

other aspects are also affected such as health, intra-household bargaining power, and marriage (see Williams et al. (2012), among others). This chapter is the first contribution to focus on marriage dynamics as a result of the Arab Spring.

In particular, marriage is a source of security to women, even more so in patriarchal societies where they are financially dependent on men. As noted by a recent report of the UN Women (2017) in the chapter 3 regarding Egypt:

"Almost 80 per cent of men are the main breadwinners for their family. More than 60 per cent of male respondents worried about their ability to provide their families with daily necessities, among other concerns."

In the context of protests, political and economic instability, marriage becomes even more of a necessity to counterbalance the insecurity created by Arab Spring. In this dissertation, I am interested in empirically assessing how the political unrest in Egypt during the period 2011-2014 affected women's marital outcomes (marriage and motherhood). In contrast to existing evidence that focuses only on the short-term effects of the Egyptian Arab Spring, this chapter draws on sufficient historical background to capture both the short- and medium-term effects. My findings document that the Arab Spring in Egypt accelerates marriages and entry into motherhood in rural areas, while this effect does not hold in urban ones. Interestingly, I show that this acceleration in entry into marriage and motherhood is higher in conservative areas. This is consistent with the security hypothesis: parents would want to marry their daughters earlier to transfer responsibility to the spouse when economic conditions are uncertain.

Outline of the thesis

The thesis consists of three chapters in the form of research papers. The first two chapters focus on university openings, and the third chapter examines the impact of the violent events of the Arab Spring in Egypt on women's marital outcomes.

Chapter 1: The impact of university openings on human capital formation

This chapter is co-authored with Fabian Gouret (CY Cergy Paris Université) and Thomas Brodaty. We study the effect of the opening of new universities on educational attainment and focus on the reform of the French higher education system where seven universities have been opened in places where no universities existed before. To this end, data from the CEREQ *Génération* surveys (1992, 1998, 2001, 2004 and 2007) are mobilized. Each wave of the survey is a representative sample of young individuals who left the French education system the year of the wave. The data provides detailed information on the last degree obtained by each individual and a set of individual characteristics (gender, father's labor market status, parents' place of birth and age of entry into middle school). Our identification strategy is based on difference-in-differences where the 1992 wave is considered the pre-treatment wave, and the remaining waves (1998, 2001, 2004 and 2007) are the post-treatment waves. We assign individuals into treatment and control groups based on their location at the date of entry into middle school. The individuals treated are those who were living at the time of their entry into middle school in a county where a university opening has opened. Individuals in the control group are those who lived, at the time of entry into middle school, in counties where there were no new university openings during the studied period. Different definitions of the control groups are used to address potential spillover and displacement effects. Our main findings are that opening a new university increases the probability to attain at least two years of higher education by about 10 percentage points. Opening new universities also increases the probability to attain at least four years of higher education by about 4 percentage points. These results are found in areas that are initially under educated compared to the rest of the country. We do not find an effect of university openings in counties surrounding Paris characterized by a high supply of higher education compared to the rest of France. We argue in favor of a catch-up effect hypothesis to explain our results, wherein counties with a lower initial level of tertiary education catch up with the most educated counties

in the same country following the opening of universities. To strengthen our results' validity, we show that they hold even with falsified treatment counties and treatment periods.

Chapter 2: The impact of university openings on labor market outcomes

In this chapter, I study the impact of the French reform of the higher education system by opening seven new universities in France. The aim of this paper is to assess the impact of these new universities on labor market outcomes for local population. I use data from CEREQ *Génération* Surveys (waves 1992, 1998, 2001, 2004 and 2007) and “*Les Enquêtes Emploi*” (EE) the French Labor Force Surveys (waves 1995, 2001, 2004, 2007 and 2010). Cereq *Génération* Surveys provide data on each respondent's labor market outcomes three years after leaving the French education system. Individuals' educational attainment, field of studies and some other characteristics (gender, father situation on the labor market, parents' place of birth and age of entry into middle school) are also provided. I measure labor market outcomes three years after leaving the French education system by: the probability of being employed; the probability to have a stable job (permanent contract or work as official) and wages. Difference-in-differences (DID) estimates show heterogeneous results according to the location of the opening. Further, I show that opening a new university increases the probability of being employed by about 8% points in counties where individuals come from disadvantaged backgrounds, and wages increase by about 5%. No impact is found in counties that are relatively more educated and where the unemployment rate is low. This last result is consistent with the main conclusion of the first chapter: in areas where the level of higher education is initially relatively high, we have no impact on human capital formation. Two channels through which universities impact labor market outcomes are discussed: the human capital channel and the spillover channel. The human capital channel consists in the positive effect educational attainment has on improving individuals employability and earnings. The spillover channel is more complex and can be positive, negative or nil. Negative spillovers can originate from over education- skill downgrading and/or employer learning-statistical discrimination. I use alternative methods to standard DID: propensity score weighting à la [Stuart et al. \(2014\)](#), Doubly Robust DID ([Sant'Anna and Zhao, 2020](#)) and Instrumental variables, and I find consistent findings giving confidence in results validity.

Chapter 3: Entry into Marriage, Motherhood and the Arab Spring: Evidence from Egypt

This paper is co-authored with Rozenn Hotte (Université de Tour) and Philip Verwimp (Université Libre de Bruxelles). We study the effect of Arab Spring violent events in Egypt on women age at marriage and age of entry into motherhood. Arab spring is the series of protests in the Arab world that started in Tunisia on December 10, 2010 and were met by violence from governments. In Egypt, protests took place between January 25, 2011 and January 31, 2014. We use data from wave 2018 of the Egyptian Labor Market Panel Surveys (ELMPS), and combine them with data from “*Wikithawra*”. The latter provides information on victims of the Egyptian revolution (people killed, arrested and injured) and gives the exact date and location at which each violent event took place. This allows us to have a measure of the revolution intensity by computing the share of victims in each district and governorate in Egypt. Empirically, we use difference-in-differences that we complement by duration model. Two main treatment definitions are considered in our analysis; on the governorate level: a dummy variable equals to 1 for the 6 governorates with the highest share of killed individuals during the period of Egyptian revolution; and a continuous measure which is the share of killed individuals. We find that women who lived in rural areas at the time of the revolution marry earlier compared to previous cohorts, and have children earlier as well. We show that the effect is more pronounced in the most conservative areas. As robustness checks, alternative definitions of the treatment are implemented, computing the treatment at the district level, among others. We also investigate whether the results are driven by the arrival of the President Morsi (who is a tenor of the Muslim Brotherhood) and find that the results do not seem to be driven by a particular phase of the Egyptian revolution.

Chapter 1

The impact of university openings on human capital formation

Abstract

This paper presents new evidence on the impact of university openings on the acquisition of human capital by the local youth in France. We exploit seven university openings between 1991-1993 in counties where no previous universities existed that we combine with five waves from representative outflow samples of young individuals leaving the French educational system at the end of their studies. We take advantage of specific control groups to compute differences-in-differences estimates that are robust to displacement, spillover and substitution effects. Our DiD outflow estimator identifies the underlying and policy relevant inflow treatment effect on birth cohorts under mild conditions that are consistent with the data. We find that opening a new university significantly increases the probability of attaining at least a two-year post-secondary degree by about 10 percentage points in counties that are initially undereducated compared to the rest of France. Conversely, university creations which occur in relatively educated counties does not have a significant impact on the acquisition of human capital. We argue in favor of a catchup effect, in which university openings help undereducated counties to converge to the average level of higher education in a given country.

JEL classification: I21, I23, C21

Keywords: university openings, human capital, differences-in-differences, outflow sampling.

1.1 Introduction

In the years following World War II, university expansion accelerated in most countries. Documenting this expansion in 1500 regions across 78 countries between 1950 and 2010, Valero and Van Reenen (2019b) show that increases in the number of universities are significantly and strongly associated with future growth of GDP per capita. One should expect that this university effect on growth works in part through the acquisition of human capital (measured by college share or years of education), as the empirical literature on growth would suggest (e.g., Mankiw et al., 1992). However, Valero and Van Reenen (2019b, p.26) point out that adding “*the lagged growth in college share which in itself is highly significant [...] reduces the university effect [on growth of GDP per capita] from 0.0710 to 0.0672.*”¹ That is, only a very small part of the university effect on growth is mediated through an increased supply of human capital. In fact, regressing the growth in college shares on the lagged growth in the number of universities (controlling or not for a host of variables), they do find a statistically significant but negligible

1. These estimated coefficients mean that a doubling of universities in a region is associated with 7.1/6.72% higher GDP per capita.

relationship. These results may appear paradoxical, given that the opening of a new university in a region reduces the access costs, and should increase investment in higher education by the local population. Thus, Valero and Van Reenen (2019b, p.66) conclude that this small role for human capital in the university effect calls for further investigation “*in individual countries where better causal designs and richer university data [are] available*”.²

With these elements in mind, this paper presents new evidence on the impact of new university openings on the acquisition of human capital by the local youth in France, using detailed microeconomic data. France is an interesting country because, between the end of 1991 and 1993, it opened seven universities in different *départements* -counties hereafter- where no previous universities existed.³ Moreover, most French universities are state-owned, and the annual fees are relatively low.⁴ The new universities have not differed from the existing ones concerning these fees. Given that they have implied a sharp decrease in housing and transportation costs for the local youth, the investment in higher education should have increased.

Thus we exploit these creations of universities that we combine with different waves of the *Génération* survey, a large scale survey on educational attainment conducted by the *Centre d'Etudes et de Recherches sur les Qualifications* (CEREQ hereafter). Each wave is composed of individuals who left the French educational system at the end of their studies the year of the wave. Basically, the first wave (1992) concerns individuals who entered their last degree in 1991, i.e., before the creation of the new universities, and left the educational system at the end of their studies in 1992; the other waves (1998, 2001, 2004 and 2007) concern individuals who left the educational system after all new universities were created. Section 2.2.2 provides more detailed information on this survey and discusses, amongst other topics, the particular case of some new universities created at the end of 1991.

We use these different waves to compute differences-in-differences -DiD- estimates of the

2. The human capital channel through which universities may affect growth is not the sole channel studied by Valero and Van Reenen (2019b). In particular, they also study an innovation channel (measured by patents), as the growth literature would suggest. They do find a role for innovation which is (again) small in magnitude. This small effect calls for further investigation at the micro level as well, but it is beyond the scope of this paper.

3. A *département* is one of the three levels of government below the national level, between the administrative regions and the *commune*, the latter being broadly equivalent to a civil township. There are 96 *départements* in metropolitan France.

4. The annual fees are less than 300 euros now, and were less than 1000 francs in the nineties. To realize that these fees have been low, note that the median monthly wage in France in 1996 was 8540 francs; the median monthly wage in 2005 was 1528 euros. Furthermore, there have been some grants for students whose parents have had too small wages, and these students have not had to pay the fees.

effect of a new university opening in a county on different variables; in particular -but not only- the probability of attaining at least a two-year post-secondary degree, and the probability of attaining at least a four-year post-secondary degree. The key identifying assumption of the impact of a new university opening is that a county with a new university and a control group of counties would have experienced the same change in the probabilities of interest without any creation –i.e., the common trend assumption. Historical data from population Censuses before the creations suggest that this untestable assumption is plausible; see Subsection 1.4.2 for more details. Even if this assumption is not totally satisfied, our regression approach of the DiD model may permit to relax it partially by controlling for various county-specific trends that would have affected differently the change in probabilities.

Although they have not focused on the causal effect of new university openings on the acquisition of diploma per se, various authors have used the presence of a college in a county or in a city as an instrument to consistently estimate returns to education since [Card \(1995a\)](#). That is, their interest is in the causal effect of years of schooling upon earnings, not on the causal effect of a nearby college on years of schooling; their reduced forms where years of schooling is explained by the presence of a nearby college are only best linear approximations of years of schooling. However, their results suggest that proximity to college matters for years of schooling. For instance, [Card \(1995a\)](#) exploits the young men's cohort of the National Longitudinal Survey, and his different reduced forms suggest that those who live near a four-year college have on average 0.32-0.38 more years of schooling; see also [Card \(2001\)](#), [Kling \(2001a\)](#), [Moretti \(2004a\)](#), as well as [Currie and Moretti \(2003\)](#) who use proximity to college as an instrument to consistently estimate the causal effect of maternal education on birth outcomes.

To our knowledge, [Siegler \(2012a\)](#) is the sole author who tries to measure the causal effect of new university openings on the acquisition of post-secondary diploma by the local youth.⁵ He uses new university openings in Germany during 1960-1979 that he combines with the 2009 wave of the German Socio-Economic Panel (SOEP) to estimate the effect of local university access

5. A related article is [Frenette \(2009\)](#). He studies the impact of a local university on the completion of at least one full year of university or of another postsecondary institution, using the Canadian quinquennial Census of Population microdata files for the years 1981-2001. He finds that the presence of a local university in a census metropolitan area increases the university attendance by 6.4 percentage points and decreases the attendance to other postsecondary institutions by 5 percentage points. There is also a connected literature which focus on the causal effect of the proximity of high schools on individual schooling decisions in developed economies (e.g., [Falch et al., 2013](#), [Garrouste and Zaiem, 2017](#)), or on the impact of primary school openings in developing countries (e.g., [Burde and Linden, 2013](#), [Duflo, 2001](#)).

on the probability of obtaining a university degree. He finds that the probability to graduate for a young rises by 8 to 10 percentage points when a new university is established in his county. However, two problems arise in the paper, first since Siegler (2012a) uses the place of residence of each individual in 2009 instead of before the creation of a university, this constitutes a problem because of a possible mobility in-between. A second problem concerns the nature of his data: basically, it is a cross-section, and he cannot control for some county-specific trends, like, e.g., a demographic trend more pronounced in some counties.

In the current paper, each wave of the *Génération* survey is composed of individuals who left the educational system the year of the wave. By considering the five waves of this survey (1992, 1998, 2001, 2004 and 2007), we control for various county-specific trends. The county of residence at the date of entry to middle school is used to assign individuals into treatment and control groups. This variable is valid as it is a priori not influenced by the university opening. Treated individuals are those who lived when they entered in middle school in a county which experienced a university creation. Individuals in the control group are those who lived when they entered in middle school in a county that didn't experience a university creation through the studied period.

Four different control groups are systematically studied, embracing a large set of all possible controls. Neighboring counties constitute the first two control groups. They may provide a similar environment due to geographic proximity. However, as Gobillon and Magnac (2016) point out, a treated county may be closer to more distant counties in some dimensions which depart from purely geographic characteristics.⁶ Furthermore, considering neighboring counties as the control group may strongly underestimate the impact of a university creation because of two potential biases. The first one is a spillover effect: a new university may have a positive impact on students coming from neighboring counties. The second bias is a (control group) substitution effect: if some universities that were already existing in the control group are extended, or, more generally, if an increase in the education supply in the control group occurs. To solve for these biases, we consider a third control group, the Paris region (i.e., the county of Paris and its bordering counties). We expect that its youth was not impacted by the creation of

6. Note also that even with permanent unobservable differences between treated and distant counties, the latter may constitute an appropriate control group in a DiD setting if the common trend assumption is satisfied, or if we are able to control for some county-specific trends that would have affected differently the change in probabilities of enrollment into the higher education system.

new universities in the treated counties, given that the Paris region is characterized by the highest local per capita supply of higher education in France. If so, it prevents from any spillover bias. It is also more able to prevent from a substitution bias, given that no strong shock in the supply of higher education happened in the Paris region over the studied period. Finally, we investigate as a last control group all the non treated counties. It is a basis for comparisons with a control group common to all treated counties but may be affected by spillover and substitution effects.

It is also prudent to emphasize that our estimates are net from any displacement effect. Indeed, some students enrolled into a new university would have enrolled anyway in the absence of any creation; perhaps into a university in a bordering county and at a higher cost. However, the different left-hand side variables that we consider are probabilities of enrollment into *any* institution or *any* university of the higher education system; they are not probabilities of enrollment into a specific *new* university. By doing so, we cancel out potential displacement effects.

Using our various control groups, we find that opening a new university significantly increases the probability of attaining at least a two-year post-secondary degree by about 10 percentage points in counties that are initially undereducated compared to the rest of France. Conversely, university creations which occur in relatively educated counties does not have a significant impact on the acquisition of human capital. We argue in favor of a catchup effect, in which university openings help undereducated counties to converge to the average level of higher education in a given country. Finally, we propose various falsification tests, with falsified treatment counties and treatment periods, and we also address the question of reverse causality. Despite of some drawbacks, these tests make us confident in the size, significance and causal interpretation of our results.

The rest of the paper is laid out as follows. Section 2.2 describes the French higher education system and the creation of new universities which took place in this country between 1991-1993. Section 2.2.2 presents the *Génération* survey and the main variables of interest for our study. Section 1.4 provides some preliminary statistics and the methodology. Section 1.5 describes the main results. Section 1.6 provides various robustness checks and falsification tests. Following all of this, Section 2.6 concludes.

1.2 The French post-secondary education system and the creation of new universities in France (1991-1993)

1.2.1 The French higher education system before 2007

After graduating from high school and getting the *Baccalauréat* (grade 12), the final high school diploma required to pursue higher education, an individual can attend the university⁷, or attend another postsecondary tracks (usually with selective admission). These other post-secondary tracks include two-year vocational degrees called *Diplôme Universitaire Technologique* (DUT, prepared in *Institut Universitaire de Technologie* -IUT-) or *Brevet de Technicien Supérieur* (BTS, prepared in some high schools even if the BTS is a post-secondary degree), or a *Classe Préparatoire aux Grandes Ecoles*, two intensive years whose goal is to train undergraduate students for passing the competitive exams to enter in *Grandes Ecoles*. Then the *Grandes Ecoles* deliver usually after three years a five-year degree mainly in engineering or in business.

Before 2007, the studies in *all* the French universities were organized in three cycles. A first cycle of two years permitted to obtain a *Diplôme d'Etudes Universitaires Générales*; a student was allowed to repeat only one year and obtained this two-year university degree in three years. This two-year university degree was not a vocational degree but a general one in the major chosen by the student (e.g., mathematics, economics, law...); obtaining this degree was necessary to enter into the second cycle of two years. At the end of the first year of the second cycle, a student obtained a *Licence* in his major. This three-year university degree was compulsory to enter in the last year of this second cycle; this last diploma, called *Maîtrise*, corresponded to a four-year degree. This degree was necessary but not sufficient to be admitted in the selective third cycle of a university to obtain a *Diplôme d'Etudes Supérieures Spécialisées* (DESS), i.e., a one-year vocational master's degree, or a *Diplôme d'Etudes Avancées* (DEA), i.e., a master of research which was the first year of doctoral studies.

Note that like the DEUG, the two-year vocational degrees (BTS and DUT) and the two years

7. The universities are usually state-owned in France. However, there are seven private catholic institutes which provide the same diplomas and which are called "university": University and Polytechnic Federation of Lille, the Catholic University of Paris, the Catholic University of Toulouse, the Catholic Institute of Higher Studies (in the Vendée county), the Catholic University of the West (in the Maine-et-Loire county), the Catholic University of Rennes, the Catholic University of Lyon, and the Domuni University (a distance education institute located in Toulouse). The annual fees in these universities are more expensive than in the state-owned ones (between 3000 and 9000 euros depending on the wage of the parents, the major chosen, and the cycle of study).

of study in *Classes Préparatoires* corresponded to a first cycle of studies. The five-year degrees in *Grandes Ecoles* were third cycle diplomas, like the DESS, the DEA and the doctoral studies. These equivalences in terms of cycles permitted some connections between the university system, vocational colleges and *Grandes Ecoles*. For instance, good students who obtained a two-year vocational degree had the possibility to enter in a second cycle at the university.⁸

The organization in cycles with a first cycle corresponding to a two-year degree, a second cycle corresponding to a four-year degree, and a third one corresponding to a five-year degree or more was still prevalent until the French Law on Liberties and Responsibilities of Universities of August 2007.⁹ Since then, the *Licence* has corresponded to a first cycle of studies, a two-year master in universities and the five-year degree in *Grandes Ecoles* to a second cycle, and the doctoral studies to a third cycle. These changes in 2007 followed the implementation of the Bachelor's Master's Doctorate system designed by the Bologna Process whose aim was to standardize the educational system in Europe. In fact, the French universities suppressed progressively between 2004-2007 the DEUG, the *Maîtrise*, the DESS and the DEA to implement the Bologna system.

1.2.2 The creation of new universities in France (1991-1993)

On May 23 1990, the French Council of Ministers decided to adopt the plan *Université 2000* (U2000). At the beginning of the nineties, France expected 300000 additional students by the year 1994 because of a growing population between 18-24 years old (DATAR, 1998, pp.14-15).¹⁰ An objective of U2000 was also to favor higher education for some sub-populations that were less educated than the rest of the French population: it was explicitly written that students who grow up in a county without a university face a higher cost of education, and this cost had to be reduced (DATAR, 1998, p.55). For these reasons, it was decided to extend some existing universities, as well as create new ones over the period 1991-1993 in counties where no university previously existed or where the higher education supply was low.

8. According to the [Ministère de l'Éducation Nationale \(2001\)](#), 19.8% of the students who obtained a DUT in 1998 entered in a second cycle at the university the following academic year. Students with a DUT who did not have all the prerequisites to enter in a specific second cycle were usually asked to study in a first step the second year of a first cycle at the university. 8.2% of those who obtained a DUT in 1998 entered in a first cycle at the university the following academic year. The data are broadly similar for those who obtained a DUT in 1999.

9. This law has also given more financial autonomy to the universities.

10. The DATAR (*Délégation à l'Aménagement du Territoire et à l'Action Régionale*) was the French state administration in charge of land planning policies between 1963 and 2014.

Figure 1.2 depicts the growth rate (relative to 1986) of student enrollment at the university, according to its location. The growth rate of enrollment in counties where creations occurred, i.e., the treated counties, is clearly higher than the one wherein no creation occurred (i.e., in the rest of France), in absolute and relative terms. Note that enrollment is not absolutely null in the treated counties before the creations; the reason, as described below, is that there was sometimes a small satellite campus of another university in some treated counties. Concerning the evolution of student enrollment of other subgroups of counties, note that these subgroups will be the different control groups of our econometric analysis; see Section 1.4.

Seven universities were created between 1991-1993. An eighth university, the University of Southern Brittany, was created in 1995 in Morbihan, a county of French Brittany. But this university was not new *stricto sensu* because its campuses were already satellites of various other universities of French Brittany.¹¹ Given that this university was *de facto* already existing, we do not include it in our analysis. No university was created afterwards in France.

Of the seven universities of interest, four universities were created by Decree-Law of July 22, 1991, and opened in October 1991 in counties around the Paris region as shown in Figure 1.1: Cergy-Pontoise University in the Val-d'Oise county, University of Marne-la-Vallée in the Seine-et-Marne county, University of Evry Val-d'Essonne in the Essonne county, and Versailles Saint-Quentin-en-Yvelines University in the Yvelines county. These four universities are called *Universités des villes nouvelles* (UVN) because they were implemented in the four new cities that were created by the French government in 1965 to permit a multipolar urban development. According to the reports of various public administrations involved in the creation and the evaluation of these universities (DATAR, 1998, p.35; CNE, 2006a, p.I, 2006d, p.5, 2006b, p.3; *Cour des comptes* 2008, p.156), the aim of these new universities was to decongest the universities in Paris.¹² The first promotion of students began their two-year university degree

11. More precisely, the University of Rennes 1 (located in the Ille-et-Vilaine county) created in Vannes (one of the city of the Morbihan county wherein the University of Southern Brittany is implemented) a DEUG in law in 1969, a DEUG in sciences in 1990, and a DEUG in economics and management in 1992. Various *Instituts Universitaires de Technologie* were also existing since the 70s in Vannes (as satellites of the *Institut Universitaire de Technologie* of Rennes and University of Rennes 2) and provided two-year vocational degree in business and accounting and two-year vocational degree in statistics and data analysis. Lorient (another city of the Morbihan county wherein the University of Southern Brittany is also implemented) has had its own *Institut Universitaire de Technologie* since 1975 providing a two-year vocational degree in thermal engineering and energy.

12. The CNE (*Comité national d'évaluation des établissements publics à caractère scientifique, culturel et professionnel*) was a public administration to which the government had devolved power to evaluate public higher education institutes in the fields of sciences, culture and professional education between 1984 and 2006. Concerning the *Cour des comptes*, it is an administrative body charged with conducting financial and legislative audits of public

(i.e., DEUG) in October 1991, and the students of this first promotion who did not repeat any year obtained it in June 1993. Two exceptions concern Evry and Cergy-Pontoise. In the county of Essonne where the University of Evry Val-d'Essonne was created, a degree in Science had been delivered by the Orsay campus of the University of Paris-Sud since 1958. Concerning Cergy-Pontoise, the Ministry of Higher Education sent a team of professors in 1990 also from the University of Paris-Sud to create a DEUG in physics in the Val-d'Oise county. The courses were taught in the building of a *Grande Ecole* in electrical engineering (*Ecole Nationale Supérieure de l'Electronique et de ses Applications*). This DEUG became a degree of Cergy-Pontoise University in 1991.¹³

By Decree-Law of November 7, 1991, Artois University and University of the Littoral Opal Coast were officially established in two bordering counties, the Pas-de-Calais and Nord counties; these two universities opened in October 1992. Since then, Artois University has had sites in various cities of Pas-de-Calais (Arras, Béthune, Douai, Lens and Liévin), like University of the Littoral Opal Coast (Boulogne, Calais and Saint-Omer) which has also one site in a city of Nord (Dunkirk). So these two creations mainly took place in Pas-de-Calais where no previous university existed. Concerning the Nord county where the site of Dunkirk was opened, note that various universities already existed in the city of Lille.¹⁴ The different reports highlighted that the population of these two counties was under-educated: in 1985, only 25.7 percent of the concerned age-group obtained the *Baccalauréat* (Thumerelle, 1998, pp.89-90). Descriptive statistics in Table 1.1 show that under-education is even more pronounced in the PC county.

institutions.

13. Note that at the creation of Cergy-Pontoise University in 1991, the two main buildings of the university, *Les Chênes* (in Cergy) and *Saint-Martin* (in Pontoise), did not exist. The construction of *Les Chênes* was finished in October 1992 and the one of *Saint-Martin* in 1994. Given the absence of buildings at the creation of this university, the first year group of students began the first year of their DEUG in some premises of the Council of the Val-d'Oise county. These premises were unsuitable for big year group. The CNE (2006d, p.I) also highlights that the University of Evry Val-d'Essonne experienced a slow and chaotic construction of its main buildings; this university had to rent some premises of the postal service which were unsuitable for big year group.

14. These universities were: Lille 1 University of Science and Technology, Lille 2 University of Health and Law, Charles de Gaulle University-Lille 3, and the private catholic University and Polytechnic Federation of Lille.

Table 1.1 – Situation of the counties in 1992 –Main variables–

	Mean (Standard-deviation)					
	Charente-Maritime (<i>c = CM</i>)	Pas-de-Calais (<i>c = PC</i>)	Nord (<i>c = N</i>)	Villes Nouvelles (<i>c = VN</i>)	Rest of France	Paris Region
Education variables						
Attaining a two-year post-secondary degree or higher diploma	0.163 (0.370)	0.230 (0.421)	0.332 (0.471)	0.361 (0.480)	0.354 (0.478)	0.514 (0.499)
Attaining a four-year post-secondary degree or higher diploma	0.059 (0.236)	0.062 (0.241)	0.088 (0.283)	0.169 (0.375)	0.127 (0.333)	0.269 (0.444)
Last socio-professional category of the individual's father when he or she left the education system						
Farmer	0.092 (0.289)	0.039 (0.194)	0.023 (0.150)	0.005 (0.068)	0.069 (0.254)	0.003 (0.053)
Storekeeper	0.158 (0.366)	0.059 (0.236)	0.077 (0.267)	0.095 (0.293)	0.118 (0.323)	0.116 (0.320)
Executive	0.070 (0.256)	0.093 (0.291)	0.125 (0.331)	0.254 (0.435)	0.141 (0.348)	0.287 (0.452)
Technician	0.099 (0.299)	0.089 (0.285)	0.111 (0.315)	0.120 (0.325)	0.090 (0.286)	0.096 (0.295)
Employee	0.195 (0.397)	0.170 (0.376)	0.191 (0.393)	0.197 (0.398)	0.190 (0.393)	0.192 (0.394)
Blue collar	0.294 (0.456)	0.454 (0.498)	0.363 (0.481)	0.245 (0.430)	0.298 (0.457)	0.189 (0.391)
County variables						
Unemployment rate (in %)	11.2	11.9	11.6	6.004 (0.694)	8.600 (1.764)	8.409 (1.359)
Observations	281	1002	1412	1697	21455	1442

Notes: Data are taken from CEREQ Génération Surveys, wave 1992.

By Decree-Law of January 20, 1993, the University of La Rochelle was officially created in the Charente-Maritime county; the first promotion of students began their two-year university degree (i.e., DEUG) in October 1993. No previous university existed in Charente-Maritime. However, University of Poitiers (in the Vienne county) had established a small satellite campus in La Rochelle in 1974. This campus only provided a DEUG in Law. Figure 1.3 shows that under-education is also prevalent in the CM county.

The data used in Figure 1.3 are extracted from the french population census surveys. Panels (A), (B) and (C) describe the evolution of economic, social and demographic trends for a sequence of cohorts plausibly non affected by the creation, i.e., the 45 years old cohorts. These Panels show that, before the creation period, the conditions of the counties wherein the 7 universities were created were heterogeneous. Between 1986-1991, Charente-Maritime and Nord-Pas-de-Calais had higher unemployment rates, higher shares of blue collars and lower shares of higher education graduates than the UVN counties. Panel (D) describes the average birth cohort size 18 years earlier (i.e., in thousands, the per county average in a given location). It shows that the youth cohort size is higher in Nord-Pas-de-Calais than in other regions and that trends are heterogeneous according to location. Figure 1.3 leads to the conclusion that potential heterogeneous impacts of a university creation are to be expected, due to different initial conditions and potential catch-up mechanisms. Higher impacts on education may be expected in the less educated CM and PC counties while only displacement effects might be expected in the more educated UVN counties (due to the decongestion process of Paris universities, see Subsection 1.2.2).

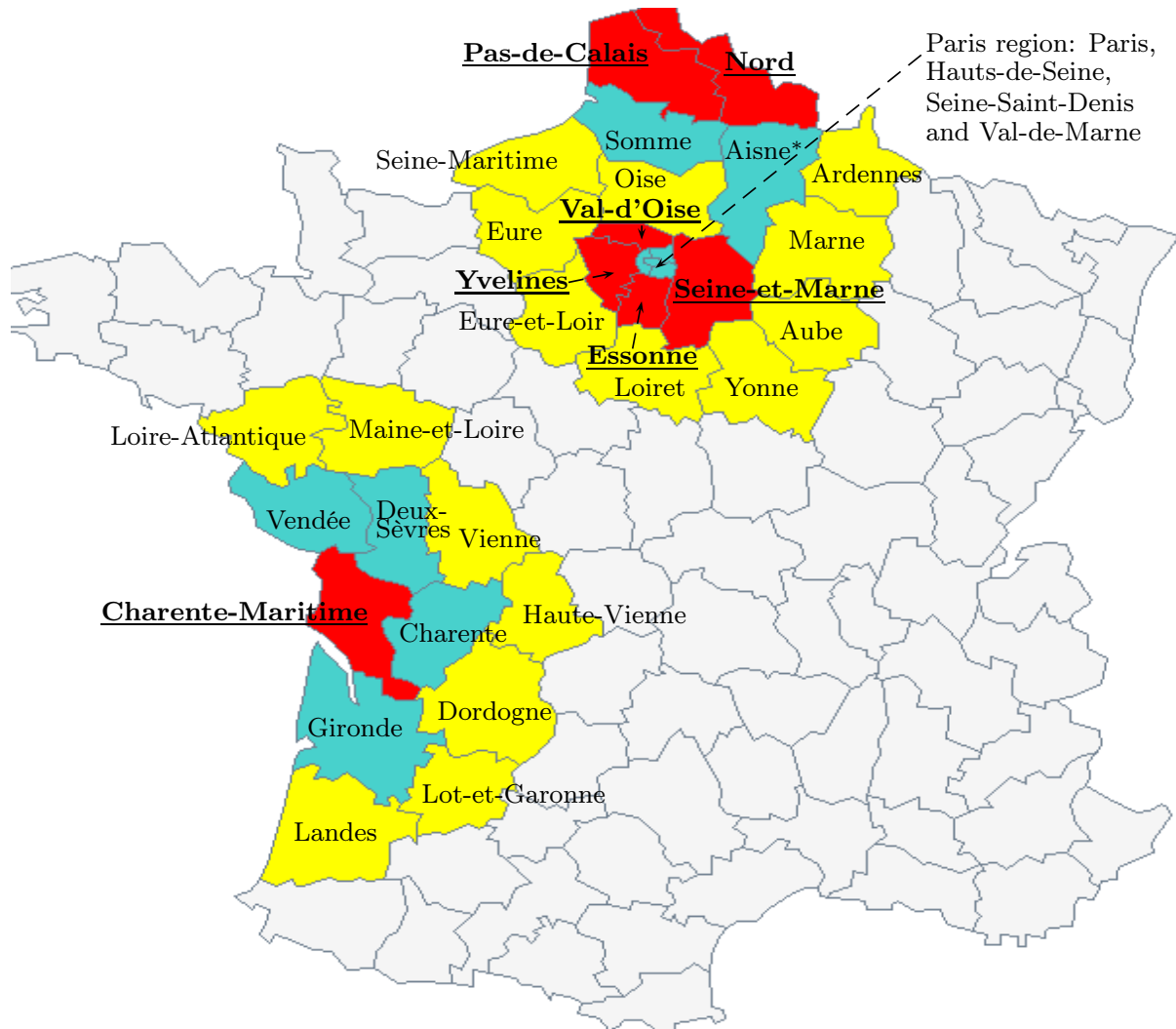
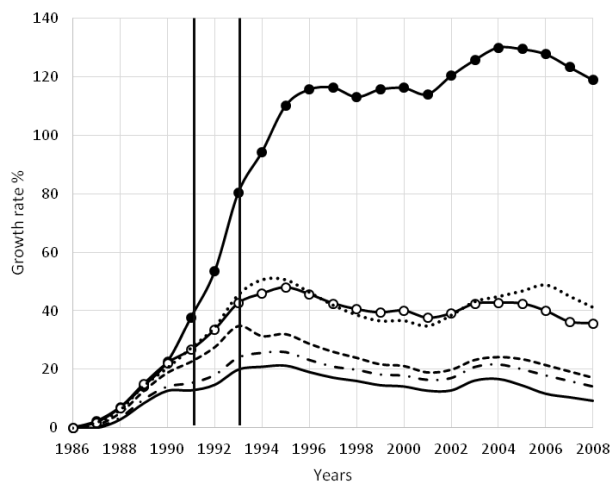
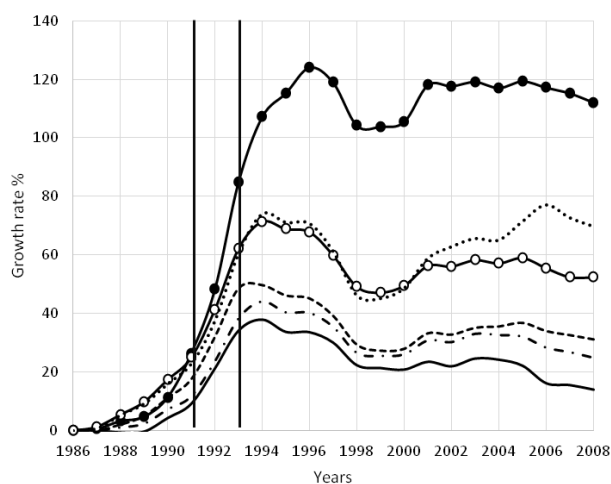


Figure 1.1 – Counties with new university openings (1991-1993)

Note: Counties where a new university opened are in red, and their name is underlined. Bordering counties are in green. They will belong to our first control group in our DiD estimates; see Sections 1.4 and 1.5. Lastly, the “extended neighboring” counties are in yellow. One exception is the Aisne county: it is in green because it is a bordering county of Nord and Pas-de-Calais, but it is also a “extended neighboring” county of the counties of the *villes nouvelles* (i.e., Val d’Oise, Seine-et-Marne, Essonne and Yvelines); see Sections 1.4.2 and 1.4.3 for more details.



(A) in absolute value



(B) relative to birth cohort $t - 18$

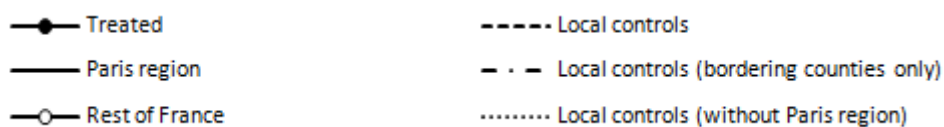
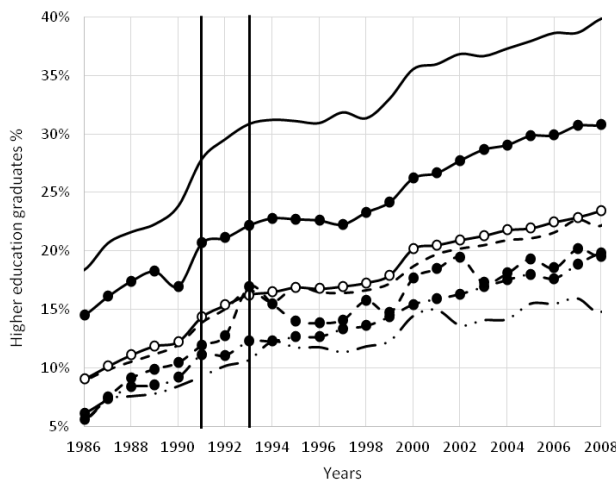
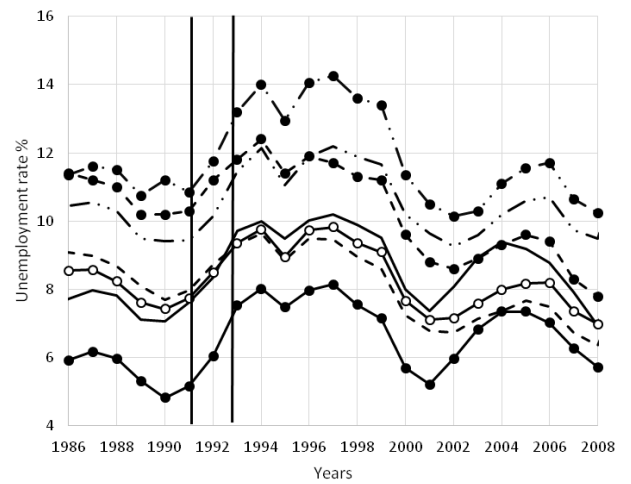


Figure 1.2 – Growth rate of student enrollment at the university (relative to 1986, according to university location)

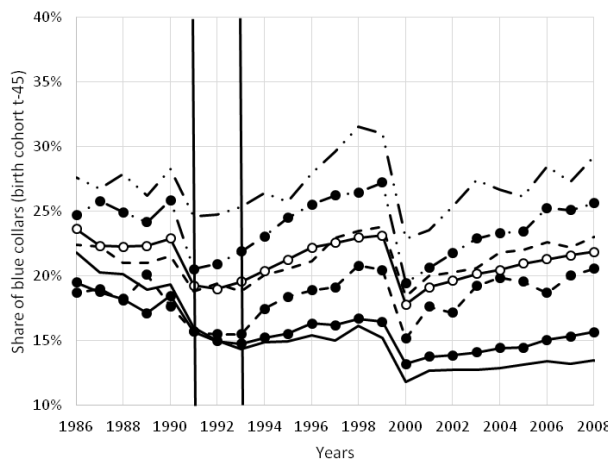
Source: Enrollment in French universities, statistical tables from 1986 to 2008, Evaluation and Prospective Department, Ministry of Education and Research (in French: “Tableaux statistiques du nombre d’inscrits dans les universités françaises, éditions annuelles de 1986 à 2008, Direction de l’Evaluation et de la Prospective, Ministère de l’Education Nationale, Enseignement Supérieur et Recherche”)



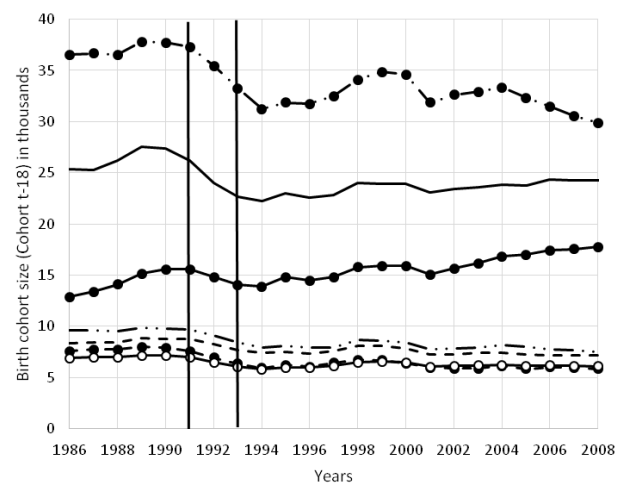
(A) Share of higher education graduates (birth cohort $t - 45$)



(B) Unemployment rate (birth cohort $t - 45$)



(C) Share of blue collars (birth cohort $t - 45$)



(D) Birth cohort size (cohort $t - 18$)

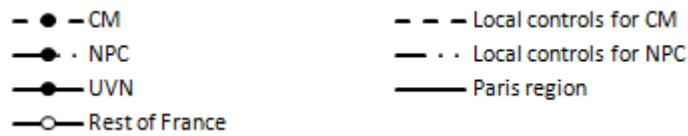


Figure 1.3 – Economic and social trends according to location (relative to population size)

Source: Standardized data from population census surveys, 1968 to 2010, National Institute of Statistics and Economic Studies (in French: “Données harmonisées des recensements de la population de 1968 à 2010, Institut National de la Statistique et des Etudes Economiques”)

1.3 The main data: The “*Génération*” survey

To analyze the effects of new university openings on the acquisition of human capital by the local youth, we use five waves of the *Génération* survey, a large scale survey realized by the CEREQ; each wave is a representative sample (at the national, regional, and county levels) of young individuals who left the educational system at the end of their studies the year of the wave.¹⁵ More precisely, we will exploit the following waves of this survey:

- *Génération 92*, the first wave of this survey, which is a sample of 26359 young individuals who left the educational system between January 1st and December 31st, 1992; the interviews were conducted between 1997-1998.
- *Génération 98* which is a sample of 55345 young individuals who left the educational system between January 1st and December 31st, 1998; the interviews were conducted in 2001.
- *Génération 01* which is a sample of 13987 young individuals who left the educational system between January 1st and December 31st, 2001; the interviews were conducted in 2004.
- *Génération 04* which is a sample of 33655 young individuals who left the educational system between January 1st and December 31st, 2004; the interviews were conducted in 2007.
- *Génération 07* which is a sample of 24879 young individuals who left the educational system between January 1st and December 31st, 2007; the interviews were conducted in 2010.

Thus a total number of 154225 individuals were interviewed in the five cross-sections of the *Génération* survey. We exclude from this sample respondents from the Overseas regions. We will exploit various subsamples of this survey, as we will describe in Section 1.4. In this Section, we describe the information contained in all the waves that we will exploit for our analysis.

All these waves contain information on each individual’s last degree attainment. We also know if each individual obtained the *Baccalauréat* and attended the university directly after. All

15. The CEREQ is a research agency working under the aegis of the Ministry of education; the website of the CEREQ (<http://www.cereq.fr>) describes the basic features of this survey which has already been used by different researchers; see, e.g., [Beffy et al. \(2012\)](#) and [Brodaty et al. \(2014\)](#) who estimate structural models of educational choice under uncertainty to understand how risk aversion affects educational choices, and [Belzil and Penas \(2010\)](#) who study education and early career of French second-generation immigrants.

these waves also provide a key piece of information: the county of residence of the individual at the age of entry in middle school. It is an exogenous variable in the sense that it is a priori not influenced by the university opening. Thus, it is the variable that we will use to assign an individual to the treatment or control group in the different DiD that we will estimate.

In our analysis, we are mainly interested in the average treatment effect (ATE) of a new university opening on the probability of attaining a higher education diploma for the population of the county. As noticed in Subsection 1.2.1, there were some connections between the university system, vocational colleges and *Grandes Ecoles*. For instance, someone with a two-year university degree (i.e., a DEUG) could have entered in a *Grande Ecole*. Such hypothetical individual appears as someone who left the educational system with a degree from a *Grande Ecole* in the *Génération* survey, but the proximity of the university was fundamental to begin post-secondary education; see also Footnote 8. Because of these connections, our two main left-hand side variables will be:

- $Y_{i,c,t}^{2+}$ which is equal to one if individual i who was living in county c when he entered in middle school leaves the educational system in t with at least a two-year degree attainment (i.e., DEUG, BTS, DUT, *Licence*, *Maîtrise*, DESS, DEA, degree from a *Grande Ecole*, or PhD), zero otherwise.
- $Y_{i,c,t}^{4+}$ which is equal to one if individual i who was living in county c when he entered in middle school leaves the educational system in t with at least a four-year degree attainment (i.e., *Maîtrise*, DESS, DEA, degree from a *Grande Ecole*, or PhD), zero otherwise.

We focus on two- and four-year post-secondary degree or higher diploma because these levels of education corresponded to first and second cycle diplomas, until 2007 as described in Subsection 1.2.1. We will consider two additional left-hand side variables to focus exclusively on individuals who attend the university at one moment during their education:

- $Y_{i,c,t}^O$ which is equal to one if individual i who was living in county c when he entered in middle school obtains the *Baccalauréat* and attends the university directly after, zero otherwise.
- $Y_{i,c,t}^U$ which is equal to one if individual i who was living in county c when he entered in middle school leaves the educational system in t attaining a university degree as last diploma (i.e., DEUG, *Licence*, *Maîtrise*, DESS, DEA, or PhD), zero otherwise.

All these waves also provide a set of individual variables: if he or she was 12 or more years old when he or she entered in middle school, the individual’s gender, if the individual’s father had an economic activity or not when he or she left the education system (i.e., if his or her father is employed or has his own business), the last socio-professional category of the individual’s father when the individual left the education system (farmer, storekeeper, executive, technician, blue collar, employee, other), if his or her two parents were born in France, if his or her two parents were born abroad, if only one of his parent was born in France or if he has only one parent.

Before to step any further, three remarks are in order.

Remark (new universities created in 1991 and pre-existing satellites) In the UVN counties, the new universities were officially created at the end of 1991; at the same time the young individuals sampled in the *ex ante* survey *Génération 92* enrolled in their last schooling year before they left the educational system in 1992. Assume that a significant fraction of students that should have left the educational system in 1992 in the absence of university creation, decided to enrol into a new university in 1991 (at the date of creation) and obtained a higher degree later in that university. Then, in this context, the outflow proportion of higher education graduates might be overestimated in the treated counties, both before and after treatment, and the DiD estimator might be biased. Although the creation of other new universities was posterior to 1991¹⁶, as a small number of campuses or satellites were pre-existing in 1991 (see Section 1.2.2). This might also appear to the reader as a potential source of bias for the DiDs¹⁷.

However, the following observations tend to prove that these potential sources of bias are likely to be insignificant. First, in the absence of treatment effect, the bias is mechanically null. This is what we expect in the case of the UVN counties where decongestion and displacement effects are prevailing. Second, even if student enrolment into a new university is not absolutely nil in 1991, it is limited to small year groups in a small number of (pre-existing) campuses or satellites. This is consistent with the growth rates of student enrolment into a university (observed in Figure 1.2)

16. The University of La Rochelle in Charente-Maritime opened in October 1993, ie. two years after that the individuals sampled in *Génération 92* enrolled in 1991. The Artois and Littoral Opal Coast Universities opened in October 1992, so one year after.

17. Our DiDs might also be biased if we would observe student graduations in a new university in the wave 1992 of the survey. However, a minimum of two years after university enrolment is necessary to graduate the first university degree (the DEUG). As a consequence, student graduation cannot be observed before June 1993 in any new university, which is clearly after 1992.

that are not yet significantly different between the treated and control counties in 1991, compared to the following years.

Remark (assignment to treatment and county of residence) The survey *Génération* provides the county of residence at the age of entry in middle school, that we use to assign an individual to the treatment or control groups. Pupils usually enroll into middle school at the age of 11, half way between the year of birth and the potential enrolment into higher education. This is a good compromise between exogeneity (not too close to university enrolment) and explanatory power (not too far from university enrolment). The French Territories Observatory (OT, 2018) reports that, on average every year, the inter-county residential mobility concerns around 3% of the 0-15 years old. This proportion decreases with age: almost 4% for the 1-5 years old, 3% for the 6-10 years old and 2% for the 11-15 years old. A simple calculation shows that the overall probability of inter-county residential mobility might concern up to 45% of the 0-15 years old and still up to 35% for the 0-10 years old alone. As a consequence, using the residential location at birth available in the usual census surveys, rather than the residential location at middle school entry, would induce an unacceptable level of measurement errors.¹⁸ The report also highlights that inter-county mobility rises again to 5% for the 16-20 years old, due to the legal age of compulsory school (16 years old), the legal age of majority (18 years old), and the related decision to pursue in higher education usually taken between 18-20 years old, after high school graduation. From that point of view, residential location at middle school is relevant to guarantee exogenous assignment.

Remark (re-entry into schooling) The young individuals sampled in the *Génération* survey left the educational system for at least one year. However, a reader might highlight that a person does not always leave the educational system for good. For instance, a person may not attend a university immediately after his *Baccalauréat*, but two years after a period of professional experience in a firm. Such hypothetical person appears in the data as someone who leaves the educational system with high school completion, while he should appear as someone with potentially a vocational degree or a university degree. However, since the main interviews of the *Génération* survey are realized three years after the year of the wave, it is possible to know

18. Note also that the location available in the French census surveys is not exactly the county of residence at birth, but the county of birth from the civil-status registries. The French National Institute of Statistics and Economic Studies (Insee, 2017) reports that on average, about 12-13 % of births occur in a county different from the mother county of residence, which would induce an additional source of measurement error.

(except for the wave 92) if a person goes back in the educational system during this period. We observe that only 4.65% of those who left the french educational system after 1998 declared that they went back more than three months in the educational system in the three following years. They are 3.57% to declare more than 9 months of study which is the standard duration for a school year. So this restriction in the data is of no practical consequence.

1.4 Objective, method and descriptive statistics

Our main objective is to measure the ATE δ_c^{j+} of a new university opening in county c on the probability that an individual who was living in this county when he entered in middle school attains at least a j -year post-secondary degree. That is, if $\mathbb{E}[Y_{i,c,t}^{j+}(1)|c, t]$ is the probability that an individual who was living in county c when he entered in middle school attains at least a j -year post-secondary degree in year t if county c were to have experienced the creation of a new university, and $\mathbb{E}[Y_{i,c,t}^{j+}(0)|c, t]$ the probability that such an individual attains at least a j -year post-secondary degree in year t if county c were not to have experienced the creation of a new university, then the ATE in county c is:

$$\delta_c^{j+} = \mathbb{E}[Y_{i,c,t}^{j+}(1)|c, t] - \mathbb{E}[Y_{i,c,t}^{j+}(0)|c, t] \quad (1)$$

It is important to understand that $Y_{i,c,t}^{j+}(1)$ and $Y_{i,c,t}^{j+}(0)$ are potential outcomes. The data $Y_{i,c,t}^{j+}$ only reveal one of these two mutually exclusive outcomes, and any analysis is based on assumption on the counterfactual (i.e., on the outcome which is not observed). Stronger the assumption, higher the potential bias on δ_c^{j+} , so one would like an approach based on credible assumption.

Our strategy is to consider the different outflow waves of the *Génération* survey (92, 98, 01, 04 and 07) and estimate for each ATE δ_c^{j+} of interest different DiD regressions in different subsamples. Each subsample always includes the individuals living in the county of interest when they entered in middle school (the treatment group) and those who were living in counties belonging to a specific control group. Under standard common trend assumptions and other mild conditions discussed in Section 2.2.2, these outflow DiD estimators identify ATEs in the underlying inflows of new entrants in middle school that we use for treatment assignment.

Subsection 1.4.1 presents the DiD regression for a given control group. To keep our task

manageable and to fully understand, we take as an example the University of La Rochelle created in 1993 in the Charente-Maritime county, and assume that the individuals who were living in the bordering counties in middle school are an adequate control group; that is, the DiD regression permits to obtain an unbiased measure of the ATE. Subsection 1.4.2 explains that, although bordering counties form a natural basis for comparison, this control group may lead to imprecise DiD estimates. We then describe alternative control groups which permit to solve these biases. Finally, Subsection 1.4.3 exposes the ATEs that will be estimated for the other areas which experienced new university openings.

1.4.1 ATEs and DiDs for Charente-Maritime

Let us focus on the University of La Rochelle created in 1993 in the Charente-Maritime county ($c = CM$), and consider that we want to measure the ATE δ_{CM}^{j+} in Equation 1 of this new university opening. We consider for the moment that the sample is composed of the individuals who were living in the county of Charente-Maritime when they entered in middle school (the treatment group), as well as those who were living in the bordering counties, i.e., Charente, Vendée, Gironde and Deux-Sèvres (the control group).¹⁹ We use this sample to estimate the following model, and obtain the DiD β_{CM}^{j+} as a measure of the treatment effect δ_{CM}^{j+} :

$$Y_{i,c,t}^{j+} = \alpha^{j+} + \varphi^{j+} \times CM_c + \lambda^{j+} \times d_t + \beta_{CM}^{j+} \times (CM_c \times d_t) + X_{c,t} \times \zeta^{j+} + Z_{i,c,t} \times \theta^{j+} + \phi^{j+} \times D_{2004} + \vartheta^{j+} \times D_{2007} + \epsilon_{i,c,t}^{j+} \quad (2)$$

where CM_c is a dummy for individuals in Charente-Maritime, d_t is a time dummy that switches on for individuals who left the education system in 1998, 2001, 2004 or 2007 (i.e., after the university creation of 1993), $X_{c,t}$ includes county-specific time-varying variables, and $Z_{i,c,t}$ is a vector of individual attributes. $\epsilon_{i,c,t}^{j+}$ is a random error term. D_{2004} and D_{2007} are year-dummies for 2004 and 2007. We include them because if the organization in cycles with a first cycle corresponding to a two-year post-secondary degree and a second cycle corresponding to a four-year degree was still prevalent until 2007, the progressive suppression of the DEUG and

19. Of the 154225 individuals interviewed in the five cross-sections of the *Génération* survey, 9541 were living in Charente-Maritime or one of its bordering counties when they entered in middle school. Missing values leave us with a sample of 9527 individuals: 1723 in Charente-Maritime (281 in wave 1992, 481 wave 1998, 98 in wave 2001, 277 in wave 2004 and 586 in wave 2007), and 7804 in its neighboring counties (1085 in wave 1992, 2561 in wave 1998, 616 in wave 2001, 1569 in wave 2004 and 1973 in wave 2007).

the *Maîtrise* by the universities between 2004 and 2007 may generate some noise; see Section 1.2.1.²⁰

As it is well known, if we do not include the county-specific time-varying variables $X_{c,t}$ in Equation 2, the DiD β_{CM}^{j+} is an unbiased measure of the ATE of interest δ_{CM}^{j+} (i.e., $\beta_{CM}^{j+} = \delta_{CM}^{j+}$) under the assumption that in the absence of any university creation, Charente-Maritime and its bordering counties would have experienced the same change in the probabilities of attaining at least a j -year post-secondary degree. As shown in Figures 1.2 and 1.3, this untestable assumption is plausible. First Figure 1.2 shows that until 1990, i.e., before the period of university creation, the evolution of enrollment into French universities is very similar between the treated and the control counties. This has to be compared with the gap prevailing after the period of creation between treated and control counties in their respective university enrollments. Second, Figure 1.3 shows that some important county-specific time-varying variables, like unemployment rate or the share of blue collars, that would have affected differently the change in probabilities, follow similar patterns in the treated and bordering counties. Concerning the newborns in $t - 18$, Charente-Maritime and its bordering counties follow similar patterns (but it is not the case for other counties which experienced a creation like the Nord-Pas-de-Calais). To avoid any doubts, we include a quadratic profile for the newborns in $t - 18$ (i.e., newborns in $t - 18$ and (newborns in $t - 18$)²) and for the unemployment rate of the county labor force (not only the 45 years old cohort). We also include the percentage of farmers. These county-specific time-varying variables may permit to relax the common trend assumption.²¹

1.4.2 Solving for multicollinearity, spillover and substitution effects using alternative groups of control for Charente-Maritime

Although the DiD regression of Equation 2 permits to relax (partially) the common trend assumption, various problems might occur if we consider the bordering counties as a control group. To solve for these problems, we consider alternative control groups.

Multicollinearity. First, the five waves of the *Génération* survey (92, 98, 01, 04 and 07) permit to control for county-specific time-varying variables, but the sample size is limited in

20. Note that the exclusion of these year dummies does not change the results.

21. Note that contrary to county-specific time-varying variables, the individual variables $Z_{i,c,t}$ are not a potential source of bias on β_{CM}^{j+} . But they may generate a more precise estimate as noticed by Angrist and Pischke (2009, p.237). The individual variables included in the regressions are those described in Section 2.2.2.

terms of counties: it includes only Charente-Maritime and its four bordering counties. Thus, it may lead to imprecise estimates for β_{CM}^{j+} due to imperfect but high multicollinearity between $(CM_c \times d_t)$ and some of the county-specific time-varying variables $X_{c,t}$ in Equation 2. To solve for this potential problem, we increase the sample size by considering a sample which includes: Charente-Maritime, its 4 bordering counties and the 7 counties around the bordering counties (Loire-Atlantique, Maine-et-Loire, Vienne, Haute-Vienne, Dordogne, Lot-et-Garonne and Landes). These seven additional counties are in yellow around Charente, Vendée, Gironde and Deux-Sèvres in Figure 1.1. This extended sample composed of 12 counties is called the sample with local controls. The term “local” is used because even if these additional counties do not border Charente-Maritime, they remain very close.

Spillover and substitution effects. Although the sample with local controls may solve for problems of multicollinearity, the DiD β_{CM}^{j+} in Equation 2 may still be a biased and underestimated measure of the ATE of interest δ_{CM}^{j+} . First, the bordering counties, or, more generally, the local controls, could be affected by Charente-Maritime due to spillover effects, i.e., the fact that the creation of the University of La Rochelle may have a positive impact on students coming from these control groups. The waves 98, 01, 04 and 07 of the *Génération* survey provides the last institution wherein each individual was studying before he left the education system. Using this information, Table 1.2 reports that 1.86 percent of the individuals from the local controls of Charente-Maritime who left the university between 1998 and 2007 were studying in the University of La Rochelle, and they represented 29.84 percent of the students enrolled in that university. The first percentage (1.86) is only a *potential* spillover effect, given that it may capture in reality a displacement effect. Indeed, we do not know if these 1.86 percent of individuals from the local controls of Charente-Maritime really benefited from the University of La Rochelle; perhaps they would have attained a similar level of education in another institution of the higher education system without the creation of this university. The second percentage suggests however that the University of La Rochelle is an important institution for these individuals to attain a university degree because they represent 29.84 percent of its students. If these percentages capture spillover effects, our DiD β_{CM}^{j+} with the local controls as a control group will underestimate the ATE δ_{CM}^{j+} ceteris paribus.²² Consequently, we will consider a sample with Charente-Maritime

22. On the contrary, if these percentages capture displacement effects, our DiD β_{CM}^{j+} with the local controls as a control group will be an unbiased measure of the ATE δ_{CM}^{j+} ceteris paribus. This is due to the fact that our probabilities of interest cancel out any potential displacement effect because they are probabilities of attaining a

and the Paris region (i.e., Paris, Hauts-de-Seine, Seine-Saint-Denis and Val-de-Marne) as a control group. As shown in Table 1.2, only 0.07 percent of the individuals from the Paris region who left the university between 1998 and 2007 were studying in the University of La Rochelle, and they represented only 1.24 percent of its students. Thus, its youth is clearly not impacted by the creation of the University of La Rochelle. Furthermore, Figure 1.2 suggests that the Paris region is the area which experienced the lowest shock in the supply of higher education. If so, it is the best control group to avoid a substitution bias, as mentioned in the Introduction.

Note that Figures 1.2 and 1.3 also suggest that the Paris region is a control group a priori as good as the bordering counties or the local controls. Indeed, Figure 1.2 shows that before 1990, the evolution of enrollment into a university in the Paris region is very similar to the growth rate in the treated counties. Thus, the plausibility of the common trend assumption is not rejected neither for the local controls, nor for the Paris region. Concerning Figure 1.3, it shows that the Paris region differs in level from the Charente-Maritime county in terms of county-specific time-varying variables. However, it follows a similar pattern that is compatible with the common trend assumption (see, e.g., the share of higher education graduates in Figure 1.3).

Finally, we will investigate as a last control group for Charente-Maritime all the 89 non-treated counties in France. This sample is called the sample with the rest of France as a control group. Like the sample with local controls, it prevents from problems of multicollinearity between $(CM_c \times d_t)$ and some of the county-specific time-varying variables $X_{c,t}$ in Equation 2. This control group does not fully prevent from a spillover bias, given that it includes the local controls. However the spillover bias will be reduced de facto because all the other individuals from this control group are too far from Charente-Maritime to benefit from the University of La Rochelle. Concerning a possible substitution bias, the rest of France, like the local controls, does not prevent from it: it experienced the second highest increase in the supply of higher education according to Figure 1.2.

1.4.3 ATEs and DiDs for the other treated counties

We expose here the other ATEs that will be estimated for the other areas which experienced new university openings, and their sensitivity to the problems mentioned in Subsection 1.4.2.

Note that it is not possible to estimate a treatment effect for each university because some j -year post-secondary degree in any institution of the higher education system.

Table 1.2 – *Assessing potential spillover and displacement effects*

	Charente-Maritime $c = CM$	Nord-Pas-de-Calais or Pas-de-Calais $c = NPC$ or $c = PC$	Villes Nouvelles $c = VN$
Treated	26.77 [38.05]	23.93 [87.32]	20.01 [56.18]
Local controls (bordering counties only)	3.59 [23.04] ^(i.)	2.11 [1.76]	6.08 [16.89]
Local controls	1.86 [29.84]	1.38 [4.00]	5.56 [25.67]
Paris region	0.07 [1.24]	0.23 [1.00]	6.08 [16.89]
Rest of France	0.45 [59.42]	0.31 [10.84]	1.84 [41.77]
Paris county	0.07 [0.40]	0.28 [0.39]	3.93 [3.60]

Notes: (i.) “3.59 [23.04]” means that 3.59 percent of the individuals from the bordering counties of Charente-Maritime who left the university between 1998 and 2007 were studying in the University of La Rochelle; they represent 23.04 percent of all enrolled students into that university.

of them opened in the same county or counties, or too close area. As noticed in Subsection 1.2.2, Artois University and University of the Littoral Opal Coast are in the Pas-de-Calais county ($c = PC$), but the University of the Littoral Opal Coast has also one site in Dunkirk in the Nord county ($c = N$). For these two counties, we will first consider an homogenous effect; we denote the ATE of these two new universities openings in the Nord-Pas-de-Calais region δ_{NPC}^{j+} . Recall that before the creations of these new universities, there was no university in the Pas-de-Calais county, but various in the Nord county, in particular in Lille; see Subsection 1.2.2. To be certain that our effects are not due to existing universities in the Nord county, we will also consider the ATE δ_{PC}^{j+} of these new universities considering only the Pas-de-Calais county. In both cases, the DiD regressions with only bordering counties as a control group will include the counties of Aisne and Somme as shown in Figure 1.1. The corresponding samples include only four counties to estimate the DiD β_{NPC}^{j+} and three for the DiD β_{PC}^{j+} ; it may lead to imprecise estimates due to multicollinearity. Thus, like for Charente-Maritime, we consider samples with local controls, i.e., the bordering counties as well as those which are around (Seine-Maritime, Oise, Marne and Ardennes). These four additional counties are those in yellow around Aisne and Somme in Figure 1.1. We will also consider regressions with the Paris region as a control group to avoid spillover and substitution biases. However, according to Table 1.2, 1.38 percent of the individuals from the local controls who left the university between 1998 and 2007 were studying in the two new universities of the (Nord)-Pas-de-Calais, and they represented only 4 percent of the students

enrolled in these universities. Again, we do not know if these percentages reflect spillover or displacement effects. But even if they reflect spillover effects, they are much lower than in the case of Charente-Maritime. Thus, the main advantage here of the Paris region as a control group is to reduce possible substitution bias. Finally, the rest of France as a control group will be also considered.

Concerning the four universities of the *villes nouvelles* which were created in the same area around the Paris region, we will consider an homogenous effect. The ATE of these new universities in these counties of the *villes nouvelles* is denoted δ_{VN}^{j+} . The bordering counties are those of the Paris region. Thus, the DiD regressions with bordering counties as a control group and the DiD regressions with the Paris region as a control group will be the same. There is a priori less problem of multicollinearity than with the DiD regressions for Pas-de-Calais or for Charente-Maritime because there are 8 counties (against 3 and 5, respectively): the 4 treated counties (Val-d'Oise, Seine-et-Marne, Essonne and Yvelines) and the 4 counties belonging to the control group (i.e., Paris, Hauts-de-Seine, Seine-Saint-Denis and Val-de-Marne). As shown in Table 1.2, 6.86 percent of the individuals from the bordering counties who left the university between 1998 and 2007 were studying in the universities of the *villes nouvelles*. This percentage may reflect only displacement effects: there are so many institutions of the higher education system in the Paris region that these individuals will have found easily another institution to attain a similar level of education without the creations of the universities of the *villes nouvelles*. However, if this percentage reflects a spillover bias, then the Paris region may not be the best control group for the counties of the *villes nouvelles*. To avoid any doubt, we will consider only the Paris county as a control group. According to Table 1.2, 3.93 percent of the individuals from this county were studying in the universities of the *villes nouvelles*, and they represented only 3.60 percent of the students enrolled in these universities. Like for Charente-Maritime and (Nord)-Pas-de-Calais, we will also consider the rest of France as a control group, as well as sample with local controls, i.e., the bordering counties and those which are around (Eure, Eure-et-Loir, Loiret, Yonne, Aube, Marne, Oise and Aisne). These eight additional counties are those in yellow around the counties of the *villes nouvelles* as well as the Aisne county in Figure 1.1.²³

23. The Aisne county is in green in Figure 1.1 because it is also a bordering county of the (Nord)-Pas-de-Calais.

1.5 Main results

Table 1.3 provides the main results, i.e., the DiD estimates on the probability of i) attaining at least a two-year post-secondary degree ($\widehat{\beta}_c^{2+}$) and ii) attaining at least a four-year post-secondary degree ($\widehat{\beta}_c^{4+}$). This Table also provides the DiD estimates on the probability of iii) attending the university directly after the *Baccalauréat* ($\widehat{\beta}_c^O$), and iv) obtaining a university degree as last diploma ($\widehat{\beta}_c^U$), two mechanisms that may explain the impacts of interest. This table only provides the DiD estimate obtained in each regression. Note (iv.) at the bottom of Table 1.3 provides the set of county-specific time-varying variables and individual variables which are included in each regression.

Note that this table provides cluster-robust standard errors. The clustering is on county. So, for instance, for Charente-Maritime, when we consider bordering counties as a control group, there are five clusters ($C = 5$): Charente-Maritime and its four bordering counties.²⁴ For the t -tests based on our cluster-robust standard errors, we follow Donald and Lang (2007) and Cameron and Miller (2015, p.346) by considering $C - 2$ degrees of freedom, with C the number of clusters in the regression, and 2 the number of regressors that are invariant within cluster; Note (ii.) at the bottom of Table 1.3 provides additional details. This is more conservative than the $C - 1$ degrees of freedom provided by computer packages, like, e.g., Stata.

Finally, this Table also reports a variance inflation factor (VIF) for the different DiD estimates as a diagnostic statistic of imprecise estimates due to multicollinearity. Values in excess of 10 are suggested as indicative of the problem.²⁵

Comparing the impact of new university openings with regard to their location, Table 1.3 shows that almost no significant impact is observed in the counties of the *villes nouvelles*. The only exception concerns the probability of obtaining a university degree as last diploma: the impact of the universities of the *villes nouvelles* is always significant and varies between 3 and 5.1 percentage points depending on the control group considered. However, recall that if the Paris region is usually a good control group to avoid spillover effects, Table 1.2 suggests that it

24. We do not consider a cluster on county-year pairs because, for example, the error for Charente-Maritime in 1992 is likely to be correlated with the error for Charente-Maritime in 1998.

25. Note that the VIF of a regressor is a measure of collinearity with the other regressors of a regression. Different regressions with different left-hand side variables but identical sample and regressors will provide identical VIFs. For instance, for Charente-Maritime, when the control group is the bordering counties, the four DiD estimates $\widehat{\beta}_{CM}^{2+}$, $\widehat{\beta}_{CM}^{4+}$, $\widehat{\beta}_{CM}^O$ and $\widehat{\beta}_{CM}^U$ must have the same VIF, given that the sample and the regressors are identical. That's why Table 1.3 reports one VIF by sample.

may not be the case for the counties of the *villes nouvelles*. Thus, and as explained in Subsection 1.4.3, we have considered only the Paris county as a control group in Table 1.4.²⁶ The DiD estimates are very similar to those with the Paris region as a control group, and do not differ significantly from zero, except, again, for the probability of obtaining a university degree as last diploma.

In contrast, and as we will explain in detail below, we do find for Charente-Maritime and (Nord)-Pas-de-Calais positive and significant impacts of the new university openings, varying around 10 percentage points for the probability of attaining at least a two-year post-secondary degree, and around 5 percentage points for the probability of attaining at least a four-year post-secondary degree. These strong differences with the *villes nouvelles* are not surprising: as already noticed in Subsection 1.2.2, the different public administrations involved in the creation and the evaluation of these universities highlighted that the four universities of the *villes nouvelles* were created to decongest Paris universities, but not to favor investment in higher education; the population was educated like in the rest of France as shown in Table 1.1. Conversely, the population was undereducated in the counties of Charente-Maritime and Pas-de-Calais. Hence, it suggests that this higher impact in lower educated counties reflects an inter-counties catch-up effect.

Comparing the results in Table 1.3 according to the choice of the control group is also informative about the biases at stake. Results for the Nord-Pas-de-Calais are markedly stronger when we consider only bordering counties as a control group: the probability of attaining at least a two-year post-secondary degree increases by 22 percentage points. However, recall that the sites of the two new universities in the Nord-Pas-de-Calais are mainly in the Pas-de-Calais county where no previous universities existed, while various universities were already present in the Nord county, as noticed in Subsection 1.4.2. When we exclude the Nord county from the

26. In Table 1.4, the squared terms of the county-specific time-varying variables are excluded (i.e., (unemployment rate)² and (newborns in $t - 18$)²) to avoid too strong problems of multicollinearity. Indeed, considering the Paris county and the *villes nouvelles*, the two “newborns $t - 18$ ” terms are highly correlated, the coefficient of correlation being 0.9949, like the two “unemployment rate” terms (0.9953). Thus, in the regressions of our different left-hand side variables, including unemployment rate, (unemployment rate)², newborns in $t - 18$, (newborns in $t - 18$)² as control variables generate a VIF of 2609 for (newborns in $t - 18$)², 1449 for (newborns in $t - 18$), above 120 for the two unemployment rate terms, and 18.40 for the DiD estimates. The high correlations between the two “newborns $t - 18$ ” terms and the two “unemployment rate” terms highlight that the squared terms are useless; hence we exclude them. It permits to reduce the VIF for the DiD estimate to 13.77. To have a point of comparison with the Paris region as a control group, Table 1.4 also provides the DiD estimates with this control group excluding the squared terms of the county-specific time-varying variables.

regression and consider only Pas-de-Calais as a treated county, the probability of attaining at least a two-year post-secondary degree increases by 12.3 percentage points. The VIF (28.75) of this DiD estimate suggests imprecision, but the results with the other control groups, although slightly lower, are broadly similar and always significant ranging between 8 and 11.1 percentage points.²⁷ Note also that the DiD estimates with the local controls and with the Paris region as a control group are very similar (10.2 and 11.1 percentage points). The Paris region is the best control group to avoid spillover and substitution biases. The lack of difference between the DiD estimates confirms what we were expecting in Subsection 1.4.3: the DiD estimate $\widehat{\beta}_{PC}^{2+}$ with only the local controls for Pas-de-Calais does not strongly underestimate the ATE δ_{PC}^{2+} because there is no or few spillover effects. The lack of difference between the DiD estimates also suggests that the substitution bias is small when we consider the local controls, given that the results are similar to those with the Paris region as a control group.

Concerning Charente-Maritime, there is a priori no strong problem of multicollinearity when we consider only the bordering counties as a control group. The DiD estimate on the probability of attaining at least a two-year post-secondary degree is 3.1 percentage points and not significant. It reaches 6.4 percentage points and becomes significant when we consider the local controls as a control group. Nevertheless, recall that these DiD estimates for Charente-Maritime may be plagued by a spillover bias. When we consider the Paris region as a control group, the DiD estimate on the probability of attaining at least a two-year post-secondary degree becomes 13.6 percentage points and is statistically significant. The result is similar when we consider the rest of France as a control group (13 percentage points).

When we consider the mechanisms behind these results, i.e., the probability of attending the university directly after the *Baccalauréat* or the probability of obtaining a university degree as last diploma, we also do find significant results. More precisely, the new universities have a significant impact on the probability of attending the university directly after the *Baccalauréat* for Charente-Maritime (respectively 6.4 and 7.4 percentage points with the Paris region and the rest of France as a control group), as well as for Pas-de-Calais (2.5 and 3.3 percentage points with the Paris region and the rest of France as a control group). These results can be

27. This problem of imprecise estimates is particularly striking when we consider the probability of attaining a four-year post-secondary degree. We find the incredible result that the impact of the new university openings in Pas-de-Calais is 14.2 percentage points when the control group is composed of its two bordering counties. With the other control groups, the impact is always significant ranging between 3.20 and 5 percentage points.

explained by the fact that the university is not selective contrary to other post-secondary tracks after the *Baccalauréat*. Hence, the university is an opportunity for individuals who obtain the *Baccalauréat* to attain for sure a higher education level. Concerning the probability of obtaining a university degree as last diploma, the results are particularly significant in Charente-Maritime; the estimates are 8.6 and 10 percentage points if we consider the Paris region and the rest of France as a control group, respectively. This result can be explained again by the fact that the university is not selective after the *Baccalauréat*, and it ensures some individuals to reach a higher education level. Furthermore, other individuals, after having enrolled and graduated in other initial tracks, may also have chosen to enroll in an advanced university degree to complete their education.

Table 1.3 – DiD estimates of new university openings (concluded on next page)

	Charente-Maritime (<i>c = CM</i>)	Nord-Pas-de-Calais (<i>c = NPC</i>)	Pas-de-Calais (<i>c = PC</i>)	Villes Nouvelles (<i>c = VN</i>)
[A] Local controls (bordering counties only)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.031 (0.015)	0.222** (0.027)	0.123** (0.004)	0.049 (0.027)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	-0.007 (0.011)	0.021 (0.008)	0.141* (0.022)	0.030* (0.015)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	-0.016 (0.031)	0.192** (0.039)	0.298** (0.007)	0.015 (0.016)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	-0.009 (0.007)	0.162** (0.037)	0.225** (0.005)	0.051* (0.021)
Observations	9527	15408	8113	18708
Degrees of freedom	3	2	1	6
VIF	7.08	13.50	28.75	8.77
[B] Local controls				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.064*** (0.019)	0.098* (0.045)	0.102** (0.031)	0.047 (0.029)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.013 (0.013)	0.028 (0.019)	0.042** (0.014)	0.015 (0.013)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.017 (0.019)	0.067 (0.045)	0.064 (0.043)	0.019 (0.015)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.029 (0.018)	0.069 (0.043)	0.094** (0.035)	0.042** (0.017)
Observations	20599	23765	16470	29582
Degrees of freedom	10	6	5	14
VIF	6.30	8.47	8.45	7.72
[C] Paris region (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.136*** (0.009)	0.121*** (0.026)	0.111*** (0.014)	0.049 (0.027)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.055** (0.011)	0.054** (0.017)	0.050** (0.010)	0.030* (0.015)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.064*** (0.004)	0.034* (0.013)	0.025* (0.009)	0.015 (0.016)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.086*** (0.010)	0.042* (0.018)	0.040 (0.018)	0.051* (0.021)
Observations	10232	20577	13282	18708
Degrees of freedom	3	4	3	6
VIF	6.06	10.51	7.49	8.77

Table 1.3 – *DiD estimates of new university openings* (continued from previous page)

	Charente-Maritime ($c = CM$)	Nord-Pas-de-Calais ($c = NPC$)	Pas-de-Calais ($c = PC$)	Villes Nouvelles ($c = VN$)
[D] Rest of France (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.130*** (0.009)	0.054*** (0.015)	0.080*** (0.010)	0.015 (0.026)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.043*** (0.004)	0.037*** (0.012)	0.032*** (0.005)	0.008 (0.011)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.074*** (0.008)	0.026*** (0.008)	0.033*** (0.008)	0.005 (0.012)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.100*** (0.009)	0.037*** (0.011)	0.051*** (0.009)	0.029** (0.013)
Observations	128244	138589	131294	136720
Degrees of freedom	88	89	88	91
VIF	5.50	5.96	5.54	6.81

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, we consider $C - 2$ degrees of freedom, with C the number of clusters in the regression of interest and 2 the number of regressors which are invariant within cluster. For instance, in Part [A] of the Table, the regression used to estimate the DiD β_{CM}^{2+} in Charente-Maritime (see Equation 2), there are five clusters ($C = 5$): Charente-Maritime -the county wherein a university creation took place- and the four neighboring counties -the control group- which include Charente, Vendée, Gironde and Deux-Sèvres. Equation 2 shows that there are two time-invariant regressors within cluster: the intercept and the dummy CM_c , so there are 3 degrees of freedom. The critical value from the t distribution with 3 degrees of freedom for a two-sided test at 10 percent level is 2.353, while the t ratio for $\widehat{\beta}_{CM}^{2+}$ is $0.031/0.015 \approx 2.08$. Given that $2.08 < 2.353$, the estimated coefficient is not statistically significant at the 10 percent level.

(iii.) Each regression corresponds to a regression similar to the one presented in Equation 2. Each regression includes the set of county-specific time-varying variables, as well as individual variables, and two year-dummies (for the years 2004 and 2007).

The set of county-specific time-varying variables are: unemployment rate, (unemployment rate)², newborns in $t - 18$, (newborns in $t - 18$)², and the percentage of farmers.

The set of individual variables are: a dummy variable equal to one if he or she was 12 or more years old when he or 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 he or 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 his/her parents –two parents born in France, two parents born abroad (the category of reference is only one of his parent was born in France or if he has only one parent)–.

Table 1.4 – *DiD estimates of the Universities of the villes nouvelles: Paris region versus Paris county as a control group*

Villes Nouvelles ($c = VN$)		
	Paris region	Paris county
DiD estimates on the probability of attaining at least a two-year post-secondary degree		
β_c^{2+}	0.050* (0.026)	0.052 (0.042)
DiD estimates on the probability of attaining at least a four-year post-secondary degree		
β_c^{4+}	0.026 (0.014)	0.036 (0.027)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after		
β_c^0	0.008 (0.010)	0.025 (0.024)
DiD estimates on the probability of attaining a university degree as last diploma		
β_c^U	0.031* (0.015)	0.104** (0.026)
Observations	18708	12512
Degrees of freedom	6	3
VIF	7.95	13.77

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).

(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. We consider $C - 2$ degrees of freedom, where C is the number of clusters in each regression; see Note (ii.) in Table 1.3.

(iii.) Each regression includes a set of county-specific time-varying variables, as well as individual variables, and two year-dummies (for the years 2004 and 2007).

The set of county-specific time-varying variables are: unemployment rate, newborns in $t - 18$, and the percentage of farmers. Contrary to the regressions in Table 1.3, the squared terms of the county-specific time-varying variables are excluded (i.e., $(\text{unemployment rate})^2$ and $(\text{newborns in } t - 18)^2$) to avoid too strong problems of multicollinearity; see Footnote 26.

The set of individual variables is similar to the one of Table 1.3; see Note (iii.) in Table 1.3.

1.6 Robustness checks and falsification tests

Subsection 1.6.1 provides various robustness checks to see if the results of Section 1.5 are robust. Subsection 1.6.2 describes various falsification tests which have been conducted.

1.6.1 Robustness checks

Using the Paris region as a control group is the best way to avoid a (control group) substitution bias because it is the area which experienced the lowest shock in the supply of higher education as shown in Figure 1.2. However, one can argue that if this shock is lower than in other control groups, it is not nil. We thus have re-estimated the different regressions of Table 1.3, adding an additional county-specific time-varying variable: the number of IUT, i.e., the number of institutions which provide the two-year vocational degree called DUT; see Subsection 1.2.1.²⁸

28. The data on the number of IUT are collected from the 2006 Central Data Base of French Educational Institutions realized by the Evaluation and Prospective Department of the Ministry of Education (in French: *Base*

This variable permits to capture the potential increase in the education supply in the control group, at least the increase in two-year vocational institutes. If there is a substitution bias, adding the supply of IUT as a control variable should increase our DiD estimates. Table 1.5 provides the results. We do not include the results with the bordering counties as control groups because the result for the Pas-de-Calais was plagued by multicollinearity (the VIF for the DiD estimates was 54.14), and as we already know, the results for Charente-Maritime is widely reduced due to a spillover bias. Including this IUT variable does not change the main results, except for Charente-Maritime and Pas-de-Calais when the Paris region is the control group: the DiD estimates on the probability of attaining at least a two-year post-secondary degree are 17.1 and 18.9 percentage points, and slightly higher than the results in Table 1.3 (13.1 and 11.1, respectively); the DiD estimates are also slightly higher for the probability of attaining at least a four-year post-secondary degree.

A second robustness check consists in the re-estimation of the models without the individuals who left the educational system in 1998. One may argue that this wave is not distant enough from the period of the creations. If so, our DiD effects in Table 1.3 may be underestimated. However, Table 1.6 shows that the DiD estimates are very similar if we exclude the wave 98. It confirms what we observe in Figure 1.2: most of the impacts of the new universities happened a few years only after their creation.

Table 1.5 – DiD estimates of new university openings, controlling for substitution bias

	Charente-Maritime ($c = CM$)	Nord-Pas-de-Calais ($c = NPC$)	Pas-de-Calais ($c = PC$)	Villes Nouvelles ($c = VN$)
[A] Local controls				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.058* (0.029)	0.097* (0.047)	0.115** (0.044)	0.050 (0.030)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.011 (0.017)	0.036* (0.017)	0.031 (0.021)	0.015 (0.013)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.016 (0.020)	0.068 (0.047)	0.069 (0.047)	0.017 (0.014)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.027 (0.021)	0.073 (0.041)	0.097** (0.037)	0.037** (0.015)
Observations	20599	23765	16470	29582
Degrees of freedom	10	6	5	14
VIF	6.31	9.60	11.56	7.78
[B] Paris region (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.171*** (0.010)	0.137*** (0.028)	0.189*** (0.017)	0.056 (0.032)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.073** (0.012)	0.061** (0.017)	0.098** (0.027)	0.031* (0.014)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.055*** (0.006)	0.044** (0.013)	0.022 (0.021)	0.015 (0.017)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.068* (0.022)	0.049 (0.024)	-0.012 (0.026)	0.047* (0.019)
Observations	10232	20577	13282	18708
Degrees of freedom	3	4	3	6
VIF	6.60	10.71	13.86	8.87
[C] Rest of France (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.130*** (0.009)	0.059*** (0.021)	0.093*** (0.012)	0.015 (0.027)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.043*** (0.004)	0.042*** (0.007)	0.040*** (0.007)	0.008 (0.011)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.074*** (0.008)	0.028*** (0.009)	0.037*** (0.011)	0.005 (0.012)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.100*** (0.009)	0.041*** (0.014)	0.058*** (0.011)	0.029** (0.013)
Observations	128244	138589	131294	136720
Degrees of freedom	88	89	88	91
VIF	5.50	5.98	5.65	6.81

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).

(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. We consider $C - 2$ degrees of freedom, where C is the number of clusters in each regression; see Note (ii.) in Table 1.3.

(iii.) Each regression includes the set of county-specific time-varying variables, as well as individual variables, and two year-dummies (for the years 2004 and 2007); see Note (iii.) in Table 1.3. It also includes an additional county-specific time-varying variable: the number of DUT.

Table 1.6 – DiD estimates of new university openings, excluding the wave Génération 98
(concluded on next page)

	Charente-Maritime ($c = CM$)	Nord-Pas-de-Calais ($c = NPC$)	Pas-de-Calais ($c = PC$)	Villes Nouvelles ($c = VN$)
[A] Local controls (bordering counties only)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	-0.040 (0.031)	0.139*** (0.012)	0.081 (0.049)	0.059* (0.025)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	-0.002 (0.008)	0.080** (0.012)	0.186** (0.003)	0.030* (0.014)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.019 (0.044)	0.209** (0.027)	0.176 (0.060)	0.008 (0.016)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	-0.142** (0.031)	0.136* (0.040)	0.187 (0.050)	0.042* (0.019)
Observations	6485	10710	5465	11870
Degrees of freedom	3	2	1	6
VIF	11.93	57.16	80.14	7.36
[B] Local controls				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.020 (0.035)	0.104* (0.042)	0.101** (0.033)	0.047 (0.029)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	-0.016 (0.016)	0.052** (0.020)	0.063** (0.022)	0.015 (0.014)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.034 (0.022)	0.067 (0.053)	0.063 (0.056)	0.010 (0.016)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	-0.042 (0.046)	0.085 (0.053)	0.104* (0.048)	0.033* (0.018)
Observations	13390	16201	10956	18692
Degrees of freedom	10	6	5	14
VIF	7.10	8.45	8.35	6.26
[C] Paris region (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.127*** (0.015)	0.077** (0.022)	0.095** (0.017)	0.059* (0.025)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.046 (0.023)	0.024 (0.016)	0.037 (0.025)	0.030* (0.014)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.070** (0.016)	-0.010 (0.012)	0.014 (0.012)	0.008 (0.016)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.070** (0.012)	0.001 (0.023)	0.029 (0.013)	0.042* (0.019)
Observations	6590	13964	8719	11870
Degrees of freedom	3	4	3	6
VIF	5.24	12.25	7.87	7.36

Table 1.6 – *DiD estimates of new university openings, excluding the wave Génération 98*
(continued from previous page)

	Charente-Maritime ($c = CM$)	Nord-Pas-de-Calais ($c = NPC$)	Pas-de-Calais ($c = PC$)	Villes Nouvelles ($c = VN$)
[D] Rest of France (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.138*** (0.012)	0.065*** (0.015)	0.085*** (0.013)	0.017 (0.027)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.051*** (0.007)	0.045*** (0.017)	0.033*** (0.007)	0.008 (0.013)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^D	0.094*** (0.010)	0.030*** (0.009)	0.039*** (0.009)	0.006 (0.012)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.102*** (0.012)	0.044*** (0.013)	0.055*** (0.011)	0.028** (0.013)
Observations	81155	88529	83284	86435
Degrees of freedom	88	89	88	91
VIF	4.30	4.92	4.40	5.43

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).

(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. We consider $C - 2$ degrees of freedom, where C is the number of clusters in each regression; see Note (ii.) in Table 1.3.

(iii.) Each regression includes the set of county-specific time-varying variables, as well as individual variables, and two year-dummies (for the years 2004 and 2007); see Note (iii.) in Table 1.3.

1.6.2 Falsification tests

We have considered various falsification tests. Subsection 1.6.2 describes some tests which have been conducted with falsified treatment groups. Subsection 1.6.2 considers some tests with some falsified treatment periods. Finally, Subsection 1.6.2 studies the reverse causality hypothesis, i.e., the possibility that university openings are not the cause but the consequence of an increasing higher education demand.

Falsified treatment groups

Concerning the DiD regressions with a falsified treatment group, we have considered a “placebo” DiD wherein, excepted the bordering counties to avoid any spillover effect, the local controls for Charente-Maritime and Pas-de-Calais are assumed to be the treatment groups.²⁹ We

29. More precisely, it means that the falsified treatment group in the case of Charente-Maritime is composed of Loire-Atlantique, Maine-et-Loire, Vienne, Haute-Vienne, Dordogne, Lot-et-Garonne and Landes (i.e., the seven counties in yellow around the bordering counties of Charente-Maritime as described in Figure 1.1). In the case of Pas-de-Calais, the falsified treatment group is composed of Seine-Maritime, Oise, Marne and Ardennes (i.e., the four counties in yellow around the bordering counties of Pas-de-Calais as described in Figure 1.1).

have assigned to these falsified treatment groups the Paris region as a common control group to limit any potential substitution bias. For robustness, we have also estimated these regressions with only the Paris county as a control group. We expect the absence of any significant impact for two reasons. First, the falsified treatment groups are a priori not significantly impacted by the openings of new universities in the (true) treated counties. Second, the Paris region as a distant group ensures that no interaction happened. If we do find significant impact, it would cast suspicion on the validity of our control groups. However, as shown in Table 1.7, only 2 out of the 16 implemented regressions give estimates significantly different from zero. Furthermore, these 2 estimates are significant only at the 10 percent level.

Table 1.7 – Falsification tests: Falsified treatment groups

	Control group			
	Paris region		Paris county	
	Charente-Maritime ($c = CM$)	Pas-de-Calais ($c = PC$)	Charente-Maritime ($c = CM$)	Pas-de-Calais ($c = PC$)
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.042 (0.028)	0.056* (0.026)	0.016 (0.029)	0.065* (0.023)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.003 (0.017)	0.038 (0.029)	-0.039 (0.028)	-0.002 (0.021)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^D	0.020 (0.020)	0.026 (0.020)	0.016 (0.022)	0.031 (0.022)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.010 (0.024)	0.025 (0.026)	0.006 (0.028)	0.032 (0.021)
Observations	19581	16866	13385	10670
Degrees of freedom	9	6	6	3

Notes: (i.) Cluster-robust standard errors in parentheses (clustering on county).

(ii.) *, ** and *** refer to significance at the 10, 5 and 1% levels respectively. We consider $C - 2$ degrees of freedom, where C is the number of clusters in each regression; see Note (ii.) in Table 1.3.

(iii.) The falsified treatment group in the case of Charente-Maritime is composed of Loire-Atlantique, Maine-et-Loire, Vienne, Haute-Vienne, Dordogne, Lot-et-Garonne and Landes (i.e., the seven counties in yellow around the bordering counties of Charente-Maritime as described in Figure 1.1). In the case of Pas-de-Calais, the falsified treatment group is composed of Seine-Maritime, Oise, Marne and Ardennes (i.e., the four counties in yellow around the bordering counties of Pas-de-Calais as described in Figure 1.1).

(iv.) Each regression includes the set of county-specific time-varying variables, as well as individual variables, and two year-dummies (for the years 2004 and 2007); see Note (iii.) in Table 1.3.

Falsified treatment periods

Two falsification tests with a falsified treatment period have been implemented. First, we have considered a “placebo” DiD wherein we have assumed that 2004 was the pretreatment period, and 2007 the treatment period. We have also considered a second “placebo” DiD wherein we have exploited an additional wave, *Génération 10*, which concerns young individuals who

left the educational system between January 1st and December 31st, 2010.³⁰ In this second placebo, 2007 is assumed to be the pretreatment period, and 2010 the treatment period.

These falsification tests are confronted to various drawbacks which have to be mentioned. First, they remain posterior to the treatment assignment period. Thus the treated counties could still be affected by the treatment. Second, the French education system experienced various important changes at the end of 2007 as already noticed. Another important reform was implemented between 2008 and 2010, the French new Post-Secondary Enrollment System called *Admission Post-Bac*; its aim was to rationalize enrollment of the secondary education graduates into the various tracks in higher education. All these reforms could have affected differently the various counties and bias our falsification tests. Third, these falsification tests rely on estimations that do not control for time-varying county-specific variables because of the limited number of waves. Despite of these limitations, our two falsification tests are conducted over a period beginning more than fifteen years after the new university openings; recalling Figure 1.2, the impact of these creations is mainly observed during the nineties.

Tables 1.8 and 1.9 provide the results. The DiD estimates are not statistically significant, or at least much less significant than those of the original estimates in Table 1.3. To have an idea of that magnitude, note that 32 out of the 64 estimates in Table 1.3 are statistically significant at the 5 percent level, i.e., 50 percent. In the corresponding falsification tests, only 12 out of 64 in 2007-2010 (18.7 percent) and 10 out of 64 in 2004-2007 (15.6 percent) are. Most importantly, even when they differ significantly from zero, the estimates in the falsification tests are generally much lower in absolute value than those in Table 1.3. The most convincing results are obtained when the Paris region is used as a control group.

30. This wave was not considered in the estimates of Table 1.3, given that important reforms concerning the financial autonomy and the cycles of studies were implemented at the end of 2007; see Subsection 1.2.1.

Table 1.8 – Falsification tests: 2007-2010 (concluded on next page)

	Charente-Maritime ($c = CM$)	Nord-Pas-de-Calais ($c = NPC$)	Pas-de-Calais ($c = PC$)	Villes Nouvelles ($c = VN$)
[A] Local controls (bordering counties only)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.014 (0.022)	-0.073 (0.033)	-0.050 (0.033)	0.021 (0.018)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.006 (0.015)	0.026 (0.042)	0.055 (0.043)	-0.028 (0.031)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.028 (0.018)	0.011 (0.059)	0.042 (0.053)	0.017 (0.025)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.030 (0.019)	-0.015 (0.022)	0.023 (0.010)	0.024 (0.032)
Observations	3974	7229	3813	7231
Degrees of freedom	3	2	1	6
[B] Local controls				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	0.014 (0.017)	-0.049** (0.020)	-0.032 (0.018)	-0.007 (0.017)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.025 (0.030)	-0.005 (0.031)	0.017 (0.030)	-0.027 (0.019)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.050*** (0.014)	0.011 (0.026)	0.039 (0.021)	0.007 (0.026)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.051** (0.022)	-0.020 (0.019)	0.012 (0.010)	-0.016 (0.029)
Observations	7771	10389	6973	11791
Degrees of freedom	10	6	5	14
[C] Paris region (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	-0.004 (0.027)	0.006 (0.021)	0.024 (0.018)	0.021 (0.018)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.024 (0.030)	-0.003 (0.031)	0.017 (0.029)	-0.028 (0.031)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.045** (0.011)	0.018 (0.017)	0.045** (0.009)	0.017 (0.025)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.104** (0.019)	0.037 (0.023)	0.066** (0.018)	0.024 (0.032)
Observations	4172	8742	5326	7231
Degrees of freedom	3	4	3	6

Table 1.8 – Falsification tests: 2007-2010 (continued from previous page)

	Charente-Maritime (<i>c = CM</i>)	Nord-Pas-de-Calais (<i>c = NPC</i>)	Pas-de-Calais (<i>c = PC</i>)	Villes Nouvelles (<i>c = VN</i>)
[D] Rest of France (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	-0.031*** (0.009)	-0.013 (0.012)	0.003 (0.009)	0.008 (0.014)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	-0.002 (0.007)	-0.014 (0.013)	0.008 (0.007)	-0.029** (0.013)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	0.026*** (0.006)	-0.001 (0.014)	0.026*** (0.006)	0.001 (0.022)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.035*** (0.008)	-0.023 (0.016)	0.006 (0.007)	-0.034 (0.023)
Observations	46040	50610	47194	49099
Degrees of freedom	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. We consider $C - 2$ degrees of freedom, where C is the number of clusters in each regression; see Note (ii.) in Table 1.3.

(iii.) Each regression includes the set of individual variables; see Note (iii.) in Table 1.3.

Table 1.9 – Falsification tests: 2004-2007 (concluded on next page)

	Charente-Maritime (<i>c = CM</i>)	Nord-Pas-de-Calais (<i>c = NPC</i>)	Pas-de-Calais (<i>c = PC</i>)	Villes Nouvelles (<i>c = VN</i>)
[A] Local controls (bordering counties only)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	-0.042 (0.035)	0.093 (0.035)	0.118 (0.028)	0.036 (0.032)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	-0.012 (0.015)	-0.022 (0.014)	-0.031 (0.015)	0.050 (0.053)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	-0.067** (0.013)	-0.022 (0.036)	-0.044 (0.030)	-0.006 (0.033)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.0014 (0.037)	0.033 (0.025)	0.007 (0.022)	0.022 (0.042)
Observations	4405	6131	3020	6687
Degrees of freedom	3	2	1	6
[B] Local controls				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	-0.057** (0.018)	0.039 (0.026)	0.062** (0.022)	0.045* (0.023)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	-0.033 (0.025)	-0.003 (0.023)	-0.012 (0.022)	0.031 (0.029)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	-0.081*** (0.014)	-0.040* (0.019)	-0.062** (0.017)	-0.016 (0.029)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	-0.020 (0.022)	-0.019 (0.022)	-0.043* (0.019)	0.019 (0.033)
Observations	8825	8673	5562	10286
Degrees of freedom	10	6	5	14
[C] Paris region (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	-0.033 (0.028)	0.032 (0.032)	0.055 (0.029)	0.036 (0.032)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.029 (0.053)	0.025 (0.051)	0.014 (0.051)	0.050 (0.053)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	-0.059* (0.019)	-0.007 (0.023)	-0.029 (0.021)	-0.006 (0.033)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	-0.004 (0.030)	-0.003 (0.031)	-0.026 (0.029)	0.022 (0.042)
Observations	3746	7888	4777	6687
Degrees of freedom	3	4	3	6

Table 1.9 – Falsification tests: 2004-2007 (continued from previous page)

	Charente-Maritime ($c = CM$)	Nord-Pas-de-Calais ($c = NPC$)	Pas-de-Calais ($c = PC$)	Villes Nouvelles ($c = VN$)
[D] Rest of France (common control)				
DiD estimates on the probability of attaining at least a two-year post-secondary degree				
β_c^{2+}	-0.013 (0.009)	0.038** (0.018)	0.063*** (0.009)	0.037** (0.018)
DiD estimates on the probability of attaining at least a four-year post-secondary degree				
β_c^{4+}	0.011 (0.008)	0.003 (0.008)	-0.005 (0.008)	0.023 (0.017)
DiD estimates on the probability of obtaining the <i>Baccalauréat</i> and attending the university directly after				
β_c^O	-0.048*** (0.006)	-0.003 (0.011)	-0.023*** (0.007)	-0.004 (0.025)
DiD estimates on the probability of attaining a university degree as last diploma				
β_c^U	0.013* (0.008)	0.005 (0.012)	-0.016* (0.008)	0.029 (0.028)
Observations	47944	52086	48975	50885
Degrees of freedom	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. We consider $C - 2$ degrees of freedom, where C is the number of clusters in each regression; see Note (ii.) in Table 1.3.

(iii.) Each regression includes the set of individual variables; see Note (iii.) in Table 1.3.

Reverse causality

The last falsification tests deal with the possibility that the university openings occurred in counties where the demand for higher education was already increasing. If so, university openings are not the cause but the consequence of an increasing higher education demand. We cannot *a priori* completely rule out this reverse causality hypothesis because the decision of university openings was not totally independent of the demography and the increasing flow of high school graduates; see Section 1.2.2. If the university openings are really the cause, and not the consequence, then the demand for higher education, prior to the university openings, should not have any significant impact on the educational outcomes. Thus, to test the reverse causality assumption, we studied the impacts of some pre-trends in the demand for higher education, prior to the university openings, on the four outcomes of interest. More precisely, we have replaced the treatment dummies in the treatment-period interaction terms (e.g., the CM_c term of $(CM_c \times d_t)$ in Equation 2) with two different pre-trend measures: the variation in the number of students who obtained the *Baccalauréat général* over the period 1982-1990, and this variation over the period 1986-1990; we have also considered the growth rates of these variables. The students who obtain

the *Baccalauréat général* represent a very large majority of enrollments at the university.³¹ Note that these pre-trends are county-specific but they are not time-varying variables. Thus, to avoid multicollinearity and weak identification, the estimations were not implemented with the local control groups, but only with the rest of France as a control group. Table 1.10 shows that only 4 out of the 64 estimates are significantly different from zero at the 5 percent level. The four significant impacts concern only the probability of attending the university directly after the *Baccalauréat*³², and only the absolute growth measures of the pre-trends. This confirms that the increasing number of high school graduates, expected at the beginning of the 90s, has presumably participated to the decision process of creation. But none of the other probabilities of attaining at least a two-year or a four-year post-secondary degree, or a university degree as last diploma were impacted by the pre-trends. All in all, these falsification tests suggest that reverse causality, alone, cannot explain the positive and significant impacts estimated in Section 1.5 on the educational outcomes.

31. The *Baccalauréat général* is the non-vocational high school diploma; it is the main access to university and other studies in higher education. It is also possible to enter university with one of the two other *Baccalauréats*, the *Baccalauréat technologique* or the *Baccalauréat professionnel*. However these two *Baccalauréats* are mainly dedicated to pursue short technical studies or to enter professional life.

32. A simple calculation suggests an impact around 1.5 percentage points in Charente-Maritime and 5 percentage points in Pas-de-Calais on the probability of attending the university directly after High School. Indeed, multiplying the observed variation in the number of students who obtained the *Baccalauréat général* over the period 1986-1990 (in thousands, 0.663 for Charente-Maritime and 1.968 for Pas-de-Calais) by their respective estimated coefficient (in Table 1.10, $\widehat{\beta}_{CM}^O = 0.025$ and $\widehat{\beta}_{PC}^O = 0.026$), we obtain an estimated impact of 1.65 percentage points for Charente-Maritime and 5.17 for Pas-de-Calais. These estimates may be compared to the estimated effects in Part [D] of Table 1.3 which are 7.4 and 3.3 percentage points, respectively. This suggests that Pas-de-Calais would be more subject than Charente-Maritime to reverse causality in the probability of attending the university directly after High School. Proceeding similarly for the period 1982-1990, the results are very close: we find an estimated impact of 1.44 percentage points for Charente-Maritime and 4.03 for Pas-de-Calais (the variations in thousands in the number of students who obtained the *Baccalauréat général* are 0.687 for Charente-Maritime and 1.832 for Pas-de-Calais and the estimated coefficients $\widehat{\beta}_{CM}^O = 0.021$ and $\widehat{\beta}_{PC}^O = 0.022$).

Table 1.10 – Reverse causality tests: Are university openings caused by the demand for education?

	Charente-Maritime (<i>c</i> = <i>CM</i>)	Nord-Pas-de-Calais (<i>c</i> = <i>NPC</i>)	Pas-de-Calais (<i>c</i> = <i>PC</i>)	Villes Nouvelles (<i>c</i> = <i>VN</i>)
Pre-trend measure: Growth in the number of high school graduates between 1982-1990				
β_c^{2+}	0.013 (0.014)	0.004 (0.007)	0.017 (0.013)	0.012 (0.012)
β_c^{4+}	-0.008 (0.014)	-0.008 (0.010)	-0.007 (0.013)	-0.006 (0.012)
β_c^O	0.021** (0.008)	0.005 (0.005)	0.022*** (0.008)	0.014* (0.008)
β_c^U	0.009 (0.010)	-0.002 (0.008)	0.012 (0.010)	0.007 (0.009)
Pre-trend measure: Growth in the number of high school graduates between 1986-1990				
β_c^{2+}	0.021 (0.016)	0.012 (0.009)	0.027* (0.015)	0.018 (0.014)
β_c^{4+}	-0.006 (0.017)	-0.007 (0.012)	-0.003 (0.016)	-0.003 (0.015)
β_c^O	0.025** (0.009)	0.011* (0.0063)	0.026*** (0.009)	0.017* (0.009)
β_c^U	0.015 (0.012)	0.002 (0.008)	0.019 (0.011)	0.012 (0.011)
Pre-trend measure: Growth rate in the number of high school graduates between 1982-1990				
β_c^{2+}	0.000 (0.026)	-0.013 (0.024)	0.000 (0.026)	-0.010 (0.025)
β_c^{4+}	-0.007 (0.019)	-0.029 (0.020)	-0.006 (0.019)	-0.015 (0.018)
β_c^O	0.003 (0.019)	-0.008 (0.019)	0.003 (0.019)	-0.003 (0.018)
β_c^U	-0.018 (0.021)	-0.039* (0.021)	-0.018 (0.021)	-0.020 (0.019)
Pre-trend measure: Growth rate in the number of high school graduates between 1986-1990				
β_c^{2+}	-0.004 (0.039)	-0.016 (0.037)	-0.000 (0.039)	-0.019 (0.037)
β_c^{4+}	-0.024 (0.031)	-0.046 (0.030)	-0.022 (0.031)	-0.034 (0.029)
β_c^O	-0.000 (0.028)	-0.009 (0.027)	0.001 (0.028)	-0.008 (0.026)
β_c^U	-0.026 (0.031)	-0.047 (0.030)	-0.023 (0.031)	-0.033 (0.029)
Observations	128244	138589	131294	136720
Degrees of freedom	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. We consider $C - 2$ degrees of freedom, where C is the number of clusters in each regression; see Note (ii.) in Table 1.3.

(iii.) Each regression includes the set of county-specific time-varying variables, as well as individual variables, and two year-dummies (for the years 2004 and 2007); see Note (iii.) in Table 1.3.

1.7 Conclusion

This paper provides new evidence on the impact of university openings on the acquisition of human capital. New university openings have increased the probability of attaining at least a two years post-secondary degree by about 10 percentage points and the probability of attaining at least four years post-secondary degree by about 5 percentage points in counties which were

initially undereducated compared to the rest of France. Conversely, new university openings in relatively higher educated counties have not had a significant impact on the acquisition of human capital. This does not mean that opening a new university in a relatively higher educated county is useless: in the case of the counties of the *villes nouvelles* around the Paris region, and as noticed in Subsection 1.2.2, the aim of these creations was mainly to decongest Paris universities according to the public administrations involved in the creations and the evaluations of these universities. These results suggest that university openings help in reducing the gap between undereducated counties and the average level of higher education in a given country. We refer to this regional disparities reduction in the level of higher education as a catch-up effect.

In practice, the choice of an adequate control and treatment group appeared as a central task in the evaluation, to avoid potential spillover and substitution biases, but also to prevent from important measurement errors in the treatment location. A precise definition of the outcomes, net of displacement effects (ie. mainly net of decongestion effects), appeared essential. We also highlighted that childhood residential mobility may induce huge measurement errors in the treatment location, whereas student mobility may give rise to important displacement and spillover effects. All this underlines the importance of human capital mobility in these evaluations, and the advantage of micro data to cope with (at least some of) these difficulties. The impact of human capital mobility on these statistical biases would deserves further investigation, with more detailed micro data.

Our empirical findings provide a potential explanation for the results presented by [Valero and Van Reenen \(2019b\)](#), which highlights that the human capital channel has only a small role in the universities effect on growth. In this paper, we show that the impact of university openings is heterogeneous according to the place where the opening occurs. In developed regions, university openings are mostly aimed at decongesting existing universities and hence opening a new university does not have an impact on human capital. In contrast, the impact is large in developing regions as the main goal of university openings is to increase the supply of higher education. Therefore, the fact that the different reasons for opening a university are not taken into account in [Valero and Van Reenen \(2019b\)](#) in the 1500 regions studied, could explain the small relationship they find between university growth and college share. Finally, we also provided in this paper a first step for exploring the effects of university openings on labor market outcomes, such as employment and wages, through the skills channel, which deserves further investigation.

Chapter 2

The impact of university openings on labor market outcomes

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 percentage 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.

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

Keywords: Human capital, university openings, labor market outcomes

2.1 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 \(2012b\)](#) 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. In the first chapter, I also find a causal impact between university openings and human capital investments. They use individual data from CEREQ *Génération* 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 \(2019a\)](#) 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 \(1995b\)](#) 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 \(2001b\)](#) and [Moretti \(2004b\)](#). In this paper, I also explore this potentially causal link between College/University and schooling and labor market outcomes. What differentiates

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

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 labor 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" channel that underpins most of the empirical work done since Mincer (1974), where labor 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 labor 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 characterized 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 nil 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 2.2 provides the background of this study and presents the data. Section 2.3 presents the empirical framework. Section 2.4 gives the main results. Section 2.5 provides some robustness checks and I finally conclude in section 2.6.

2.2 Context and Data

2.2.1 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.

2.2.2 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.

CEREQ *Génération* surveys

CEREQ *Génération* surveys are large scale surveys realized by the CEREQ³ 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 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

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

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

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 2.1). 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.

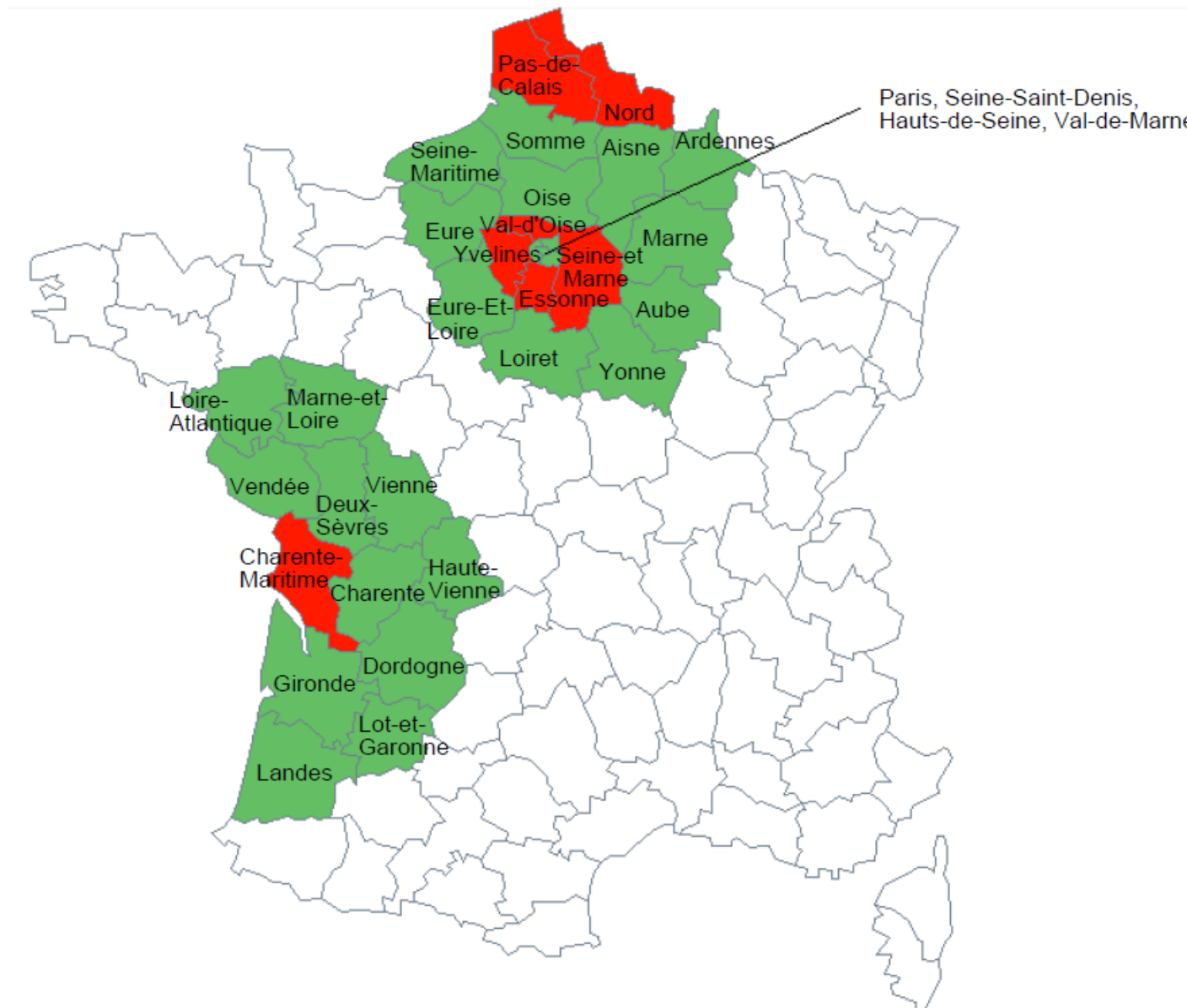


Figure 2.1 – Individuals assignment into treated and control groups

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, Eure-et-Loire, Loiret, Yonne, Aube, Marne and Aisne.

Les Enquêtes Emploi

Les Enquêtes Emploi realized by INSEE⁵ 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

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.

from the 1st 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.

2.3 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):

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

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

$$S = g(UO) = \lambda + \gamma_2 \times UO \quad (2)$$

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

$$Y = f(g(UO), UO) = \eta + \alpha \times UO \quad (3)$$

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)

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 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).

2.3.1 Total effect of university openings

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

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

$Y_{i,c,t}$ 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 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 labor market outcomes of those

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).

who have not attended university.

2.3.2 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.

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

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" γ_1 (i.e. which is here simply the DID coefficient purged from the contribution of EDUC in eq (5)) according to:

$$\hat{\alpha} = \gamma_1 + \gamma_2 \times \hat{\beta} \quad (6)$$

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):

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

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.

2.4 Results

In Table 2.1 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 neighboring counties to those where a university opening occurred (see Figure 2.1 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 2.1 shows that the opening of a new university in Charente-Maritime increases the probability to be employed by about 8 percentage points and decreases this probability by about 4 percentage points in Pas-de-Calais. Employment stability increases by 6 percentage points as a result of university openings in Charente-Maritime and decreases by 5 percentage points in Villes-Nouvelles. Finally, wages increase by about 5% in Charente-Maritime.

Before explaining the previous results, Table 1.1 shown in the first chapter, 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 1.1 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. In the first chapter, I show that opening a new university increases the probability to attain at least 2 years of higher education by about 10 percentage points, nonetheless results in Table 2.1 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 2.2 shows results. The negative impact on unemployment decreases and becomes almost non significant when controlling for the field of

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.

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 2.3 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 percentage points. This previous impact is negative in Nord and Pas-de-Calais at the level of 2 percentage points. Moreover, university openings seem to decrease the probability to work as a blue collar in Nord and Pas-de-Calais by about 3 percentage points. Results are consistent when I control for the field of the last degree attained (Table 2.4). 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 1.1 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 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 2.5, 2.6 and 2.7. 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 2.5, 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 unknown 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 2.8 gives the estimates of γ_2 , the effect of university opening on human capital in Equation 7. The results from Tables 2.5 and 2.6, when combined with Panel A results of Table 2.8 allow to illustrate the Gelbach decomposition formula described in Equation 6. In the case of Table 2.7 where the outcome variable is log monthly wages, results are to be combined with Panel B of Table 2.8 in order to illustrate the Gelbach decomposition formula. For example, from Table 2.5 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 2.8, 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:

$$\hat{\alpha} = \hat{\gamma}_1 + \hat{\gamma}_2 \times \hat{\beta} = 0.064 + 0.474 * 0.026 = 0.076$$

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 2.1 and 2.5).

Table 2.1 – DID estimates ($\hat{\alpha}$)

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

Table 2.2 – 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*** (0.016)	-0.022 (0.012)	-0.0215 (0.012)	0.027* (0.013)
# obs	20433	23608	16368	29429
Stable employment	0.055*** (0.014)	-0.025 (0.013)	-0.027* (0.013)	-0.047** (0.016)
# obs	20433	23608	16368	29429
log (wages)	0.060*** (0.013)	0.017** (0.005)	0.016** (0.006)	-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 2.1, 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 2.3 – DID estimates ($\hat{\alpha}$) by type of job

Dep var:	Charente-Maritime	Nord_Pas-de-Calais	Pas-de-Calais	Villes-Nouvelles
Executives	0.057*** (0.012)	-0.007 (0.011)	-0.022** (0.008)	0.020* (0.011)
Employee	0.008 (0.010)	-0.015 (0.009)	-0.010 (0.009)	-0.051** (0.022)
Blue collar	0.001 (0.011)	-0.021* (0.011)	-0.030** (0.009)	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 (2.1)

Table 2.4 – 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*** (0.012)	-0.001 (0.015)	-0.025* (0.012)	0.018 (0.012)
Employee	0.003 (0.010)	-0.023* (0.011)	-0.011 (0.010)	-0.044** (0.018)
Blue collar	-0.002 (0.010)	-0.019** (0.007)	-0.037*** (0.008)	0.007 (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 (2.1)

Table 2.5 – *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	$\hat{\alpha} = 0.076^{***}$ (0.015)	$\hat{\gamma}_1 = 0.064^{***}$ (0.016)	0.012*** (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	$\hat{\alpha} = -0.041^{**}$ (0.016)	$\hat{\gamma}_1 = -0.058^{***}$ (0.003)	0.017** (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	$\hat{\alpha} = -0.039^*$ (0.016)	$\hat{\gamma}_1 = -0.060^{***}$ (0.015)	0.021*** (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	$\hat{\alpha} = -0.029^*$ (0.015)	$\hat{\gamma}_1 = -0.041^{***}$ (0.014)	0.012* (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 2.1

Table 2.6 – 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	$\hat{\alpha} = 0.062^{***}$ (0.012)	$\hat{\gamma}_1 = 0.044^{***}$ (0.013)	0.018*** (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	$\hat{\alpha} = -0.035^{***}$ (0.014)	$\hat{\gamma}_1 = -0.054^{**}$ (0.016)	0.019** (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	$\hat{\alpha} = -0.039^{**}$ (0.014)	$\hat{\gamma}_1 = -0.063^{***}$ (0.014)	0.023*** (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	$\hat{\alpha} = -0.051^{**}$ (0.021)	$\hat{\gamma}_1 = -0.066^{***}$ (0.016)	0.014* (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 2.1

Table 2.7 – *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	$\hat{\alpha} = 0.048^{***}$ (0.012)	$\hat{\gamma}_1 = -0.002$ (0.008)	0.050*** (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	$\hat{\alpha} = 0.017^*$ (0.007)	$\hat{\gamma}_1 = -0.013$ (0.010)	0.031*** (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	$\hat{\alpha} = 0.010$ (0.006)	$\hat{\gamma}_1 = -0.022^{**}$ (0.006)	0.032*** (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	$\hat{\alpha} = -0.006$ (0.016)	$\hat{\gamma}_1 = -0.031^{**}$ (0.011)	0.025** (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 2.1

Table 2.8 – 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*** (0.067)	0.300* (0.122)	0.367** (0.092)	0.310* (0.167)
# obs	20433	23608	16368	29429
Panel B: Restricting the sample to individuals who are employed				
$\hat{\gamma}_2$	0.759*** (0.110)	0.488*** (0.0656)	0.504*** (0.057)	0.361** (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 2.1.

2.5 Robustness

2.5.1 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 sector of industry in the set of control variables I rule out this hypothesis. Results are depicted in Table 2.9. 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.

2.5.2 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 neighboring 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 neighboring 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 neighboring counties where university openings did not occur (Seine-Saint-Denis, Val-De-Marne and Hauts-De-Seine). Results are depicted in Table 2.10. 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.

2.5.3 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 might be unobserved determinants of labor market outcomes that are correlated to schooling, hence schooling is not necessarily 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 2.13, 2.14 and 2.15. 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 2.5, 2.6 and 2.7.

2.5.4 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 known 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 group 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 2.11. Estimates are similar to main results estimates in Table 2.1, 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 2.12. 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 2.2), i.e., the negative effect of UO on employment is not statistically different from zero.

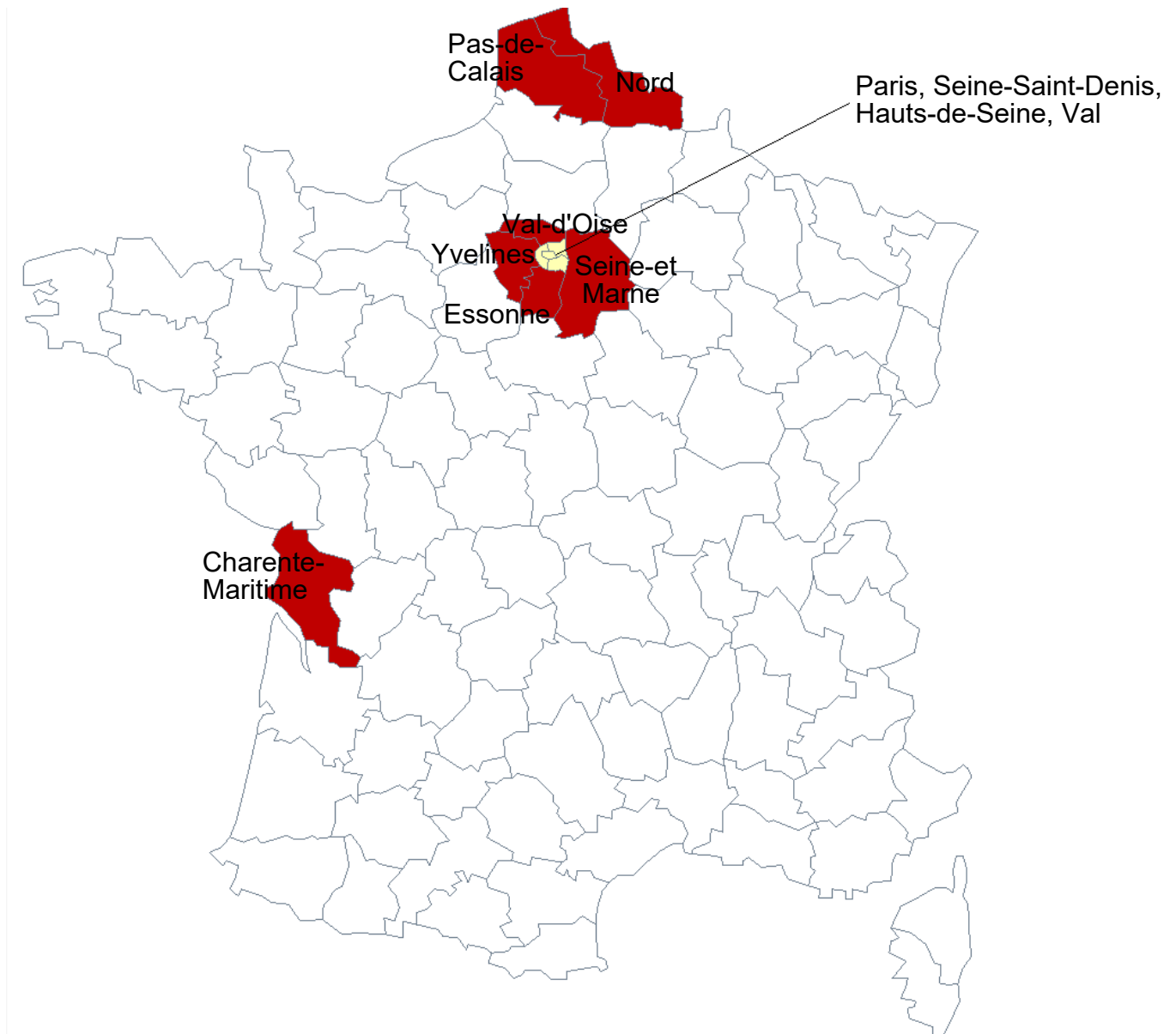


Figure 2.2 – Individuals assignment into treated and control groups (alternative control group1)

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.

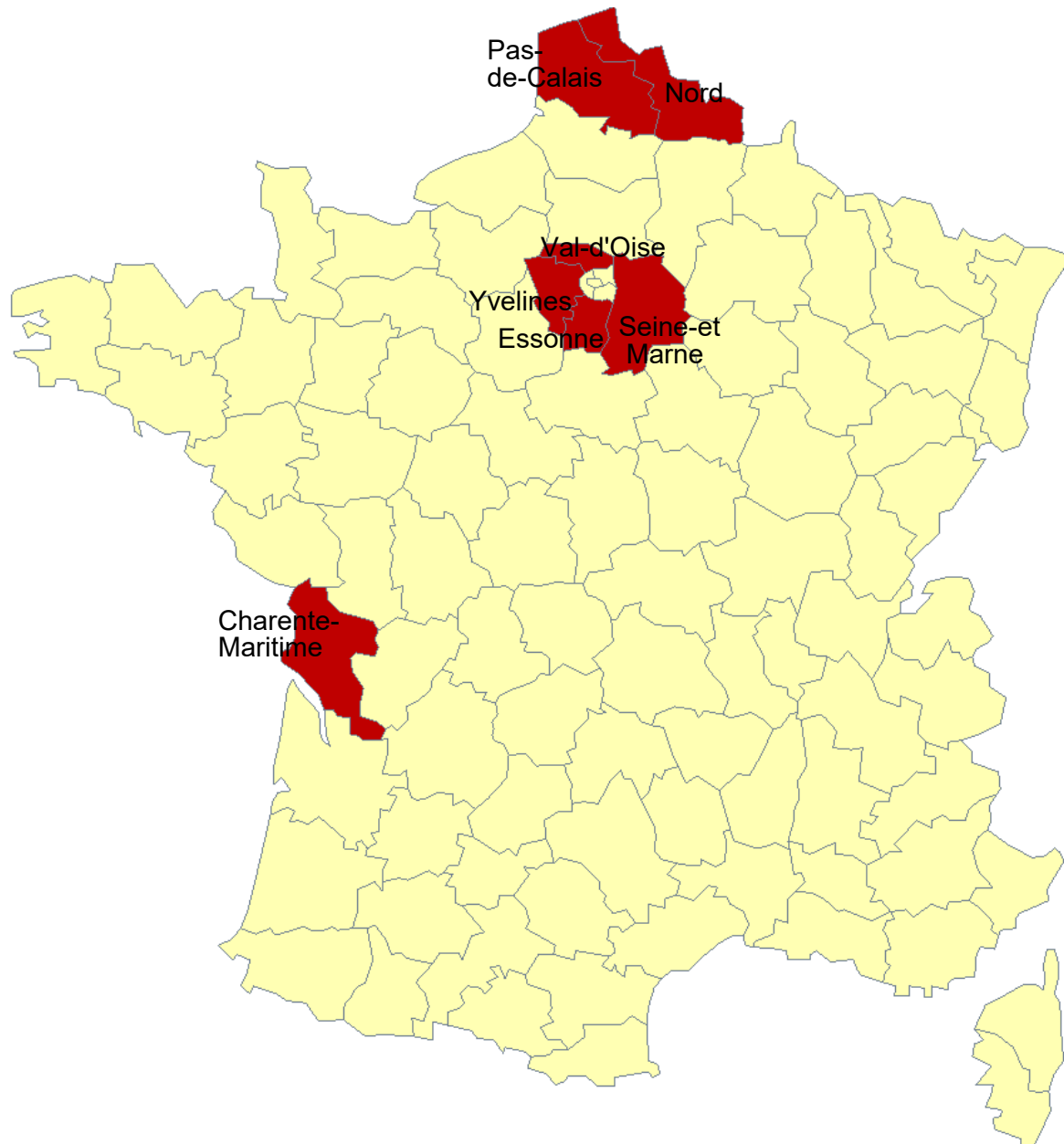


Figure 2.3 – *Individuals assignment into treated and control groups (alternative control group2)*

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 2.9 – 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*** (0.013)	0.022** (0.008)	0.0132* (0.006)	-0.005 (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 2.1, 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 2.10 – 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*** (0.018)	0.046** (0.016)	0.029 (0.021)	0.004 (0.020)
# obs	10038	20181	13031	18332
Stable employment	0.025 (0.022)	-0.047* (0.021)	-0.084** (0.022)	0.011 (0.020)
# obs	10038	20181	13031	18332
log (wages)	0.051** (0.014)	0.036* (0.015)	0.027 (0.016)	0.004 (0.016)
# obs	5738	11178	7464	11095
Df	3	4	3	6
Remaining counties of France as a control group				
Employment	0.088*** (0.007)	-0.010 (0.008)	-0.016** (0.007)	-0.057*** (0.013)
# obs	124812	134955	127805	133106
Stable employment	0.023*** (0.005)	-0.047*** (0.006)	-0.050*** (0.005)	-0.002 (0.010)
# obs	124812	134955	127805	133106
log (wages)	0.045*** (0.004)	0.031** (0.013)	0.014*** (0.005)	-0.020* (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 2.1

Table 2.11 – 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*** (0.015)	-0.034 (0.020)	-0.046* (0.020)	-0.032** (0.014)
# obs	20433	23608	16368	29429
Stable employment	0.082*** (0.011)	-0.026 (0.016)	-0.034 (0.018)	-0.051** (0.021)
# obs	20433	23608	16368	29429
log (wages)	0.048*** (0.011)	0.021** (0.007)	0.016** (0.006)	-0.002 (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 2.1.

(iii) weights used are propensity score weights w_i multiplied by the sample weights à la Ridgeway et al. (2015).

Table 2.12 – 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*** [0.000]	0.036*** [0.001]	-0.005 [0.622]	-0.074*** [0.000]
# obs	20433	23608	16368	29429
Stable employment	0.031*** [0.000]	0.062*** [0.000]	0.019* [0.091]	-0.084*** [0.000]
# obs	20433	23608	16368	29429
Log(wages)	0.045*** [0.000]	0.034*** [0.001]	-0.010 [0.256]	-0.035*** [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, executive, 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 2.13 – 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 Nou- velles
First stage				
Grade repetition before 6th grade	-1.961*** (0.061)	-1,722*** (0.066)	-1.666*** (0.068)	-0.041*** (0.013)
Treat x Post	0.474*** (0.067)	0.300** (0.122)	0.367*** (0.092)	0.301* (0.167)
Second stage				
Treat x Post	0.058*** (0.016)	-0.057*** (0.015)	-0.059*** (0.014)	-1.828*** (0.065)
educational attainment	0.039*** (0.005)	0.056*** (0.007)	0.054*** (0.012)	0.039*** (0.006)
#obs	20,433	23,608	16,368	29,429
Df	10	6	5	14

Notes:(i) Grade repetition 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 2.1.

Table 2.14 – 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 Nou- velles
First stage				
Grade repetition before 6th grade	-1.961*** (0.061)	-1.722*** (0.066)	-1.666*** (0.068)	-1.828*** (0.065)
Treat x Post	0.474*** (0.067)	0.300** (0.122)	0.367*** (0.092)	0.310* (0.167)
Second stage				
Treat x Post	0.035** (0.014)	-0.054*** (0.015)	-0.063*** (0.013)	-0.065*** (0.014)
educational attainment	0.057*** (0.006)	0.062*** (0.004)	0.063*** (0.007)	0.050*** (0.006)
#obs	20,433	23,608	16,368	29,429
Df	10	6	5	14

Notes:(i) Grade repetition 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 2.1.

Table 2.15 – 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 Nou- velles
First stage				
Grade repetition before 6th grade	-2.022*** (0.075)	-1.831*** (0.069)	-1.807*** (0.093)	-1.882*** (0.094)
Treat x Post	0.759*** (0.110)	0.488*** (0.066)	0.504*** (0.057)	0.361** (0.132)
Second stage				
Treat x Post	0.006 (0.011)	-0.017* (0.008)	-0.026*** (0.006)	-0.031** (0.011)
educational attainment	0,072*** (0.006)	0.071*** (0.004)	0.071*** (0.006)	0.077*** (0.004)
#obs	10,978	12,854	9,172	17,545
Df	10	6	5	14

Notes:(i) Grade repetition 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 2.1.

2.6 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 percentage 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.

Appendices

2.A 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:

$$w_i = \frac{e_1(X_i)}{e_g(X_i)} \quad (8)$$

$e_1(X_i)$ is the probability for individual i to be in group 1, and $e_g(X_i)$ 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, executive, 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. $W_i = w_i * \text{pondef}$) 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:

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

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:

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

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

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

With

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

$$w_{0,1}(Treat, Post, X) = Post \times (1 - Treat) \times \frac{e_1(X_i)}{e_3(X_i)}$$

$$w_{0,0}(Treat, Post, X) = (1 - Treat) \times \frac{e_1(X_i)}{e_4(X_i)}$$

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:

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

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 ($Treat = 1$ and $Post = 1$) and individuals in treatment group before the treatment ($Treat = 1$ and $post = 0$) respectively.

2.B 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

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

on vocational training in the technological and tertiary sectors. (CNE Évrý 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, experimental 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 DEUG¹³ (a two year university degree), a DEA¹⁴ (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

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.

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 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 techniques of physical and sports activities) at Lévin.

This multipolarity, makes it difficult to communicate across disciplines and reinforces the local specialization 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 Béthune 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 specializations, 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))

2.C 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 2.16, 2.17 and 2.18) and are consistent with results using larger definition of educational attainment measured by years of schooling (Tables 2.5, 2.6 and 2.7). The effect of UO on labor market outcomes that goes through human capital channel is always positive and statistically significant.

Table 2.16 – 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	$\hat{\alpha} = 0.076^{***}$ (0.015)	$\hat{\gamma}_1 = 0.067^{***}$ (0.014)	0.009*** (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	$\hat{\alpha} = -0.041^{**}$ (0.016)	$\hat{\gamma}_1 = -0.060^{***}$ (0.019)	0.020** (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	$\hat{\alpha} = -0.039^*$ (0.015)	$\hat{\gamma}_1 = -0.065^{**}$ (0.018)	0.026*** (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	$\hat{\alpha} = -0.029^*$ (0.015)	$\hat{\gamma}_1 = -0.037^{***}$ (0.014)	0.008** (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 2.1

Table 2.17 – *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	$\hat{\alpha} = 0.062^{***}$ (0.012)	$\hat{\gamma}_1 = 0.045^{***}$ (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	$\hat{\alpha} = -0.035^{**}$ (0.014)	$\hat{\gamma}_1 = -0.058^{**}$ (0.018)	0.023 ^{**} (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	$\hat{\alpha} = -0.039^{**}$ (0.014)	$\hat{\gamma}_1 = -0.069^{***}$ (0.016)	0.030 ^{***} (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	$\hat{\alpha} = -0.051^{**}$ (0.021)	$\hat{\gamma}_1 = -0.060^{***}$ (0.018)	0.010 ^{**} (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 2.1

Table 2.18 – 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	$\hat{\alpha} = 0.048^{***}$ (0.012)	$\hat{\gamma}_1 = 0.012$ (0.011)	0.036*** (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	$\hat{\alpha} = 0.017^*$ (0.007)	$\hat{\gamma}_1 = -0.007$ (0.010)	0.024*** (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	$\hat{\alpha} = 0.010$ (0.006)	$\hat{\gamma}_1 = -0.016$ (0.009)	0.026*** (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	$\hat{\alpha} = -0.006$ (0.016)	$\hat{\gamma}_1 = -0.031^{**}$ (0.011)	0.025** (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 2.1

Chapter 3

Entry into Marriage, Motherhood and the Arab Spring: Evidence from Egypt

Abstract

The Egyptian Revolution has been shown to have triggered important economic and social changes, including the reduction of gender inequalities. However, few has been said on the impact of the Arab Spring on the age of entry into marriage and motherhood, which are key issues for women's welfare. To shed light on this question, we combine a dataset of the Egyptian Revolution, with the 2018 wave of the Egyptian Labor Market Panel Survey. We rely on quasi-experimental geographical and historical variations in the level of violence, to build a difference-in-differences analysis. Our main findings are that women residing in rural areas who were aged between 16 and 20 at the time of the Egyptian Revolution, marry earlier than the previous cohorts, and have children earlier as well. These results tend to moderate previous evidence on the increase in women's empowerment following the Arab Spring.

JEL classification:J12, J16, 055

Keywords: Arab Spring, Marriage, Fertility, Egypt

3.1 Introduction

The Egyptian Revolution, during which almost 1000 civilians were killed and more than 6000 injured, has been shown to have triggered important economic and social changes.

El-Mallakh et al. (2018) have shown that the Arab Spring has decreased intra-household differences in labor force participation, and Bargain et al. (2019) have highlighted that these events have improved women's empowerment by increasing their final say within the household regarding decisions on health and expenditures. However, the Revolution's positive impact on women's situation in the labor market is not that obvious. Conversely, Hendy (2015) argues that the January 25th Revolution has had a negative effect on women's status in the labor market. Labor force participation has decreased, and unemployment has increased. These contrasting findings invite us to investigate the determinants of women's participation in the labor market, mainly marriage and fertility. This investigation is all the more important in a context where very few married women work.

Few has been said on the impact of the Arab Spring on the age of entry into marriage and motherhood, which are not only key determinants for women's participation in labor market

but also essential for women's welfare. Beyond the health problems and lower educational attainment when women marry and give birth at young age, women's age at marriage impact their bargaining power within the household. [Moeeni \(2021\)](#) for instance shows that women that are less educated have lower bargaining power. Now that the Revolution took place almost 10 years ago, we are able to look at its impact on these outcomes.

To shed light on the question, we combine a dataset of the Egyptian Revolution, which provides information on the number of people killed, injured and arrested during the entire period of the Egyptian Arab Spring, with the 2018 wave of the Egyptian Labor Market Panel Survey. We rely on quasi-experimental geographical and historical variations in the level of violence, to build a difference-in-differences analysis. We use variation over space and over time, in particular women's place of residence and their birth cohort. We use binary outcomes, defined for the whole sample (such as being married or not before a certain age), and continuous variables such as age at marriage and age at first child. Since the last variables are right censored, we use duration models to take this into account.

Our main findings are that women residing in rural areas who were aged between 16 and 20 at the time of the Egyptian Revolution, marry earlier than the previous cohorts, and have children earlier as well. These results tend potentially to moderate previous evidence on the increase in women's empowerment following the Arab Spring. The discrepancies with previous research may be explained by the fact that we are looking at longer term outcomes. The impact we also find concerns a specific type of women, the ones residing in rural areas. This paper's findings raise important questions in terms of the repercussions that the decrease in the age of marriage could have on women's economic activity and bargaining power within the household.

3.2 Literature review

Conflicts and entry of women into marriage and fertility The Egyptian revolution counted 1000 civilian casualties and is responsible for the injury of at least 6000. As our research exploits the occurrence of violence in Egypt, we contribute to the literature on the consequences of conflict for women's welfare in general and their marital and fertility decisions in particular. There is no clear-cut conclusion in the literature in terms of the direction and magnitude of the impact of conflict on marriage and motherhood. [Shemyakina et al. \(2009\)](#) provide evidence

of the armed conflict in Tajikistan between 1992-1998 and show that an economic crisis when associated with armed conflict delays age at first marriage for women. [Jayaraman et al. \(2009\)](#) also support the hypothesis that conflict delay age at marriage, by considering the 1994 genocide in Rwanda. They find that the conflict delays marriage in regions where people experience more siblings' deaths, but accelerates the entry into fertility in places with higher levels of under-five mortality. They conclude that the conflict affected age at marriage through disruption in the age structure and the sex ratio in the aftermath of the genocide, and the decomposition of kinship in the case of women who lost their siblings. On the contrary, in the context of Cambodia, [Saing and Kazianga \(2019\)](#) find that bombing reduced age at marriage, potentially due to reduced education attainment or old-age security motives. Taking evidence from Nepal decade-long conflict, [Williams et al. \(2012\)](#) find that conflicts increase the likelihood of marriages, most probably since single people were more at risk of imposed conscription and forced labor. In addition, conflicts have also been shown to increase fertility, with negative consequences for women and their children ([Nepal et al. \(2018\)](#) in the context of Nepal). Therefore, the impact of conflicts on age at marriage and age at first child seem to depend on the context.

Egyptian Context In the particular context of the Egyptian revolution conflict, the consequences on women's welfare are not that obvious. On the one hand, [El-Mallakh et al. \(2018\)](#) show that the Revolution has decreased intra-household differences through increasing labor force participation of women compared to men in both private and public sectors. In another paper, [Bargain et al. \(2019\)](#) show that 2011 Arab Spring protests in Egypt improved women empowerment by increasing their final say within the household regarding decisions on health expenditures. Women also showed reduced tolerance towards domestic violence in places mostly affected by the protests. However, these studies focus on the short term (one year after the revolution) and do not look at the impact on age at marriage. The positive impact on women's welfare could for instance be compatible with a delay in age at marriage and entry into motherhood, if the Egyptian Revolution introduced a change in norms pertaining to women's position and status in the society. On the other hand, [Hendy \(2015\)](#) shows that the 2011 Egyptian revolution had a negative impact on women's status in the labor market. This mixed evidence in the literature on the impact of Arab Spring on women's welfare calls for further investigation. In this paper, we attempt to provide a part of the answer to support one of the two strands of literature.

Marriage and fertility in Egypt Egypt witnessed a reversal of the trend in the total fertility rate (TFR) before the 2011 Revolution. It first decreased in 2000 and continued to decrease till 2007, after which it increased again. Goujon and Al Zalak (2018), show that according to the Egypt Demographic and Health Survey (EDHS) the TFR increased from 3 to 3.5 children per woman between 2008 - 2014. And the most rapid increase was observed in rural areas. Goujon and Al Zalak (2018) argue that this rise in fertility is more likely to be the consequence of an increase in women unemployment. Krafft (2020) observes a decrease in the public sector jobs that coincides with the increase in fertility. The public sector in Egypt is much more attractive to women compared to the private sector since it is easier to balance with marital life in terms of lower working hours and higher job security. Nevertheless, she suggests that the decrease in public sector employment is unlikely to be the main reason for the rise in fertility observed in Egypt starting 2008. She explains that the reasons behind the rise in fertility in Egypt is a complex question including different factors and is subject of debate and discussion. This paper attempts to fill this gap in the literature in the Arab Spring by investigating to which extent the Arab Spring has affected the fertility trend. We focus on age at marriage and age of entry into fertility, as these are less studied aspects in the literature reflecting women's well-being.

3.3 Conflict setting

The Egyptian Revolution refers to a long series of protests and upheavals, marked by its starting date of January 25, 2011 that coincides with "Police day" yearly celebrations. Thousands of protesters inspired by the Tunisian revolution gathered in "Tahreer square" in the capital city Cairo and several Egyptian cities at the same time. The protesters demonstrated peacefully demanding the government to leave by chanting slogans as "The people want to bring down the regime". A curfew was imposed in Egypt three days after the beginning of protests; however, this did not stop protesters so that many of them got arrested, injured or killed in an attempt from the police and army to contain the situation and put an end to demonstrations. 18 days later, on February 11, 2011 President Hosni Mubarak resigned, thereby ending the first phase of violence. The Supreme Council of the Armed Forces ruled Egypt until elections were held and Mohamed Morsi, candidate of the Muslim brotherhood was elected in June 2012.

During the rule of the Supreme Council, violence did not stop as clashes between Egyptian

soldiers and civilians continued in a series of events leading to more victims. The victory of removing the head of the ancient regime was only the beginning for Egyptian protesters. They demanded the resignation of remaining regime figures and to stand trials for all the corrupt former officials and those responsible for the mass killings during the 18 days leading to Mubarak's resignation. Protesters eventually demanded the dissolution of the Supreme Council as they were frustrated about the slow pace of economic reform and the slow process of investigations to prosecute former regime figures.

President Morsi has faced fierce opposition, after releasing a declaration that immunizes his decrees from challenge in an attempt to pass an Islamic oriented constitution. The declaration was criticized by the constitution party leader and led secular and liberal groups to leave the constituent party. Violent protests took place across Egypt demanding the dissolution of the constituent party and the reversal of Morsi's declaration. By the first anniversary of Morsi's election on 30th of June 2013, millions of Egyptians protested asking he steps down from office. 3 days later he was overthrown, and a civilian senior jurist Adly Mansour was appointed interim president. Morsi and his supporters, mainly from Islamic parties refused his removal. Sit-ins were organized but quickly met security forces interventions leaving hundreds of killed civilians. On March 26, 2014, Abdel Fattah el-Sissi the head of the Egyptian Armed Forces, resigned from the military, and announced himself candidate for presidential election. He took office as president on 8th of June 2014.

As the geography of the protests will play an important role in our identification strategy, we want to highlight it here: the protests took place mainly in big cities. Nevertheless, protests cover almost all the Egyptian territory and rural areas were also impacted. We investigate the potential correlation of the protest with the pre-existing economic situation in Table 3.1. The number of people killed by governorate is not correlated with the share of unemployment before the Revolution (in 2006). In the next section, we show the evolution of our major outcomes, according to the intensity of the Revolution, thereby highlighting that there were no differences in pre-Arab Spring trends between high and low treated areas.

Table 3.1 – Geography of the Revolution

Descriptive statistics	(1)	(2)	(3)
	Mean	Mean	Diff.
	Treated	Not treated	
Unemployment level	0.11	0.09	-0.02
% Rural	0.15	0.74	0.58***
Secondary education or higher	0.81	0.76	-0.05
Number of governorates	6	15	21

Note: Column (1) reports the mean of each variable listed for the 6 more impacted governorates. Column (2) reports the mean of each variable listed for the other governorates. Column (3) reports the difference between these two groups and the significance level from a t-test.

Significance levels are denoted as follows: + $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.4 Data

3.4.1 The ELMPS

The Egyptian Labor Market Panel Survey is a nationally representative panel survey who was first conducted in 1998. Three waves followed in 2006, 2012 and then in 2018, directed by the Economic Research Forum (ERF) in cooperation with Egypt's Central Agency for Public Mobilization and Statistics (CAPMAS). A total of 30,698 women have been interviewed.

In this paper, we use the 2018 wave, conducted 7 years after the Revolution, which gives us a long enough period to study the consequences of the Arab Spring on marriage and fertility. The 2018 wave of the ELMPS provides the complete record of marital and fertility life for interviewed women. We are also able to locate women at the time of revolution as the data provides the place of residence at birth, current place of residence and all the changes in residence in between. Our two main outcomes of interest are the age at marriage and the age at first child.

3.4.2 The Egyptian Revolution database

Our second source is "Wikithawra", a database on the Egyptian Revolution that reports the list of individuals who were arrested, injured or killed during the entire duration of the Egyptian Revolution, i.e., January 25 2011, till January 31, 2014. Detailed information on the exact location of each person arrested, killed or injured, is given along with the date of the event, which allows to properly locate killed individuals into governorates and districts. We use information

on killed individuals only as this is the most accurate one, parts of the reported data on arrested and injured persons is missing. Also, it is less likely to mistakenly report a killed individual in comparison to an injured one. In addition, previous papers such as [El-Mallakh et al. \(2018\)](#) and [Bargain et al. \(2019\)](#) also use only the data on fatalities, allowing for a better comparison .

3.4.3 Definition of the Treatment

We use two definitions of the treatment. The first one, that we call Treatment 1, is a commonly used measure of the intensity of violence: the number of people killed during the whole period of the Revolution (January 25, 2011 - January 31, 2014), relative to the number of people living in the governorate¹. The second one, Treatment 2, is a dummy taking the value 1 for the 6 governorates with the highest share of killed people, and zero otherwise. We use this second definition to capture a potential non-linearity in the effect of violence, it could for example be that only the most violent areas drive the effect. The map in [Figure 3.1](#) and the data presented in [Figure 3.2](#) visualize that data. The map shows the geographical distribution of the event. The red parts are the governorates with the highest share of killed individuals. In [Figure 3.2](#) we choose these 6 governorates as their share of killed people is significantly higher compared to the other governorates. We present alternative measures of the treatment in the robustness checks, were we compute treatment measures at the district level.

We present alternative measures of the treatment in the robustness checks. We compute treatment measures at the district level as well.

1. we use population census that reports the number of individuals by governorate

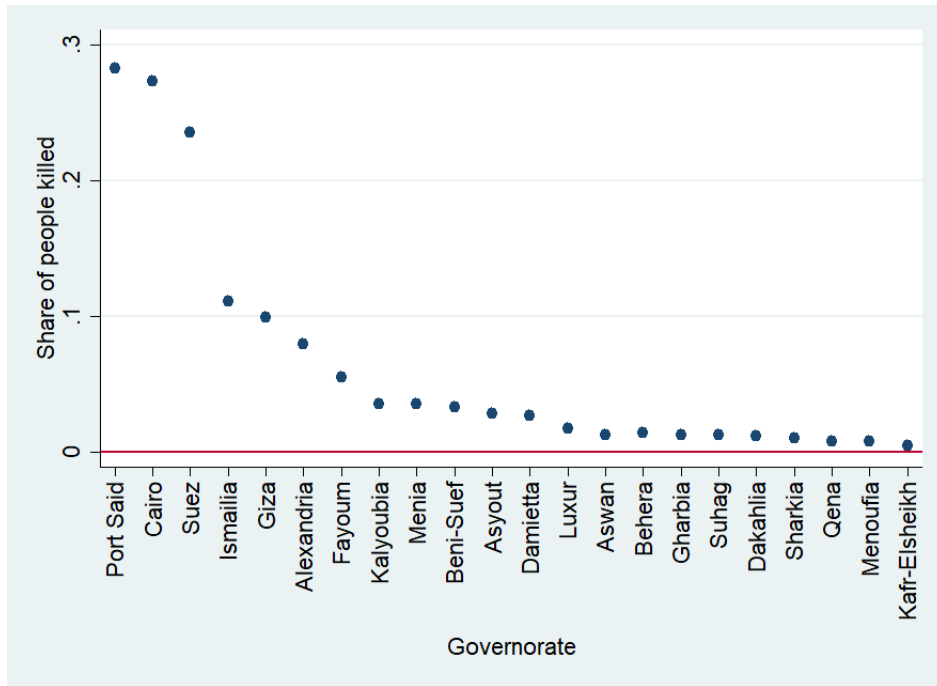


Figure 3.1 – Number of people killed during the protests for 1000 inhabitants by governorates

Note: The figure presents the number of people killed for 1000 individuals by governorate.

Source: Egyptian Revolution database

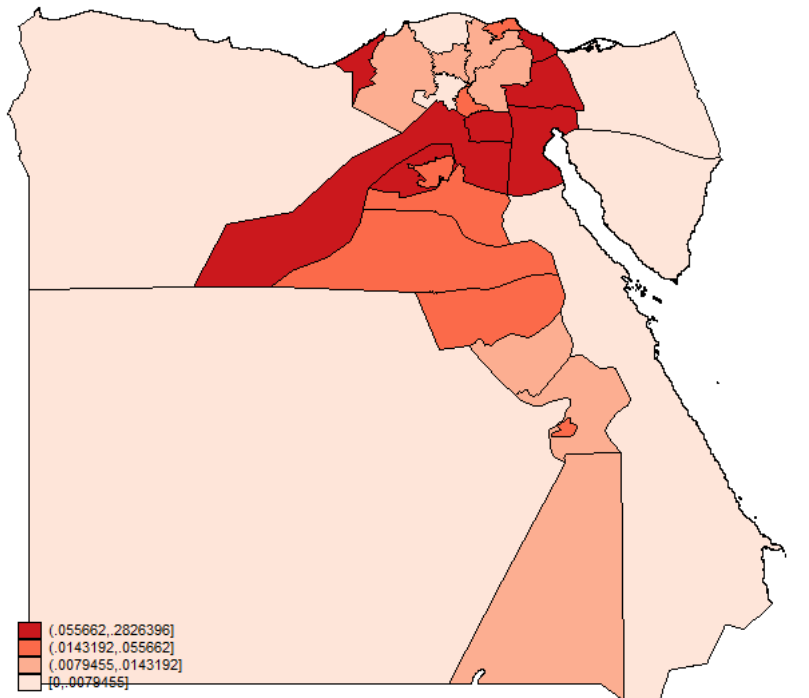


Figure 3.2 – Geographic distribution of people killed during the protests for 1000 inhabitants by governorates

Note: The map presents the number of people killed for 1000 individuals by governorate.

Source: Egyptian Revolution database

3.4.4 Descriptive Statistics

Table 3.2 presents some descriptive statistics for women interviewed in our sample by type of residence in January 2011 at the time of Revolution (rural /urban). Table 3.2 shows that the average age at marriage and age at first birth is slightly lower for rural women compared to urban women. The education level is also lower for rural women, and the violence intensity measure by the share of people killed during the Revolution is lower in rural areas.

Table 3.2 – Descriptive statistics

	All women	Rural women	Urban women
Age at first marriage	20.945 (4.496)	20.505 (4.343)	22.639 (4.687)
Age at first child	22.390 (4.258)	21.979 (4.144)	23.949 (4.353)
No schooling	0.248 (0.432)	0.279 (0.448)	0.137 (0.344)
Secondary education and above	0.323 (0.468)	0.346 (0.476)	0.529 (0.499)
Mother secondary education and above	0.282 (0.450)	0.187 (0.390)	0.324 (0.468)
Father secondary education and above	0.345 (0.475)	0.254 (0.436)	0.392 (0.488)
Currently working	0.670 (0.470)	0.655 (0.475)	0.710 (0.454)
Share of killed people	0.046 (0.066)	.0265 (.026)	0.113 (0.108)
Cohort 16-20	2562	1983	509
Cohort 26-30	2257	1667	524

Note: First column reports the mean for some of the variables characterizing women in the whole sample. Second column is for women in rural areas and third column is for women in urban areas. Standard errors are in parenthesis.

Source: ELMPS Egypt 2018.

3.5 Estimation approach

3.5.1 Graphical Evidence

Figure 3.3 plots the average age at marriage, according to the age of the respondent in January 2011 (beginning of the Egyptian Revolution) and compares women who lived in the 6 more affected governorates (in terms of the number of people killed) with women in the less affected governorates. The figure shows that women who were 27 years or older in January 2011 (meaning born before 1994) married at 23 years on average in the more affected areas, whereas this was on average 21 years old in less affected areas. The gap between the two areas is stable, when comparing with the older cohort of women being 35 years old in 2011. On the contrary, for younger cohorts, the age at marriage seems to decrease everywhere, but more strongly in highly affected areas. As a result, the gap between highly and the less affected areas seems to drastically reduce.

This figure however only tells part of the story. Indeed, age at marriage is a right-censored variable, since it is only known for individuals who were already married at the time of the survey. Younger cohorts have a higher probability not to be married at this time, and it is even more the case in urban areas. The computation of age at marriage averages for these younger cohorts will mechanically downward bias the results, even more strongly for urban areas where the age at marriage is higher relative to rural areas (see Table 3.2). There are two ways to correct for this type of error: either use duration model (which we will do in Subsection 3.5.3) or look at variables that are not right-censored, such as to be married before a certain threshold, as presented in the next subsection.

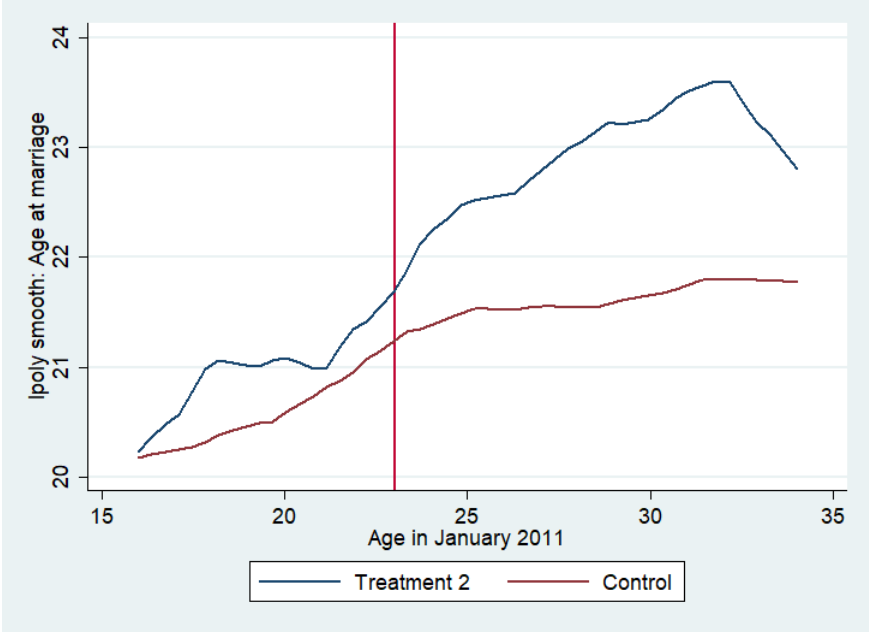


Figure 3.3 – Average age at marriage

Note: The figure presents the average age at marriage according to the age in January 2011.

Source: ELMPS 2018

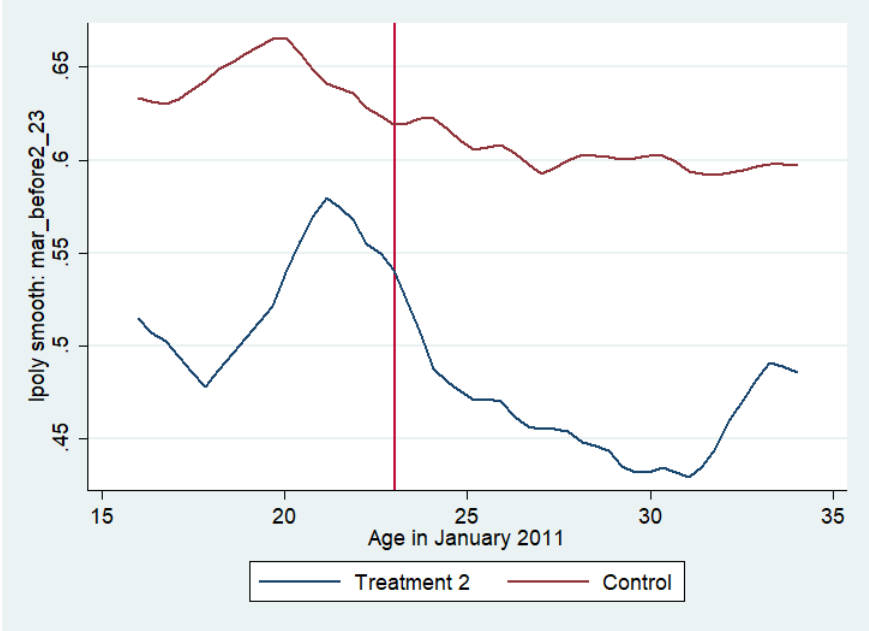


Figure 3.4 – Probability of marriage before 23

Note: The figure presents the probability of marriage before 23 according to the age in January 2011.

Source: ELMPS 2018

3.5.2 Difference-in-Differences

We use a standard difference-in-differences method. We compare individuals according to their governorate of residence and their birth cohort: women being between 16-20 years old at the time of the Revolution are considered as treated, while the control group is composed of women between 26 and 30 years old. We choose this age group since women start to marry at 18 years old (in 2018, 6% of women of 17 years old are married, versus 15% of women of 18 years old). At the age of 26, 70% of a cohort is married. Therefore, women who are older than 26 at the time of the Egyptian Revolution are far less susceptible to be affected in terms of entry into marriage.

We estimate the following model:

$$y_{igc} = a_0 + \beta * Post_i + \theta * Treat_g + \delta * Treat_g * Post_i + \alpha_g + \eta X_i + \varepsilon_{igc} \quad (1)$$

where y_{imc} is the outcome of interest for individual i residing in governorate g and born in year c ; a_0 is a constant; α_g is a governorate of residence fixed effect. $Post_i$ is a dummy variable equal to 1 if the individual was between 16 and 20 years old in January 2011. It is equal to 0 if she was between 26 and 30 (women who are in general already married at that time). $Treat_g$ is a measure of the treatment: we use alternatively treatment different definitions, as presented in Section 3.4. We also add a set of individual controls X_i , which includes whether the education of the individual has some education and, her residence and is living in rural or urban areas in 2011 or not. To build the type of residence variable (rural or urban) in 2011, we use retrospective information on the kism (district) of residence collected in the survey, that we match with information coming from the census data. More precisely, we compute the average level of rurality by district using the 2006 census data, that are more representative and of higher size than the ELMPS data. We match this variable with the kism (district) of residence of the individual in 2011, that we built thanks to information on migration present in the ELMPS data: changes of kism are reported ².

For each set of results, we provide placebo estimations, comparing women who were between

2. There is one exception in the report of changes of kism. Enumerators were advised that “Moving within the same governorate from urban to urban or rural to rural does not count as a move.” We need therefore to do an approximation (not changing the kism of residence when the type of area is the same), but since the type of residence did not change, we think that the issue is limited. We present also results using the current place and type of residence at the time of the survey (2018) in Appendix: results are in line with our main results. In this alternative strategy, we do not have to do the earlier mentioned approximation but changes of residence following the Arab spring can bias the results (we discuss concerns relative to migration in Section 3.8).

26 and 30 at the time of the Revolution, to women who were between 36 and 40. The age at marriage of these women is not supposed to be affected by the Revolution, since they were already married. This is a way to test the parallel trend assumption.

We look at different outcomes: age at marriage and age at first birth. Though, classical OLS with these variables is biased since women in the treatment group are not all married. We therefore consider an alternative variable, the probability to be married before 23 years old. This variable is known at the time of the survey for every woman being 16 years old or older at the time of the Revolution (in 2011). Figure 3.5 plots the probability to be married before the age of 23 years old, according to the age at the beginning of the Egyptian Revolution. Women have a higher probability to be married before 23 years old in less treated areas; but the gap with highly treated areas drastically reduces for women who were younger than 23 years old at the time of the Revolution. This time, the results are not biased by the censorship of the variable, and thus offer evidence of the represent a first hint of a potential impact of the Egyptian Revolution on the age of entry into marriage for women.

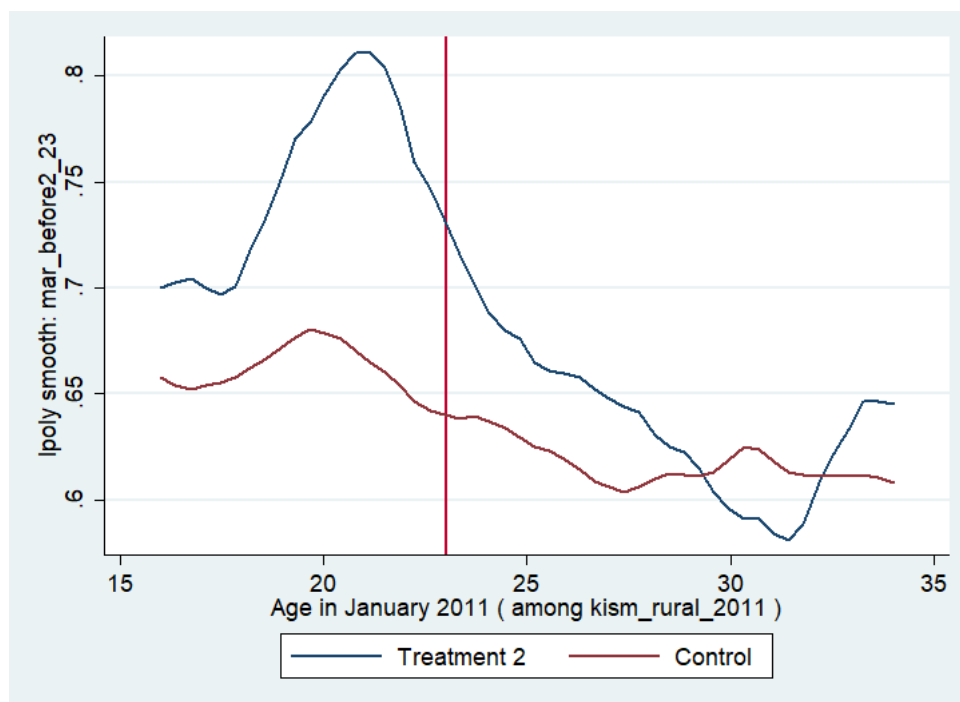


Figure 3.5 – Probability of marriage before 23 for rural women

Note: The figure presents the probability of marriage before 23 according to the age in January 2011 in rural areas.

Source: ELMPS 2018

3.5.3 Duration model

Since two of our main outcome variables are right-censored (age at marriage and age at first birth), we use a duration model to correct for this bias.

We use a discrete-time duration model to test whether being exposed to the Revolution accelerates the entry into marital and fertile life. It is common in the literature, to use a Cox proportional hazard model (Cox (1972); Lancaster (1990)). Here, we choose the discrete time duration model for two main reasons.

First, in models such as the Cox model, time is continuous and there should not be any simultaneous events. A duration, t_i , that leads to the studied event should be associated with one observation, i , if the clock for measuring the duration is sufficiently precise. Yet the Labor Market Panel Survey (as well as many other household surveys) collects time-discretized data. Therefore, many events can happen at the same time (women born the same year and the same month having their child at the same age, for instance). Second, and more importantly, the Cox model relies on the parallel hazard assumption. In the context of this paper, it would mean that the ratio of the risk of experiencing a marriage is constant between treated and untreated women at every moment within the studied duration. But here, we cannot assert that the risk of experiencing the first marriage in the studied duration is the same, whatever the geolocation (and therefore the intensity of the treatment) and the age of the women. The discrete-time model allows us to introduce time as a covariate, and therefore reduces this issue. We use the most commonly used function in this setting, the logistic regression:

$$\log\left(\frac{p_{t_{igc}}}{1 - p_{t_{igc}}}\right) = a_0 + a_k * t + a_{pk} * t^2 + \alpha_g + \beta * Post_i + \theta * Treat_g + \delta * TREAT_g * Post_i + \gamma X_i \quad (2)$$

where $p_{t_{igc}}$ is the probability of experimenting the event (marriage or first birth) and t is the number of years passed since the respondent's birth. For this analysis, the data is reshaped: there is one observation per year and per woman until she get married or, if she does not, until she is surveyed. We keep only observations for which the age of the woman is at least 11 years old, since there is almost no marriage before this age. As in the previous model, X_i includes whether the individual has an education or not; $TREAT$ is the binary exposure to treatment according to

the birth cohort; and α_g is the governorate of residence (in 2011) fixed effect. We present the odd-ratios (and not the coefficients) in the tables displaying the results.

As a robustness check, we present in appendix the results using the more traditional Cox Proportional Hazard model, according to which the hazard function satisfies:

$$h_{imct} = h_0(t) * \exp(\beta * Post_i + \theta * Treat_g + \delta * Treat_g * Post_i + \alpha_g + \eta X_i) \quad (3)$$

where h_{imct} represents the hazard function, determined by the same covariates we use in the OLS model and $h_0(t)$ is the baseline hazard. Variables with positive coefficients (the β values) are associated with increased hazard and decreased survival times, i.e., as the predictor increases the hazard of the event and the predicted survival duration decreases. The results (in hazard-ratios) using a Cox Model are very similar to the ones found with the discrete time duration model.

3.6 Regression results

3.6.1 Linear Probability Models

Age at marriage being right-censored, using this variable with an OLS model would bias the results. Age at marriage is only known for women who are already married, this constitutes a bias in the difference-in-differences setup. In the control group women are aged between (33-37 at the year of the survey) so almost all of them are married, however in the treatment group, women are younger (aged between 23-27 at the year of the survey) and not all of them are married yet. As women who are already married in the treatment group are those who got married at a young age, this would artificially lower the age of marriage and lead to overestimate the effect of revolution on age at marriage.³ We consider an alternative variable, the probability to be married before 23 years old. We choose this age threshold, because it is the threshold for which, at the time of the survey, in 2018, this variable is known for every respondent considered in the age interval of the specification (ages between 16-30 in January 2011). We could not look

3. Table 3.11 in the Appendix displays the estimates of the OLS estimation of the age at marriage. Columns (1), (3) and (5) present the results using the first definition of the treatment, meaning the share of people killed. The coefficient of interest is not significant no matter the sample considered (all women, women in rural areas or in urban areas). If we use the second definition of the treatment (dummy variable for the six governorates with the highest share of killed people), the coefficient is negative and significant, regardless of the sample definition. The results of the placebo are overall never significant. Note that these results are potentially biased.

at the probability to be married before 25 years because women who are 16 or 17 in 2011, at the time of the Revolution, are only 23 or 24 at the time of the survey, and we do not know yet whether they will be married before 25. Table 3.3 displays the results.⁴ Results with the first treatment definition, i.e., the share of persons killed at the level of the governorate are depicted in columns (1), (3) and (5). Results with the second definition of treatment, i.e., a dummy variable that takes the value 1 for the six governorates with the highest share of persons killed and 0 for remaining governorates, are depicted in columns (2), (4) and (6). The coefficient of interest for the first treatment definition is not statistically significant when considering the whole sample or for urban areas. It is positive and significant at the 5% level for rural areas⁵, meaning that the Arab Spring has accelerated the marriage for this specific sample. We find the same results when using the second definition of treatment: not statistically significant when using the entire sample and for urban areas, but positive and significant at the 5% level for rural areas sample (column(4)). The latter result suggests that the probability of being married before 23 increased by 12.8 percentage points in rural areas following the Arab Spring.

When considering the probability to have a child before 23 years old, Table 3.4 does not show a statistically significant effect at the usual thresholds, it is marginally significant at the 15% threshold. As these results only consider this specific threshold, we turn to a duration model for a more complete picture, in the next section.

4. We show results using logit model in tables 3.18 and 3.19 of the Appendix. We find that Logit results are consistent with LPM.

5. The rural characteristic is endogenous, since migration could for instance have followed the Arab Spring. We check therefore the migratory status of rural women using the census (see Section 3.7.4)). We find that migration of young women is relatively rare, and it is even more the case in rural areas.

Table 3.3 – LPM Model: Marriage before 23 years old

	All			Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	
Comparing people between 16-20 in January 2011 versus 26-30							
<i>Treatment 1</i> : Share of killed people	0.319** (0.140)		18.12*** (0.393)		-0.521*** (0.0310)		
People aged between 16 and 20 in January 2011 (vs 26-30)	0.0157 (0.0168)	0.00232 (0.0140)	-0.00833 (0.0252)	0.0196 (0.0171)	-0.0529+ (0.0345)	-0.110*** (0.0303)	
Aged 16-20*Share of killed people	-0.198 (0.177)		1.447** (0.718)		-0.0271 (0.126)		
<i>Treatment 2</i> : 6 more impacted governorates		0.0557* (0.0322)		0.0860** (0.0315)		-0.154*** (0.0129)	
Aged 16-20*6 more impacted governorates		0.00555 (0.0630)		0.128** (0.0621)		0.0747+ (0.0478)	
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	
N	4640	4640	3647	3647	993	993	
r2	0.07	0.07	0.03	0.03	0.04	0.04	
Comparing people between 26-30 in January 2011 versus 36-40							
<i>Treatment 1</i> : Share of killed people	0.310** (0.124)		8.458*** (0.874)		0.413*** (0.0926)		
People aged between 26 and 30 in January 2011 (vs 36-40)	0.0599** (0.0249)	0.0524** (0.0255)	0.0503 (0.0377)	0.0403 (0.0300)	0.0861** (0.0410)	0.0963** (0.0461)	
Aged 26-30*Share of killed people	-0.214* (0.115)		-0.245 (0.851)		-0.315* (0.158)		
<i>Treatment 2</i> : 6 more impacted governorates		0.0584* (0.0317)		0.0233 (0.0246)		0.102*** (0.0301)	
Aged 26-30*6 more impacted governorates		-0.0202 (0.0307)		0.0471 (0.0372)		-0.0759 (0.0528)	
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	
N	3334	3334	2555	2555	779	779	
r2	0.06	0.06	0.02	0.02	0.04	0.04	

Note: The dependent variable is a dummy taking the value 1 if the marriage happens before 23 years old. We control also for primary education. Panel A and panel B present the results of the double difference model. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Women.

Source: ELMPS Egypt 2018.

Table 3.4 – LPM Model: Child before 23 years old

	All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)
Comparing people between 16-20 in January 2011 versus 26-30						
<i>Treatment 1</i> : Share of killed people	0.732*** (0.158)		20.78*** (0.459)		-0.409*** (0.0562)	
People aged between 16 and 20 in January 2011 (vs 26-30)	-0.000190 (0.0182)	-0.0143 (0.0160)	-0.0143 (0.0273)	0.00437 (0.0193)	-0.0689+ (0.0440)	-0.137*** (0.0467)
Aged 16-20*Share of killed people	-0.171 (0.191)		1.044 (0.786)		0.0259 (0.194)	
<i>Treatment 2</i> : 6 more impacted governorates		0.161*** (0.0365)		0.188*** (0.0328)		-0.127*** (0.0164)
Aged 16-20*6 more impacted governorates		0.0143 (0.0613)		0.110+ (0.0636)		0.100+ (0.0611)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	4640	4640	3647	3647	993	993
r2	0.07	0.07	0.03	0.03	0.04	0.04
Comparing people between 26-30 in January 2011 versus 36-40						
<i>Treatment 1</i> : Share of killed people	0.594*** (0.152)		13.45*** (0.725)		0.186* (0.103)	
People aged between 26 and 30 in January 2011 (vs 36-40)	0.0804*** (0.0251)	0.0807*** (0.0242)	0.0773** (0.0331)	0.0774** (0.0271)	0.0312 (0.0480)	0.0774+ (0.0478)
Aged 26-30*Share of killed people	-0.267* (0.135)		0.143 (0.602)		-0.0765 (0.184)	
<i>Treatment 2</i> : 6 more impacted governorates		0.147*** (0.0345)		0.103*** (0.0218)		0.0823** (0.0357)
Aged 26-30*6 more impacted governorates		-0.0605** (0.0253)		0.0484* (0.0271)		-0.0811 (0.0582)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	3334	3334	2555	2555	779	779
r2	0.05	0.05	0.03	0.03	0.04	0.04

Note: The dependent variable is a dummy taking the value 1 if the woman had a child before 23 years old. We control also for primary education. Panel A and panel B present the results of the double difference model. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Women.

Source: ELMFPS Egypt 2018.

3.6.2 Duration Model

Table 3.5 displays the estimates of the duration model for rural women (we present results for all women in Appendix Table 3.13 and for urban women in Table 3.14 of the Appendix), whereby we estimate the determinants of the probability to be married at a given age. Considering the first definition of the treatment, the Arab Spring seems to have increased the likelihood of experiencing marriages in rural areas. The placebo estimations are not significant, which gives confidence in the validity of the results. When considering the second definition of the treatment, results are in line with the results obtained with the first treatment definition. Furthermore, with this second treatment, the placebo is also not significant. We can thus state that in treated rural areas, the Arab Spring has increased the probability to experience a marriage for women. Results for the probability to obtain a first child (columns (3) and (4)) are coherent with the probability of entering into marriage (columns (1) and (2)). For instance, to live in one of the 6 more impacted governorates, in rural areas, for the treated cohort, multiplies the odd of experiencing the event by 1.6055 relatively to the older cohort, meaning it increases the probability to experience the event. The Arab Spring has accelerated the occurrence of giving birth, for women in rural areas. Results with the Cox model are displayed in Appendix in Table 3.15, they are very in line with the discrete time duration model.

Table 3.5 – Marital outcome at the governorate level, for rural women - Discrete time duration model

	Age at marriage		Age at first child	
	(1)	(2)	(3)	(4)
		Comparing people between 16-20 in January 2011 versus 26-30		
<i>Treatment 1</i> : Share of killed people	1.86137e+12*** (3.17128e+12)		3.64043e+20*** (8.17905e+20)	
Aged 16-20*Share of killed people	211.9** (591.0)		33.08** (53.74)	
<i>Treatment 2</i> : more impacted governorates		1.460*** (0.178)		1.724*** (0.105)
Aged 16-20*6 more impacted governorates		1.605** (0.396)		1.352*** (0.153)
governorate fixed effects	Yes	Yes	Yes	Yes
N	38146	38146	43914	43914
r2_p	0.14	0.14	0.15	0.15
		Comparing people between 26-30 in January 2011 versus 36-40		
<i>Treatment 1</i> : Share of killed people	1292309.5*** (2752429.3)		1.05354e+17*** (2.93027e+17)	
Aged 26-30*Share of killed people	0.0588+ (0.109)		0.282 (0.489)	
<i>Treatment 2</i> : more impacted governorates		1.273*** (0.0825)		1.581*** (0.107)
Aged 26-30*6 more impacted governorates		0.930 (0.0893)		1.022 (0.0805)
governorate fixed effects	Yes	Yes	Yes	Yes
N	28725	28725	34546	34546
r2_p	0.07	0.07	0.13	0.13

Note: The dependent variable is age at marriage for columns (1) and (2), and age at first child for columns (3) and (4). We control also for primary education. Odd-ratios are displayed. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Rural Women.

Source: ELMPS Egypt 2018.

3.7 Robustness checks and Discussion

In this section, we conduct several robustness tests by varying the definition of the treatment and the scale of the analysis.

3.7.1 Alternative definition of the treatment

We use two alternative measures of the treatment. The first one, *Treatment 3*, is a binary variable taking the value 1 if the share of people killed is larger than the median, and the second one, *Treatment 4*, is the same measure, except that people injured and arrested are also considered. These measures have the advantage of being less correlated with the urbanization, as opposed to *Treatment 2*, which was especially concentrating big cities. Results for rural women regarding the entry into marriage (Probability to marry before age 23) hold when using these two alternative treatment measures, as shown in Table 3.6.

When considering the entry into fertility, the signs of coefficients are consistent, but results are no longer significant with *Treatment 3*. Considering *Treatment 4*, results are nevertheless marginally significant at 15% level for the probability to have a child before 23 years old.

3.7.2 At the district level

In order to strengthen our results, we perform the analysis at a finer geographic level, the district. We geo-coded every event of the "Wikithawra" database, and construct two treatment variables at this level: the share of killed people in the district, and a dummy variable. This dummy variable takes the value 1 for districts, in which the share of people killed is over the median, and 0 otherwise. Results are very consistent with the results at the governorate level (Table 3.7). Rural women residing in "high" treated districts marry earlier and have children earlier as well. We still prefer to use the governorate analysis as the main one, in order to limit the issues linked to migration (since the governorate analysis mitigates some concerns due to governorate-internal migration.⁶)

6. Within governorate migration from rural to rural or urban to urban between districts is not considered as migration in the database.

3.7.3 Alternative specification: discretized duration model

We also tested an alternative specification of the discretized duration model. In this model, there is one observation by woman and by year starting from age 11 (since there is almost no marriage and birth before this age). The outcome (marriage or birth) takes the value 0 if the woman is not married (or has not given birth) at the age considered. It takes the value 1 at the age where the woman married (or gave birth). The subsequent year, the woman leaves the sample. For women who are still not married, the outcome takes the value 0 for each age. Results are presented in Table 3.8. They are in line with our main results.

3.7.4 Threats to the Identification Strategy

The use of the difference-in-differences in this setting raises several issues.

Urbanization The protests mostly occurred in cities. One threat to our identification strategy would be that we are capturing the effect of urbanization, instead of the one of protests intensity. Indeed, the 6 most impacted governorates are big urban cities. This threat is nevertheless mitigated by the fact that our results concern above all rural areas, where violence intensity is low. Our results hold also when we use a more continuous variable of the treatment, and when we are at the district level (these are 2 different things). Moreover, we are using governorates fixed effects that control for the level of urbanization.

Impact of other changes The protests follow and create economic difficulties. Another potential threat to our identification strategy could be that the results are not driven by the protests, but by pre-existing socio-economic difficulties. We do not think that this is a major issue since the protests are not correlated with economic difficulties, as shown in Table 3.1. This table allows to assess the correlation of the protests with the pre-existing economic situation according to highly treated and less treated areas. The difference in unemployment rate between treated and less treated before the Revolution (2006) is not significantly different from zero which gives confidence in the fact that our results are not entirely drawn by pre-existing socio-economic trends.

Migration Thanks to the retrospective information on migration, the individual's place of residence at the time of the Egyptian Revolution is known and used to determine the intensity of the treatment. Meaning that even when a woman migrated after the Revolution, we assign the level of violence in her place of residence at the time of the Revolution, because she was clearly affected or exposed to it. Migration implied by the Revolution should therefore not be a threat to our strategy. Nevertheless, to build the information on the place of residence in 2011, we have to do a small approximation: since enumerators were advised that "Moving within the same governorate from urban to urban or rural to rural does not count as a move". Thus, we have to consider that there was no move when individuals moved within the same governorate, from rural to rural, or urban to urban. This is also the reason why we present our results at the governorate level as our main or 'preferred' specification and not the results at the district level: even when the last one is at a more disaggregated level it is more prone to migration bias. The use of governorate level as the location of the violence should be immune to within-governorate migration movements. To convince that this approximation is not a strong issue, we redo the analysis, using the place of residence at the time of survey (2018) as our indicator of residence. As can be seen in Table 3.16 of the Appendix, results are in line with the main ones, suggesting that migration does not strongly impact the results. When using as intensity of the treatment, the place of residence in 2018, migration could be a problem because people may not be in the same place as in 2011: there could be a measurement error, or even a bias if migration is not orthogonal to the treatment. Migration however is in fact very limited. Only 6.44% of women who were between 16 and 20 years old at the time of the Revolution have moved from districts where they were born. There is no significant difference according to whether they lived in highly treated governorates or less treated. Also, we observe that 95.43% of women in our sample live in 2018 in the same district in which they lived in 2011. The latter share is even higher (96.23%) when we consider women in the age category of interest (16-30) and in rural areas. Even if migration is not a threat to our identification strategy, it could be a channel explaining the results on marriages: it would be the case if for instance the Revolution fostered women to migrate in places where they marry sooner. We restrict the sample to only women who did not move between 2011 and 2018, results are presented in Table 3.17 in Appendix. The table shows consistent results to those in (Table 3.5), where the place of residence in 2011 is used. Overall, migration does not seem to impact our results.

What if marriage and fertility at younger age had nothing to do with the Revolution?

Even though the total fertility rate in Egypt was increasing before the Revolution, we argue that the Egyptian revolution played a role in increasing this trend further, by impacting the age at entry into marriage, a pre-requisite to have a child.

[Krafft \(2020\)](#) observes a decrease in the public sector jobs that coincides with the increase in fertility. However, the decrease of opportunities of employment in the public sector impacts more the third and fourth births, and not the first births, which is our outcome of interest in this paper. This suggests that this mechanism cannot explain the decrease in the age at marriage and age at first birth, that we document in this paper.

[Ambrosetti et al. \(2019\)](#) suggests that the trend reversal is due to a change in ideals related to family size preferences. When comparing cohorts of women in 2008 and 2015 between the age of 15 and 24, [Ambrosetti et al. \(2019\)](#) shows that there is an increase in the preference for more children and this preference is observed no matter women's level of education or place of residence. The Arab Spring could have affected preferences, but we see clearly a difference of impact between rural and urban areas, suggesting that the channel of preferences change cannot explain our results.

Overall, even if fertility had already been increasing before the Revolution, this does not run counter to our results. The revolution seems to have reinforced the trend shift. We argue that revolution played a role in reducing childbearing age in Egypt.

Table 3.6 – Marital outcome at the governorate level, rural women –Treatment 3 and Treatment 4- discrete time duration model

	Marriage before 23		Child before 23		Age at marriage		Age at first child	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Comparing people between 16-20 in January 2011 versus 26-30								
<i>Treatment 3:</i> Share killed p. above the median	-0.0811*** (0.0165)		-0.0161 (0.0218)		0.714*** (0.0493)		0.792*** (0.0407)	
Aged 16-20*Share killed p. above the median	0.0656* (0.0343)		0.0191 (0.0447)		1.240* (0.159)		1.126 (0.138)	
<i>Treatment 4:</i> Share killed/arrested p. above the median		0.150*** (0.0171)		0.175*** (0.0211)		1.224** (0.103)		1.531*** (0.104)
Aged 16-20*Share killed/arrested p. above the median		0.0749* (0.0378)		0.0730+ (0.0451)		1.228 (0.209)		1.139 (0.170)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3647	3647	3647	3647	38146	38146	43914	43914
r2_p					0.14	0.14	0.15	0.15
Comparing people between 26-30 in January 2011 versus 36-40								
<i>Treatment 3:</i> Share killed p. above the median	-0.0174 (0.0297)		-0.0251 (0.0298)		0.826** (0.0627)		0.852* (0.0772)	
Aged 26-30*Share killed p. above the median	-0.0560 (0.0559)		0.00323 (0.0496)		0.829+ (0.0984)		0.911 (0.130)	
<i>Treatment 4:</i> Share killed/arrested p. above the median		0.133*** (0.0398)		0.169*** (0.0315)		1.298*** (0.0998)		1.613*** (0.152)
Aged 26-30*Share killed/arrested p. above the median		-0.0842 (0.0625)		-0.0605 (0.0480)		0.777* (0.106)		0.845 (0.123)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2555	2555	2555	2555	28725	28725	34546	34546
r2_p					0.07	0.07	0.13	0.13

Note: The dependent variable is age at marriage for columns (1) and (2), marriage before 23 years old for columns (3) and (4), age at first child (5) and (6) and child before 23 years old for columns (7) and (8). We control also for primary education. Panel A and panel B present the results of the double difference model. For the duration model, odd-ratios are displayed. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Rural Women.

Source: ELMPS Egypt 2018.

Table 3.7 – Marital outcome at the district level - discrete time duration model, for rural women

	Age at marriage		Marriage before 23		Age at first child		Child before 23	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Comparing people between 16-20 in January 2011 versus 26-30								
Treatment 1: Share of killed people	0.473 (0.629)		-0.385 (0.516)		0.637 (0.879)		0.317 (0.711)	
Aged 16-20*Share of killed people	51.25** (87.34)		1.322*** (0.421)		62.51** (120.9)		0.780 (0.907)	
Treatment 3: Share killed above the median		0.975 (0.0773)		0.0108 (0.0352)		0.994 (0.0787)		0.0337 (0.0422)
Aged 16-20*Share killed p. above the median		1.381*** (0.134)		0.0755** (0.0359)		1.452*** (0.144)		0.0754* (0.0407)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	38146	38146	3647	3647	43914	43914	3647	3647
r2_p	0.14	0.14			0.15	0.16		
Comparing people between 26-30 in January 2011 versus 36-40								
Treatment 1: Share of killed people	0.741 (1.282)		-0.265 (0.805)		1.160 (1.802)		0.314 (0.700)	
Aged 26-30*Share of killed people	0.640 (0.881)		0.140 (0.325)		0.580 (0.487)		0.291 (0.587)	
Treatment 3: Share killed above the median		1.247* (0.152)		0.000866 (0.0368)		1.053 (0.121)		0.00559 (0.0302)
Aged 26-30*Share killed p. above the median		0.766** (0.0815)		0.0118 (0.0359)		0.934 (0.104)		0.0309 (0.0446)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	28725	28725	2555	2555	34546	34546	2555	2555
r2_p	0.07	0.07			0.13	0.13		

Note: The dependent variable is age at marriage for columns (1) and (2), marriage before 23 years old for columns (3) and (4), age at first child (5) and (6) and child before 23 years old for columns (7) and (8). We control also for primary education. For the duration model, odd-ratios are displayed. Panel A and panel B present the results of the double difference model. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Rural Women.

Source: ELMPS Egypt 2018.

Table 3.8 – Alternative specification

	Age at marriage		Age at first birth	
	(1)	(2)	(3)	(4)
	Comparing people between 16-20 in January 2011 versus 26-30			
<i>Treatment 1</i> : Share of killed people	1.105*** (0.128)		1.658*** (0.0975)	
Aged 16-20*Share of killed people	0.238+ (0.146)		0.137* (0.0712)	
<i>Treatment 2</i> : 6 more impacted governorates		0.0176** (0.00646)		0.0222*** (0.00261)
Aged 16-20*6 more impacted governorates		0.0204+ (0.0136)		0.0111** (0.00484)
governorate fixed effects	Yes	Yes	Yes	Yes
N	38135	38135	43881	43881
r ² _p				

Note: The dependent variable is age at marriage and age at first birth. We control also for primary education. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Rural Women.

Source: ELMPS Egypt 2018.

3.8 Channels and Interpretation

3.8.1 By which period of the Revolution are the results driven?

As explained in Section 3.3, the Egyptian Revolution can be roughly decomposed in 4 time periods: the first 18 days leading to Mubarak resignation, the Supreme Council of Armed Force rule until June 2012, then the governance of President Morsi and finally the period of protests against him and oppression of his supporters until Sissi's arrival on January 2014.

We put together the two first periods and the two last periods of the revolution, since the election of President Morsi from Muslim brotherhood party is a turning point in the Revolution. We look at the impact of the number of people killed, for each aggregated period separately. Results on age at marriage and age at first birth are presented in (Table 3.9)⁷. We find that results are statistically significant for both periods, and hence we can not conclude on whether the results are driven by the first period of revolution against "Hosni Mubarak regime" or the second period characterized by the arrival of Muslim brotherhood adept "president Mohamed Morsi".

3.8.2 Economic shock versus social norms and religiosity

When we show that the Arab Spring has reduced the age at marriage in rural areas, we wonder whether the effect could channel through an economic variable, meaning economic difficulties proper to this uncertain period, or to another channel, pertaining to social norms and religiosity highlighted in the first or second period of the Egyptian Revolution. Two arguments speak against the economic channel: (i) according to World Bank data, the decrease in the GDP occurs in 2017 and not earlier (ii) results are not driven by the areas that have been the most impacted economically. Results however seem to be driven by areas where people have a more conservative vote.

To further investigate to what extent results are driven by the level of conservatism, we use data from the Arab Barometer (2010-2011). We compute indicators of conservatism by governorate, such as the share of people who think women are worse than men in politics and

7. In this table, coefficients and not odd-ratios are presented, because the magnitude of the odd-ratio for the variable "share of killed people" would be too large. For this reason, only the sign of the coefficients may be commented.

the share of people who think Islamic laws should be applied in marriage and divorce. Results in (Table 3.10)⁸ suggest that the decrease in the age of marriage in rural areas seem to be driven by conservatism. The revolution contributed in decreasing age of marriage in areas that are conservative. It could have been driven in the first period, by a reaction to the ideas carried by the Egyptian Revolution and driven in the second period by the presence of Morsi as President and by the revolt against his destitution.

8. In this table, coefficients and not odd-ratios are presented, because the magnitude of the odd-ratio for the variable "share of killed people" would be too large. For this reason, only the sign of the coefficients may be commented.

Table 3.9 – Duration model: Occurrence of the event. Age at marriage and first birth according to Revolution period in rural areas - discrete time duration model

	Age at marriage		Age at first child	
	(1)	(2)	(5)	(6)
	Comparing people between 16-20 in January 2011 versus 26-30			
Share of killed people until June 2012	1901.4*** (87.32)		3021.5*** (124.6)	
People aged between 16 and 20 in January 2011 (vs 26-30)	-0.0156 (0.0699)	0.0246 (0.0641)	-0.106+ (0.0664)	-0.0425 (0.0700)
Aged 16-20*Share of killed people until June 2012	19.15** (9.975)		16.24*** (6.114)	
Share of killed people after June 2012		28.35*** (1.920)		48.08*** (2.447)
Aged 16-20*Share of killed people after June 2012		6.274* (3.437)		3.710* (2.186)
governorate fixed effects	Yes	Yes	Yes	Yes
N	38146	38146	43914	43914
r ² _p	0.14	0.14	0.15	0.15

Note: The dependent variable is age at marriage for columns (1) and (2). The dependent variable is age at first child for columns (3) and (4). The model is estimated with a discrete-time duration model (coefficients are displayed). We control also for primary education. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Rural Women.

Source: ELMPS Egypt 2018.

Table 3.10 – Impact of the revolution on the probability of marriage according to the degree of conservatism in rural areas (discrete time duration model)

	Women worse in politics belief		Not the case		Islamic rules should apply in marriage		Not the case	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
	Comparing people between 16-20 in January 2011 versus 26-30							
<i>Treatment 1</i> : Share of killed people	42.75*** (2.334)		7627.8*** (237.2)		40.63*** (0.968)			
Aged 16-20*Share of killed people	2.183 (1.782)		-0.488 (1.913)		4.304** (1.987)			
<i>Treatment 2</i> : 6 more impacted governorates		0.323*** (0.0486)		3.834*** (0.175)		0.277*** (0.0429)	2.896*** (0.188)	
Aged 16-20*6 more impacted governorates		0.302*** (0.0820)		-0.204** (0.0881)		0.326*** (0.0854)	-0.205** (0.100)	
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
N	20078	20078	17116	17116	18170	18170	19024	
r2_p	0.14	0.14	0.13	0.13	0.14	0.14	0.13	

Note: The dependent variable is age at marriage, and the model estimated is a discrete time duration model (coefficients are displayed). Results of the estimation of the model with treatment 1 for the sample of people who think not that Islamic rules should apply in marriage are not displayed because the model does not converge. We control also for primary education. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Rural Women.
Source: ELMPS Egypt 2018.

3.9 Conclusion

This paper studies the impact of the Arab spring on women's age of entry into marriage and fertility in the case of Egypt. We find that women residing in rural areas, who were aged between 16 and 20 at the time of the Egyptian Revolution, marry earlier than the previous cohorts, and have children earlier as well. These results moderate previous evidence on the increase in women's empowerment following the Arab Spring.

Our results contribute to the ambiguous context of existing literature on the impact of the Egyptian revolution on women's status and participation in the labor market. The fact that women marry earlier in rural areas corresponds with the negative impact of 2011 Revolution on labor market status shown by [Hendy \(2015\)](#), since women marrying earlier are less likely to participate in the labor market. Our results also highlight the heterogeneity of the impact of the Revolution on entry into marriage and motherhood between urban and rural areas.

Appendices

3.A Appendix

Table 3.11 – LPM Model: Age at marriage

	All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)
Comparing people between 16-20 in January 2011 versus 26-30						
<i>Treatment 1</i> : Share of killed people	1.065 (1.413)		-143.2*** (3.866)		7.945*** (0.293)	
People aged between 16 and 20 in January 2011 (vs 26-30)	-1.054*** (0.144)	-0.950*** (0.116)	-0.931*** (0.219)	-1.053*** (0.133)	-1.118** (0.397)	-0.390+ (0.242)
Aged 16-20*Share of killed people	-2.658 (1.896)		-8.254 (6.643)		-1.833 (1.673)	
<i>Treatment 2</i> : 6 more impacted governorates		0.383 (0.348)		-0.183 (0.199)		2.089*** (0.0613)
Aged 16-20*6 more impacted governorates		-0.992*** (0.347)		-1.151*** (0.380)		-1.377*** (0.374)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	4102	4102	3293	3293	809	809
r ²	0.11	0.11	0.07	0.07	0.06	0.07
Comparing people between 26-30 in January 2011 versus 36-40						
<i>Treatment 1</i> : Share of killed people	-4.314** (1.676)		-173.8*** (9.357)		-4.211*** (1.284)	
People aged between 26 and 30 in January 2011 (vs 36-40)	-0.720** (0.304)	-0.699** (0.292)	-0.815* (0.436)	-0.637* (0.318)	-0.435 (0.444)	-0.645 (0.588)
Aged 26-30*Share of killed people	2.856+ (1.754)		7.013 (7.361)		1.483 (2.012)	
<i>Treatment 2</i> : 6 more impacted governorates		-0.999** (0.374)		-0.122 (0.287)		-1.187*** (0.360)
Aged 26-30*6 more impacted governorates		0.555* (0.305)		0.0970 (0.392)		0.582 (0.601)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	3131	3131	2418	2418	713	713
r ²	0.07	0.07	0.05	0.05	0.04	0.04

Note: The dependent variable is age at marriage. We control also for primary education. Panel A and panel B present the results of the double difference model. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Women.

Source: ELMPS Egypt 2018.

Table 3.12 – LPM Model: Age at first child

	All			Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	
Comparing people between 16-20 in January 2011 versus 26-30							
<i>Treatment 1</i> : Share of killed people	-3.732*** (1.265)		-241.7*** (2.216)		6.957*** (0.308)		
People aged between 16 and 20 in January 2011 (vs 26-30)	-1.356*** (0.143)	-1.318*** (0.116)	-1.195*** (0.206)	-1.382*** (0.138)	-1.637*** (0.315)	-1.007*** (0.236)	
Aged 16-20*Share of killed people	-4.901** (2.107)		-10.40** (5.085)		-3.065* (1.552)		
<i>Treatment 2</i> : 6 more impacted governorates		-0.971*** (0.302)		-0.855*** (0.172)		1.777*** (0.0392)	
Aged 16-20*6 more impacted governorates		-1.222*** (0.248)		-1.055*** (0.355)		-1.466*** (0.302)	
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	
N	3705	3705	2978	2978	727	727	
r ²	0.13	0.14	0.09	0.09	0.10	0.11	
Comparing people between 26-30 in January 2011 versus 36-40							
<i>Treatment 1</i> : Share of killed people	-5.890** (2.168)		-160.2*** (7.872)		0.480 (1.549)		
People aged between 26 and 30 in January 2011 (vs 36-40)	-0.931*** (0.319)	-0.911*** (0.306)	-0.992** (0.413)	-0.870** (0.304)	-0.353 (0.547)	-0.315 (0.705)	
Aged 26-30*Share of killed people	2.308 (1.761)		4.869 (7.182)		-0.363 (2.391)		
<i>Treatment 2</i> : 6 more impacted governorates		-1.408*** (0.480)		-1.285*** (0.243)		0.139 (0.453)	
Aged 26-30*6 more impacted governorates		0.430 (0.354)		0.0806 (0.339)		-0.122 (0.731)	
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	
N	2955	2955	2282	2282	673	673	
r ²	0.08	0.08	0.05	0.05	0.07	0.07	

Note: The dependent variable is age at first child. We control also for primary education. Panel A and panel B present the results of the double difference model. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Women.

Source: ELMPS Egypt 2018.

Table 3.13 – Marital outcome at the governorate level, for ALL women - Discrete time Duration model

	Age at marriage		Age at first child	
	(1)	(2)	(3)	(4)
Comparing people between 16-20 in January 2011 versus 26-30				
<i>Treatment 1</i> : Share of killed people	1.587 (0.634)		6.195*** (3.026)	
Aged 16-20*Share of killed people	0.528* (0.202)		0.529+ (0.217)	
<i>Treatment 2</i> : more impacted governorates		1.079 (0.0882)		1.546*** (0.173)
Aged 16-20*6 more impacted governorates		0.936 (0.156)		0.920 (0.126)
governorate fixed effects	Yes	Yes	Yes	Yes
N	50543	50543	57641	57641
r2_p	0.14	0.14	0.15	0.15
Comparing people between 26-30 in January 2011 versus 36-40				
<i>Treatment 1</i> : Share of killed people	1.527 (0.461)		5.283*** (1.950)	
Aged 26-30*Share of killed people	0.216*** (0.0761)		0.475** (0.157)	
<i>Treatment 2</i> : more impacted governorates		1.039 (0.0591)		1.480*** (0.119)
Aged 26-30*6 more impacted governorates		0.760*** (0.0608)		0.874+ (0.0835)
governorate fixed effects	Yes	Yes	Yes	Yes
N	39206	39206	46617	46617
r2_p	0.04	0.04	0.13	0.13

Note: The dependent variable is age at marriage for columns (1) and (2), and age at first child for columns (3) and (4). We control also for primary education. Odd-ratios are displayed. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: All Women.

Source: ELMPS Egypt 2018.

Table 3.14 – Marital outcome at the governorate level, for URBAN women - Discrete time Duration Model

	Age at marriage		Age at first child	
	(1)	(2)	(3)	(4)
Comparing people between 16-20 in January 2011 versus 26-30				
<i>Treatment 1</i> : Share of killed people	0.0746*** (0.00683)		0.597*** (0.0494)	
Aged 16-20*Share of killed people	1.776** (0.531)		1.529 (0.650)	
<i>Treatment 2</i> : more impacted governorates		0.515*** (0.0203)		0.881*** (0.0195)
Aged 16-20*6 more impacted governorates		1.190+ (0.128)		1.123 (0.151)
governorate fixed effects	Yes	Yes	Yes	Yes
N	12397	12397	13727	13727
r2_p	0.15	0.15	0.16	0.16
Comparing people between 26-30 in January 2011 versus 36-40				
<i>Treatment 1</i> : Share of killed people	1.184 (0.237)		2.157*** (0.470)	
Aged 26-30*Share of killed people	0.164*** (0.0748)		0.456* (0.208)	
<i>Treatment 2</i> : more impacted governorates		1.032 (0.0600)		1.279*** (0.0906)
Aged 26-30*6 more impacted governorates		0.610*** (0.0872)		0.730** (0.0963)
governorate fixed effects	Yes	Yes	Yes	Yes
N	10481	10481	12071	12071
r2_p	0.12	0.12	0.14	0.14

Note: The dependent variable is age at marriage for columns (1) and (2), and age at first child for columns (3) and (4). We control also for primary education. Odd-ratios are displayed. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Urban Women.

Source: ELMPS Egypt 2018.

Table 3.15 – Marital outcome at the governorate level, for rural women - Cox duration model

	Age at marriage		Age at first child	
	(1)	(2)	(3)	(4)
	Comparing people between 16-20 in January 2011 versus 26-30			
<i>Treatment 1</i> : Share of killed people	7.03247e+10*** (9.66827e+10)		7.60806e+17*** (1.40982e+18)	
Aged 16-20*Share of killed people	85.05** (188.0)		22.34** (29.83)	
<i>Treatment 2</i> : more impacted governorates		1.334*** (0.134)		1.562*** (0.0788)
Aged 16-20*6 more impacted governorates		1.474** (0.284)		1.305*** (0.120)
governorate fixed effects	Yes	Yes	Yes	Yes
N	3647	3647	3647	3647
r2_p	0.00	0.00	0.00	0.00
	Comparing people between 26-30 in January 2011 versus 36-40			
<i>Treatment 1</i> : Share of killed people	288840110.0*** (580121168.6)		4.70361e+14*** (1.09894e+15)	
Aged 26-30*Share of killed people	0.0919+ (0.149)		0.271 (0.397)	
<i>Treatment 2</i> : more impacted governorates		1.164*** (0.0627)		1.455*** (0.0887)
Aged 26-30*6 more impacted governorates		0.955 (0.0675)		1.002 (0.0731)
governorate fixed effects	Yes	Yes	Yes	Yes
N	2555	2555	2555	2555
r2_p	0.00	0.00	0.00	0.00

Note: The dependent variable is age at marriage for columns (1) and (2), and age at first child for columns (3) and (4). We control also for primary education. Hazard-ratios are displayed. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Rural Women.

Source: ELMPS Egypt 2018.

Table 3.16 – Marital outcome at the governorate level, for rural women - Current place of residence, discrete time duration model

	Age at marriage		Age at first child	
	(1)	(2)	(3)	(4)
	Comparing people between 16-20 in January 2011 versus 26-30			
<i>Treatment 1</i> : Share of killed people	4.31980e+09*** (2.11501e+10)		4.60588e+17*** (3.49209e+18)	
Aged 16-20*Share of killed people	1429.8*** (3654.6)		43.25*** (74.52)	
<i>Treatment 2</i> : more impacted governorates		1.415*** (0.149)		1.737*** (0.101)
Aged 16-20*6 more impacted governorates		1.942*** (0.390)		1.435*** (0.141)
governorate fixed effects	Yes	Yes	Yes	Yes
N	30137	30137	35075	35075
r2_p	0.14	0.14	0.16	0.16
	Comparing people between 26-30 in January 2011 versus 36-40			
<i>Treatment 1</i> : Share of killed people	44827.0 (521425.2)		2.87335e+13** (4.27557e+14)	
Aged 26-30*Share of killed people	0.0242** (0.0466)		0.278 (0.487)	
<i>Treatment 2</i> : more impacted governorates		1.467*** (0.173)		1.698*** (0.168)
Aged 26-30*6 more impacted governorates		0.813 (0.131)		0.954 (0.106)
governorate fixed effects	Yes	Yes	Yes	Yes
N	23064	23064	27994	27994
r2_p	0.10	0.10	0.13	0.13

Note: The dependent variable is age at marriage for columns (1) and (2), and age at first child for columns (3) and (4). We control also for primary education. Odd-ratios are displayed. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Rural Women.

Source: ELMPS Egypt 2018.

Table 3.17 – Outcomes at the governorate level for non migrant rural women between 2011-2018, discrete time duration model

	Age at marriage		Age at first child	
	(1)	(2)	(3)	(4)
		Comparing people between 16-20 in January 2011 versus 26-30		
<i>Treatment 2: more impacted governorates</i>	1.431*** (0.181)		1.707*** (0.121)	
Aged 16-20*6 more impacted governorates	1.644** (0.418)		1.359** (0.183)	
<i>Treatment 1: Share of killed people</i>		1.47356e+12*** (2.62156e+12)		1.87266e+20*** (4.27594e+20)
Aged 16-20*Share of killed people		258.9** (755.2)		37.04** (68.57)
governorate fixed effects	Yes	Yes	Yes	Yes
N	36506	36506	42074	42074
r ² _p	0.13	0.13	0.15	0.15
		Comparing people between 26-30 in January 2011 versus 36-40		
<i>Treatment 2: more impacted governorates</i>			1.651*** (0.133)	
Aged 26-30*6 more impacted governorates			0.980 (0.0959)	
<i>Treatment 1: Share of killed people</i>				5.14912e+16*** (1.44884e+17)
Aged 26-30*Share of killed people				0.399 (0.698)
governorate fixed effects			Yes	Yes
N			33268	33268
r ² _p			0.13	0.13

Note: The dependent variable is age at marriage for columns (1) and (2), and age at first child for columns (3) and (4). We control also for primary education. Results for treatment 2 and the placebo are not displayed because the model does not converge. Odd-ratios are displayed. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Rural Women.

Source: ELMPS Egypt 2018.

Table 3.18 – Logit Model: Marriage before 23 years old

	All			Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	
Comparing people between 16-20 in January 2011 versus 26-30							
<i>Treatment 1</i> : Share of killed people	1.445** (0.623)		73.80*** (2.492)		-2.105*** (0.127)		
People aged between 16 and 20 in January 2011 (vs 26-30)	0.0687 (0.0758)	0.00921 (0.0632)	-0.0624 (0.134)	0.0906 (0.0788)	-0.223+ (0.145)	-0.459*** (0.126)	
Aged 16-20*Share of killed people	-0.851 (0.742)		8.211+ (5.041)		-0.116 (0.519)		
<i>Treatment 2</i> : more impacted governorates		0.259* (0.142)		0.448** (0.192)		-0.624*** (0.0525)	
Aged 16-20*6 more impacted governorates		0.0244 (0.272)		0.784+ (0.493)		0.309+ (0.198)	
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	
N	4640	4640	3647	3647	993	993	
r ²							
Comparing people between 26-30 in January 2011 versus 36-40							
<i>Treatment 1</i> : Share of killed people	1.364*** (0.527)		35.05*** (3.966)		1.752*** (0.394)		
People aged between 26 and 30 in January 2011 (vs 36-40)	0.264** (0.108)	0.232** (0.111)	0.223 (0.166)	0.182 (0.133)	0.364** (0.174)	0.400** (0.190)	
Aged 26-30*Share of killed people	-0.936** (0.493)		-0.994 (3.786)		-1.323*** (0.655)		
<i>Treatment 2</i> : more impacted governorates		0.261** (0.134)		0.0956 (0.111)		0.431*** (0.129)	
Aged 26-30*6 more impacted governorates		-0.0949 (0.133)		0.212 (0.169)		-0.313+ (0.218)	
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	
N	3334	3334	2555	2555	777	777	
r ²							

Note: The dependent variable is a dummy taking the value 1 if the marriage happens before 23 years old. We control also for primary education. Panel A and panel B present the results of the double difference model. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Women.

Source: ELMPS Egypt 2018.

Table 3.19 – Logit Model: Child before 23 years old

	All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)
Comparing people between 16-20 in January 2011 versus 26-30						
<i>Treatment 1</i> : Share of killed people	3.290*** (0.780)		93.95*** (2.230)		-1.835*** (0.248)	
People aged between 16 and 20 in January 2011 (vs 26-30)	0.00481 (0.0765)	-0.0580 (0.0661)	-0.0643 (0.116)	0.0186 (0.0793)	-0.331+ (0.210)	-0.631*** (0.226)
Aged 16-20*Share of killed people	-0.892 (0.969)		4.710 (3.761)		0.0159 (0.921)	
<i>Treatment 2</i> : more impacted governorates		0.711*** (0.182)		0.785*** (0.151)		-0.570*** (0.0771)
Aged 16-20*6 more impacted governorates		0.0569 (0.289)		0.517+ (0.323)		0.437+ (0.302)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	4640	4640	3647	3647	984	984
r ²						
Comparing people between 26-30 in January 2011 versus 36-40						
<i>Treatment 1</i> : Share of killed people	2.606*** (0.713)		58.18*** (3.172)		0.919** (0.480)	
People aged between 26 and 30 in January 2011 (vs 36-40)	0.338*** (0.104)	0.336*** (0.0999)	0.315*** (0.135)	0.318*** (0.111)	0.148 (0.222)	0.345* (0.210)
Aged 26-30*Share of killed people	-1.082* (0.581)		0.672 (2.439)		-0.350 (0.837)	
<i>Treatment 2</i> : more impacted governorates		0.645*** (0.164)		0.416*** (0.0888)		0.395*** (0.170)
Aged 26-30*6 more impacted governorates		-0.241** (0.108)		0.203* (0.111)		-0.362 (0.265)
governorate fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	3334	3334	2555	2555	762	762
r ²						

Note: The dependent variable is a dummy taking the value 1 if the woman had a child before 23 years old. We control also for primary education. Panel A and panel B present the results of the double difference model. Significance levels are denoted as follows: + p<0.15, * p<0.10, ** p<0.05, *** p<0.01. Sample: Women.

Source: ELMPS Egypt 2018.

General Conclusion

What have we learned?

The purpose of this thesis was to empirically study the impact of (i) university openings on education and labor market outcomes, (ii) political outbreaks on women's marital outcomes. The analyses conducted rely on extensive survey data combined with well-established causal methods, notably difference-in-differences estimation technique. Specifically, I assess the consequences of specific events, which are illustrated by a reform of the French higher education system and the violent events of the Arab Spring in Egypt.

The first block of this work has important policy implications on the value-added of opening a new university in a place where there was none before. It shows that the impact of university openings is sensitive to where the university opens. Indeed, opening a new university in places where the level of higher education is already high has no impact on educational attainment or labor market outcomes. However, in places where the local population is initially undereducated relative to the rest of the country, the opening of a new university increases human capital formation (10% points increase in the probability to attain at least two years of higher education) without translating into better labor market outcomes. Such a result can be interpreted as a catch-up effect following the opening of the universities: the least educated counties converge to the level of the most educated counties. Therefore, opening new universities can be seen as a lever to reducing regional disparities in the level of higher education in a given country.

Interestingly, the opening of a new university produces positive effects on education and labor market outcomes in areas that are jointly characterized by: (i) lack of access to higher education, (ii) predominantly blue-collar jobs, (iii) low initial levels of higher education, and (iv) high unemployment rates. In such areas, opening a new university increases the probability of being employed by about 8% points and increases wages by 5%. Finally, this block also

demonstrates that the effect of university openings on labor market outcomes operates via two channels: the human capital channel, and the spillover channel (e.g., potential overeducation and/or signalling - employer learning).

This thesis is also the first to focus on the impact of the Arab Spring on outcomes that are crucial for women's labor force participation, namely entry into marriage and motherhood. As a matter of fact, the second block contributes to the literature on the impact of conflicts on women's well-being in two important respects. On the one hand, it documents in the case of the Arab Spring in Egypt, that the uncertainties created by violent events encourage rural women to marry and have children earlier. This effect is likely due to the fact that women are turning to marriage as a safety in the context of growing conservatism, consistent with the empirical evidence that conservatism is higher in the aftermath of conflict. On the other hand, authors as [El-Mallakh et al. \(2018\)](#) and [Bargain et al. \(2019\)](#) have documented positive implications of the Arab Spring on women's empowerment, through the lens of labor market participation and household bargaining power. The findings of this chapter on marriage and motherhood somewhat moderate these positive effects.

What remains to be known?

Regarding the first block of the thesis, further research is needed to better understand the spillover effects through which universities impact labor market outcomes. The spillover channel consisting in employer learning-statistical discrimination deserves a closer attention. In this sense, an interesting question to ask would be how long it takes for employers to reveal the true productivity of workers graduating from new universities. Moreover, policymakers may want to know whether the opening of new universities encourages businesses to locate nearby. Indeed, there are many reasons to believe that businesses will want to locate near new universities, namely: (i) taking advantage of the highly skilled workforce, (ii) meeting the growing demand for consumer services (e.g., restaurants and stores for students and staff), and (iii) collaborating on Research and Development.

With regard to the second block, more effort should be devoted to analysing other aspects of women's well-being in times of conflicts. For example, women's and children's health is highly sensitive to the onset of conflict, yet to date, little is known about how the Arab Spring

have affected these outcome variables. This is particularly important given the recent political unrest in the region, such as the Syrian refugee crisis, where camps are often unhygienic putting women health at risk. Similarly, the consequences of sexual harassment experienced by women during demonstrations are still not known and hence deserve to be addressed. Health conditions and sexual harassments are not only important determinants of women labor force participation but are also important factors in income inequality within the household and women financial dependence on men. At last, there is a need to cover other MENA countries (Bahrain, Tunisia, etc.) where the effects of the Arab Spring are not well documented, which would allow for comparisons and trends assessment.

Bibliography

Abadie, Alberto (2005), “Semiparametric difference-in-differences estimators.” *The Review of Economic Studies*, 72, 1–19.

Altonji, Joseph G and Charles R Pierret (2001), “Employer learning and statistical discrimination.” *The quarterly journal of economics*, 116, 313–350.

Ambrosetti, Elena, Aurora Angeli, and Marco Novelli (2019), “Ideal Family Size and Fertility in Egypt: An Overview of Recent Trends.” *Statistica*, 79, 223–244.

Angrist, J. and Jorn-Steffen Pischke (2009), *Mostly Harmless Econometrics*. Princeton University Press.

Anselin, Luc, Attila Varga, and Zoltan Acs (1997), “Local geographic spillovers between university research and high technology innovations.” *Journal of urban economics*, 42, 422–448.

Bargain, Olivier, Delphine Boutin, and Hugues Champeaux (2019), “Women’s political participation and intrahousehold empowerment: Evidence from the Egyptian Arab Spring.” *Journal of Development Economics*, 141, 102379.

Becker, Gary (1993), “Human capital: a theoretical and empirical analysis, with special reference to education.” *New York: National Bureau of Economic Research, Columbia*.

Becker, Gary S (1962), “Investment in human capital: A theoretical analysis.” *Journal of political economy*, 70, 9–49.

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

- Beffy, Magali, Denis Fougère, and Arnaud Maurel (2012), “Choosing the field of study in postsecondary education: Do expected earnings matter?” *Review of Economics and Statistics*, 94, 334–347.
- Belzil, Christian and François Penas (2010), “Education and early career outcomes of second-generation immigrants in France.” *Labour Economics*, 17, 101–110.
- Berlingieri, Francesco, Christina Gathmann, and Verena Lauber (2017), “College Openings and Local Economic Development.” *working paper*.
- Bordón, Paola and Breno Braga (2020), “Employer learning, statistical discrimination and university prestige.” *Economics of Education Review*, 77, 101995.
- Brodaty, T., Robert Gary-Bobo, and Ana Prieto (2014), “Do risk aversion and wages explain educational choices.” *Journal of Public Economics*, 117, 125–148.
- Burde, Dana and Leigh L. Linden (2013), “Bringing education to Afghan girls: A randomized controlled trial of village-based schools.” *American Economic Journal: Applied Economics*, 5, 27–40.
- Cameron, A. Colin and Douglas L. Miller (2015), “A practitioner’s guide to cluster-robust inference.” *Journal of Human Resources*, 50, 317–372.
- Cantoni, Davide and Noam Yuchtman (2014), “Medieval universities, legal institutions, and the commercial revolution.” *The Quarterly Journal of Economics*, 129, 823–887.
- Card, David (1995a), “Using geographical variation in college proximity to estimate the return to schooling.” In *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp* (L.N. Christofides, E. K. Grant, and R. Swidinsky, eds.), 201–222, University of Toronto Press, Toronto.
- Card, David (1995b), “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.
- Card, David (2001), “Estimating the return to schooling: Progress on some persistent econometric problems.” *Econometrica*, 69, 1127–1160.

Carnevale, Anthony P, Ban Cheah, and Stephen J Rose (2011), “The college pay off.”

Chaney, Eric (2013), “Revolt on the Nile: Economic shocks, religion, and political power.” *Econometrica*, 81, 2033–2053.

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 Versailles-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.

Cour des comptes (2008), “Les universités des villes nouvelles franciliennes: bilan et perspectives.” In *Rapport public annuel* (Cour des comptes, ed.), 155–185.

- Cox, David R (1972), “Regression models and life-tables.” *Journal of the Royal Statistical Society: Series B (Methodological)*, 34, 187–202.
- Currie, Janet and Enrico Moretti (2003), “Mother’s education and the intergenerational transmission of human capital: Evidence from college openings.” *The Quarterly Journal of Economics*, 118, 1495–1532.
- 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.
- Donald, Stephen G. and Kevin Lang (2007), “Inference with Difference-in-Differences and other panel data .” *Review of Economics and Statistics*, 89, 221–233.
- Duflo, Esther (2001), “Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment.” *American Economic Review*, 91, 795–813.
- El-Mallakh, Nelly, Mathilde Maurel, and Biagio Speciale (2018), “Arab spring protests and women’s labor market outcomes: Evidence from the Egyptian revolution.” *Journal of Comparative Economics*, 46, 656–682.
- Enquêtes et études démographiques (INSEE) (2017), “Les naissances en 2016 - État civil - Insee Résultats.” Insee insee résultats - août 2017.
- Falch, Torberg, Paivi Lujala, and Bjarne Strom (2013), “Geographical constraints and educational attainment.” *Regional Science and Urban Economics*, 43, 164–176.
- Frenette, Marc (2009), “Do universities benefit local youth? Evidence from the creation of new universities.” *Economics of Education Review*, 28, 318 – 328.
- Gagnol, Laurent and Jean-Alain Héraud (2001), “Impact économique régional d’un pôle universitaire: application au cas strasbourgeois.” *Revue d’Economie Régionale & Urbaine*, 581–604.
- Garrouste, Manon and Meryam Zaiem (2017), “Geographical constraints in track choices: a French study using high school openings.” mimeo.

- Gelbach, Jonah B (2016), “When do covariates matter? And which ones, and how much?” *Journal of Labor Economics*, 34, 509–543.
- Ghazalian, Pascal L (2022), “The effects of the Arab Spring on female labour force participation in the MENA region.” *Economics of Transition and Institutional Change*.
- Gobillon, L. and Thierry Magnac (2016), “Regional policy evaluation: Interactive fixed effects and synthetic controls.” *Review of Economics and Statistics*, 98, 535–551.
- Goujon, Anne and Zakarya Al Zalak (2018), “Why has fertility been increasing in Egypt?” *Population Societies*, 1–4.
- Goux, Dominique (2003), “Une histoire de l’enquête Emploi.” *Économie et statistique*, 362, 41–57.
- Graham, Bryan S, Cristine Campos de Xavier Pinto, and Daniel Egel (2012), “Inverse probability tilting for moment condition models with missing data.” *The Review of Economic Studies*, 79, 1053–1079.
- Grenet, Julien (2010), “La date de naissance influence-t-elle les trajectoires scolaires et professionnelles?” *Revue économique*, 61, 589–598.
- Heckman, James J, Hidehiko Ichimura, and Petra E Todd (1997), “Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme.” *The review of economic studies*, 64, 605–654.
- Hendy, Rana (2015), “Women’s participation in the Egyptian labor market: 1998-2012.” *The Egyptian labor market in an era of revolution*, 147.
- Jayaraman, Anuja, Tesfayi Gebreselassie, and S Chandrasekhar (2009), “Effect of conflict on age at marriage and age at first birth in Rwanda.” *Population Research and Policy Review*, 28, 551.
- Kantor, Shawn and Alexander Whalley (2014), “Knowledge spillovers from research universities: evidence from endowment value shocks.” *Review of Economics and Statistics*, 96, 171–188.
- Kling, Jeffrey R (2001a), “Interpreting instrumental variables estimates of the returns to schooling.” *Journal of Business & Economic Statistics*, 19, 358–364.

- Kling, Jeffrey R (2001b), "Interpreting instrumental variables estimates of the returns to schooling." *Journal of Business & Economic Statistics*, 19, 358–364.
- Krafft, Caroline (2020), "Why is fertility on the rise in Egypt? The role of women's employment opportunities." *Journal of Population Economics*, 33, 1173–1218.
- Lancaster, Tony (1990), *The econometric analysis of transition data*. 17, Cambridge university press.
- Lee, Kye Woo and Kisuk Cho (2005), "Female labour force participation during economic crises in Argentina and the Republic of Korea." *Int'l Lab. Rev.*, 144, 423.
- Mankiw, N. Gregory, David Romer, and David N. Weil (1992), "A contribution to the empirics of economic growth." *The Quarterly Journal of Economics*, 107, 407–437.
- Mincer, Jacob (1958), "Investment in human capital and personal income distribution." *Journal of political economy*, 66, 281–302.
- Mincer, Jacob (1974), "Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2."
- Ministère de l'Éducation Nationale (2001), "La réussite en premier cycle universitaire." *Note d'information*, 1–6.
- Moeeni, Safoura (2021), "Married women's labor force participation and intra-household bargaining power." *Empirical Economics*, 60, 1411–1448.
- Monjon, Stephanie and Patrick Waelbroeck (2003), "Assessing spillovers from universities to firms: evidence from French firm-level data." *International Journal of Industrial Organization*, 21, 1255–1270.
- Moretti, Enrico (2004a), "Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data." *Journal of Econometrics*, 121, 175–212.
- Moretti, Enrico (2004b), "Workers' education, spillovers, and productivity: evidence from plant-level production functions." *American Economic Review*, 94, 656–690.

- Nepal, Apsara, Martin Halla, and Steven Stillman (2018), “Violent conflict and the child quantity-quality tradeoff.”
- Observatoire des Territoires (OT) (2018), “Les mobilités résidentielles en France Tendances et impacts territoriaux.” Rapport 2018 de l’observatoire des territoires.
- OECD (2022), “Unemployment rate by age group (indicator).” Doi: 10.1787/997c8750-en (Accessed on 21 April 2022).
- Psacharopoulos, George and Harry Antony Patrinos (2018), “Returns to investment in education: a decennial review of the global literature.” *Education Economics*, 1–14.
- Ridgeway, Greg, Stephanie Ann Kovalchik, Beth Ann Griffin, and Mohammed U Kabeto (2015), “Propensity score analysis with survey weighted data.” *Journal of causal inference*, 3, 237–249.
- Rother, Mr Bjoern, Ms Gaelle Pierre, Davide Lombardo, Risto Herrala, Ms Priscilla Toffano, Mr Erik Roos, Mr Allan G Auclair, and Ms Karina Manasseh (2016), *The economic impact of conflicts and the refugee crisis in the Middle East and North Africa*. International Monetary Fund.
- Saing, Chan Hang and Harounan Kazianga (2019), “The long-term impacts of violent conflicts on human capital: US bombing and, education, earnings, health, fertility and marriage in Cambodia.” *The Journal of Development Studies*, 1–16.
- Sant’Anna, Pedro HC and Jun Zhao (2020), “Doubly robust difference-in-differences estimators.” *Journal of Econometrics*, 219, 101–122.
- Sharma, Radha R, Sonam Chawla, and Charlotte M Karam (2021), “Global gender gap index: world economic forum perspective.” In *Handbook on diversity and inclusion indices*, Edward Elgar Publishing.
- Shemyakina, Olga et al. (2009), “The marriage market and Tajik armed conflict.” *Brighton, UK: Households in Conflict Network (HiCN). Working Paper*, 66.
- Siegler, Benedikt (2012a), “The effect of university openings on local human capital formation: Difference-in-differences evidence from Germany.” Working paper 124, Bavarian Graduate Program in Economics Discussion Paper.

- Siegler, Benedikt (2012b), “The Effect of University Openings on Local Human Capital Formation: Difference-in-Differences Evidence from Germany.” Technical report, BGPE Discussion Paper.
- Sloane, Peter J Mavromaras et al. (2020), “Overeducation, skill mismatches, and labor market outcomes for college graduates.” *IZA World of Labor*.
- Somani, Ravi (2021), “The returns to higher education and public employment.” *World Development*, 144, 105471.
- Stuart, Elizabeth A, Haiden A Huskamp, Kenneth Duckworth, Jeffrey Simmons, Zirui Song, Michael E Chernew, and Colleen L Barry (2014), “Using propensity scores in difference-in-differences models to estimate the effects of a policy change.” *Health Services and Outcomes Research Methodology*, 14, 166–182.
- Thumerelle, Pierre-Jean (1998), “Nord-Pas-de-Calais, un rattrapage spectaculaire.” In *Développement Universitaire et Développement Territorial: L’Impact du Plan U2000 (1990-1995)* (Délégation à l’aménagement du territoire et à l’action régionale (DATAR), ed.), 89–97, La documentation Française.
- UN Women (2017), “Understanding masculinities: Results from the International Men and Gender Equality Survey- Middle East and North Africa.” UN Women report.
- UNESCO (2020), “Towards universal access to higher education: international trends.” UNESCO report.
- Valero, Anna and John Van Reenen (2019a), “The economic impact of universities: Evidence from across the globe.” *Economics of Education Review*, 68, 53–67.
- Valero, Anna and John Van Reenen (2019b), “The economic impact of universities: Evidence from across the globe.” *Economics of Education Review*, 68, 53–67.
- Valletta, Robert 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*, 313–342, University of Chicago Press.

Vermeulen, Karel and Stijn Vansteelandt (2015), “Bias-reduced doubly robust estimation.” *Journal of the American Statistical Association*, 110, 1024–1036.

Williams, Nathalie E, Dirgha J Ghimire, William G Axinn, Elyse A Jennings, and Meeta S Pradhan (2012), “A micro-level event-centered approach to investigating armed conflict and population responses.” *Demography*, 49, 1521–1546.

