Document de travail du LEM / Discussion paper LEM
2019-21

# Geographical constraints in track choices: a French study using high school openings

**Manon Garrouste**

LEM UMR 9221 / manon.garrouste@univ-lille.fr

**Meryam Zaiemy**

Statistical service of the French Ministry of Labour (Dares)

http://lem.cnrs.fr/

# Geographical constraints in track choices: a French study using high school openings

Manon Garrouste [*], Meryam Zaiem[†]

September 16, 2019

## Abstract

In this paper we study the effect of opening a new high school on pupils' schooling at the end of lower secondary education. We use high school openings to highlight the constraint local school supply exerts on individual schooling decisions. Our working sample covers all pupils enrolled in 9th grade between the school year 2007-2008 and the school year 2012-2013 in France. Our estimation strategy (a generalized difference in differences) takes advantage of the variation in time and space of the openings of high schools to estimate the causal effect of an increase in school supply on the allocation of pupils at the end of 9th grade. We show that opening a new high school significantly increases the probability for pupils from neighboring middle schools to continue in higher secondary education. The effect is only due to new high schools which propose a vocational track. Furthermore, the effect is mainly driven by low achieving students.

*Keywords*: Education, Track choice, School openings, Difference in differences, Two-way fixed effects
*JEL classification*: I21, I22, C23

[*]Univ. Lille, CNRS, UMR 9221 - Lem - Lille Économie Management, F-59000 Lille, France. Email: manon.garrouste@univ-lille.fr

[†]Statistical service of the French Ministry of Labour (Dares)

# Introduction

Over the recent years, there has been an increasing demand for liberalization of schooling decisions, in the sense that households have been asking for more freedom in their choice of schools and curricula. This demand has been met by public policies such as the distribution of vouchers allowing to pay for a school different from the catchment area one, or for a private school; school choice reforms aiming at giving parents more liberty in the choice of school for their child(ren); or the development of alternative pedagogical methods, such as Montessori, or Waldorf education. The essential rationale behind these policies would be that individuals are constrained in their schooling decisions by catchment area systems, financial constraints, information costs, or geographical constraints. But little is known about how such constraints influence schooling decisions.

What is known is that distance to school matters. A first group of papers studies the link between schooling supply and enrollment rate. They show that the probability of going to school increases significantly when new schools are built and when the distance to school decreases (Burde and Linden, 2013; Duflo, 2001; Handa, 2002; Filmer, 2007). A second group of papers focus on the link between school accessibility and pupils' performance. They identify a negative impact of distance to school on academic achievement (Burde and Linden, 2013; Falch et al., 2013). Finally, a third group of papers points out that the local school supply is key to explain whether or not individuals pursue in higher education (Dickerson and McIntosh, 2013; Frenette, 2009; Gibbons and Vignoles, 2012; Griffith and Rothstein, 2009; Spiess and Wrohlich, 2010). The literature also shows that, to a certain extent, individuals are better off if they can choose their school. In the French context, Fack and Grenet (2012) showed that the catchment area system reform had no effect on school choice in the sense that it did not significantly increase the number of pupils asking for another school than their catchment area one. In the United States, although the context is very different, Hastings et al. (2009) find that a school choice plan in North Carolina had a significant impact on school choice but ambiguous effects on academic outcomes, and Deming et al. (2014) find that attending a first-choice school increases college attainment. Studying a Tel-Aviv school choice program, Lavy (2010) shows that choice reduces the drop-out rate and increases high school achievement. The author also finds long-term positive effects on post-secondary enrollment and earnings (Lavy, 2015). The Swedish school choice reform proved to have small but positive short-term effects on academic achievement, but no effect on long-run outcomes (Wondratschek et al., 2013).

In this paper, we try to assess how opening a new high school may alleviate constraints on pupils' schooling. Our question comprises three parts. First, are individuals constrained by local school supply? To answer this, we ask whether opening a new high school is effective in making more individuals continue in higher secondary education. Second, how local school supply shapes schooling decisions? To answer this, we analyze whether pupils' allocation change when the local supply of schooling is increased by the opening of a new high school. More precisely, we analyze pupils' allocation in different tracks at the end of lower secondary education in France. Third, who are those pupils who are constrained by the local school supply? To answer this, we look at heterogenous effects with respect to pupils pre-opening results.

The main challenge is that the relationship between school supply and schooling decisions is complex, and isolating the impact of the former on the latter is not an easy task. The reason is that pupils are not randomly located relative to schools. First, schools are not evenly distributed on the territory. In France, at the beginning of the 2013 school year, there was an average of 8 high schools for every 10 000 pupils enrolled in secondary schooling. There were 13 for 10 000 pupils in the Paris district, and more than twice less in the neighboring Versailles school district. Second, households pay attention to the school supply in the neighbourhood when choosing where to live (Epple and Romano, 2003; Barrow, 2002; Chumacero et al., 2011; Bayer et al., 2007; Fack and Grenet, 2010). Unobserved characteristics of households may explain both their location (and thus the school supply they face), and their schooling decisions. For example parents with high preferences for academic achievement are expected to locate in neighbourhoods where the school supply in abundant and of good quality, and are also those with children who have the best academic outcomes, and study the longest. Then the quantity, and quality of local school supply is not exogenous from schooling preferences.

Our contribution to the literature is twofold. First, the literature on school supply and schooling decisions mainly focuses on primary education, or higher education. Little is known about schooling decisions at the secondary level. We do think that looking at decisions at the end of middle school is important, especially in the case of France, where pupils make an important choice at the end of 9th grade. They can choose between vocational and general track, and this is also the first moment when they may drop out of school. This choice has long run consequences on both achieved level of schooling and labor market outcomes. Goux et al. (2017) show that getting more low achieving pupils to follow a vocational track after middle school leads to a significant and important reduction in grade repetition and high-school drop out for those at-the-margin students. Second, exogenous shocks in local school supply are rare, and difficult to observe. We rely on high school openings to highlight the constraint local school supply exerts on pupils' schooling decisions. We use exhaustive data on 9th grade pupils from 2007 to 2013. As we are able to precisely locate middle and high schools, we are able to observe whether a high school opened in the neighborhood of a given middle school a given year. A generalized difference in differences estimation allows us to make use of the variation in time and space of high school openings to identify the causal effect of a change in local school supply on the allocation of pupils at the end of middle school.

Our results show that opening a new high school significantly increases the probability to continue in higher secondary education, and reduces the probability of dropping out. The constrained pupils seem to be pupils who would like to follow a vocational track, and who are at-the-margin of passing the end-of-9th-grade exam.

The paper is organized as follows. We first describe the institutional context of track choices at the end of 9th grade, and the administrative process of opening a new high school. We then describe the data and the estimation strategy. Another section presents some descriptive statistics. Estimation results come in the last section and we conclude with a discussion.

# 1 Institutional Context

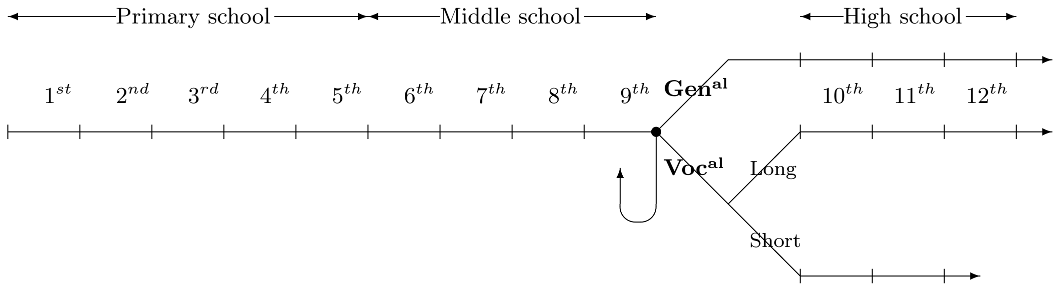
## 1.1 Track choice at the end of 9th grade

Education is compulsory in France from age 6 to age 16. Primary education lasts 5 years (from age 6 to age 10). Secondary education is divided between 4 years of lower secondary (from age 11 to age 14) in *collèges*, which are equivalent to middle schools, and 3 years of upper secondary (from age 15 to age 17) in *lycées*, equivalent to high schools.

At the end of middle school, pupils have to choose whether they would like to continue in a general or a vocational track (see Figure 1). In the general track, pupils study academic and technical subjects during three years to prepare for a general national exam (called *Baccalaurát*). The general *Baccalauréat* gives access to higher education. The vocational track provides a professional training. There are two types of vocational track. A two year track prepares for a professional certificate and a direct entry into the labor market as a skilled worker. A three year track prepares for a vocational *Baccalauréat* which gives access to qualified professions or to higher education.[1] At the end of middle school, pupils may also choose to drop out, or to repeat

[1]Both vocational tracks may be completed through a work-based training (apprenticeship).

Figure 1 – Illustration of the French school system




9th grade if their results are not sufficient to continue in high school.

The track choice procedure starts in January of 9th grade. Families have to choose between general track, vocational track, or repetition. At the end of the second term, the teaching staff responds to families' choices by providing temporary propositions of allocation. Before the end of the third and last term, families are asked to make a final choice. If their choice matches the school recommendation, the pupil is officially allocated to this track. If the school and the family disagree on the allocation, a meeting with the school headmaster is organised. If no agreement is reached at the end of the meeting, the family may resort to an appeal board. The decision of this board is final. However, whatever the decision, pupils are always free to choose to repeat 9th grade and go through the process again the next year.

After a decision about the track is made, pupils are allocated to high schools on the basis of a catchment area system. Each pupil has priority in the public general high school of her district according to where she lives. Pupils can go to another public high school through a special dispensation. If the number of dispensations exceeds the number of places in a given high school, the priority order is determined by the local education authority (*académie*) director. Allocation to a vocational high school is not based on the catchment area system, but on pupils' academic achievement. Another option is to go to a private high school, which is not subject to the catchment area system either.

## 1.2 Building new high schools

In France, the State and the three local authorities (*régions*, *départements*, and municipalities) share the responsibility for education. The State is responsible for defining the national curricula, delivering degrees, recruiting and paying teachers. *Régions* are responsible for high schools, meaning that they are responsible for the building, maintaining and functioning of high schools.

*Départements* are responsible for middle schools. And municipalities are responsible for primary schools.

Deciding to build a new high school is a long process reflecting the sharing of responsibilities between these different entities. First, on a regular basis, regions have to plan their needs in terms of middle schools and high schools, based on the demographic situation in the region and the expected number of future pupils. Second, the representative of the State at the region level approves of the region's project. If a high school needs to be built, the regional assembly then votes to allow the building. The whole building process (from selecting a service provider to realization) often takes many years. The mean duration between the regional assembly vote and the delivery of a new high school is 5 years and the cost is between 20 and 60 million euros.

The process is slightly different in the private sector since anyone can open a new private high school, though with prior notification to the local education officer (*recteur d'académie*). However, in France, almost all private schools are publicly-funded. They follow the same national curriculum as public schools (except for religious education[2]) and prepare for the same national exams, their teachers are employed by the State and local authorities are in charge of their functioning, in the exact same way as for public schools. About 20% of secondary education pupils are enrolled in a private school. 98% of them go to a publicly-funded school.

# 2 Data

To analyze the effect of opening a new high school on pupils' school choice and academic achievement, we use exhaustive micro-level data provided by the statistical service of the French Ministry of Education, both at the pupil and school levels.

We use annual exhaustive individual data sets of French secondary education pupils (called "*fichiers anonymisés d'élèves pour la recherche et les études*" or FAERE). These annual databases are composed of every pupil enrolled in a secondary school every year from 2004-2005 to 2013-2014. We focus on the 9th grade pupils enrolled in a middle school in France. Each of these pupils are observed in year t (the year of their 9th grade), and up to year t+4. The data provide the school and track of each pupil, each year. We know whether they are enrolled in a private or in a public school, whether it is a middle school or a high school, and whether it is a general or a vocational high school. For each pupil, we observe some socio-demographic characteristics: sex, age, origin, the family background through parents' occupations, and whether or not she ben-

[2]Most private schools (more than 95%) are Catholic schools.

efits from a scholarship. Pupils' scores at the end of middle school national exam (*Brevet*) are also observed. In addition, each year we know whether the pupil graduates. By the time of the analysis, the cohorts 2011 to 2013 could not be followed for four years, so that we cannot use them for the regressions on long term outcomes (graduation after 9th grade). Note that the data cover all schooling institutions but agricultural ones. Moreover, because there was a reform of the vocational track in 2007, there is a clear rupture in the data for pupils following this type of track. For that reason, estimations will be made on the cohorts enrolled in 9th grade from the 2007 school year only.

A second source of data comes from an exhaustive school-level panel data set, which provides information on every French school. Their postal address is known, so that we can observe their exact geographic location. The exact administrative date when they opened (and, if they ceased to exist, the date when they closed) is also observed.

Working with exhaustive data sets, we are able to identify, every year, high schools that appear for the first time in the data. For a given year t, a high school is considered as a new high school if some pupils are enrolled in that high school in year t while no pupil were enrolled in there the previous years. We also check that this year corresponds to the administrative date of opening. Moreover, a high school that appears only one year in the data set is not considered as an opening. As a consequence, the last cohort of the data (2013) is excluded from the working sample because we cannot know if the openings observed that year are permanent or not.

A pupil is then considered as treated if a high school opened in her middle school's neighborhood the year of her 9th grade. The treatment is thus defined at the middle school level. We tried different definitions of whether a middle school is treated or not. First, only the closest middle school to each opening high school is considered as treated. Then treated schools are extended to the two closest schools to each opening high school. Second, we used an alternative definition in which treated middle schools are those which neighborhood contains an opening high school. The neighborhood of a middle school is defined as the circle of radius $r$ centered in the middle school, where $r$ is equal to the median distance between the middle school and all high schools, weighted by the proportion of pupils going to each high school.

All these treatments are computed separately for different types of high schools. In France there are three types of high schools; those preparing for general tracks (*lycées généraux et technologiques* or LGT), those offering vocational tracks (*lycées professionnels* or LPR), and those providing both general and vocational tracks (called *lycées polyvalents*, hereafter LPO). Vocational high schools are less numerous and have a larger area of influence. Pupils going to a vocational high school have an average distance from middle school to high school of about 20 km, compared to 14 km for pupils attending a general high school. Thus we need to compute separate distances and treatments.

These definitions may be ranked from the more conservative (i.e. only the closest school is treated) to the less conservative (i.e. all schools with a new high school within their radius are treated). According to the first definition, there are as much treated middle schools as opening high schools; with the second definition, there are two treated schools for every new high school; with the third definition, there are five treated schools for every new high school; in the last case, there are about 22 treated schools for every new high school on average.

# 3 Descriptive Statistics

The main sample consists of more than 4.4 million 9th grade pupils, in about 7 000 middle schools, evenly distributed over the 6 cohorts (2007 to 2012). Among them, 60% continue in a general track, 27% go to a vocational track and 5% repeat 9th grade. The remaining 8% drop out of school or exit the data.[3]

63 new high schools opened in France over the period (Table 1). They represent about 1.6% of about 4 000 high schools. 41 were public schools and 22 were private schools. 28 were general high schools, 11 were vocational high schools and 24 were high schools providing both vocational and general tracks. On average, around 11 new high schools opened every year over the period. Figure 2 shows the locations of these new high schools. They are located in municipalities with about 160 000 inhabitants on average, compared to municipalities with an average of 180 000 inhabitants for pre-existing high schools. According to Table 2, 4 new high schools are located in rural municipalities, they represent 6.5% of new high schools, compared to 2% of pre-existing high schools being in rural areas. 24 new high schools opened in large cities with more than 200 000 but less than 2 million inhabitants, it represents 39% of opening high schools, compared to 27% of pre-existing high schools. Thus, with respect to pre-existing high schools, new high schools seem to open more often in very small or very big municipalities. To control for this, we will use the panel nature of the data. As explained later, because schools are observed at many points in time, we do not need high schools to appear randomly on

[3]Note that some pupils exit the data because they go to an agricultural school. Moreover, a few may also change identifier, so that they can no longer be tracked.

Table 1 – High schools openings in the sample by year and type

| | Number of opening high schools | | | | | |
|---|---|---|---|---|---|---|
| | Total | Public | Private | LGT | LPO | LPR |
| 2007 | 11 | 9 | 2 | 5 | 5 | 1 |
| 2008 | 10 | 8 | 2 | 4 | 3 | 3 |
| 2009 | 11 | 8 | 3 | 2 | 7 | 2 |
| 2010 | 10 | 5 | 5 | 6 | 2 | 2 |
| 2011 | 8 | 5 | 3 | 5 | 2 | 1 |
| 2012 | 13 | 6 | 7 | 6 | 5 | 2 |
| Total | 63 | 41 | 22 | 28 | 24 | 11 |
| Mean over the period | 11 | 7 | 4 | 5 | 4 | 2 |

*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2012-2013.

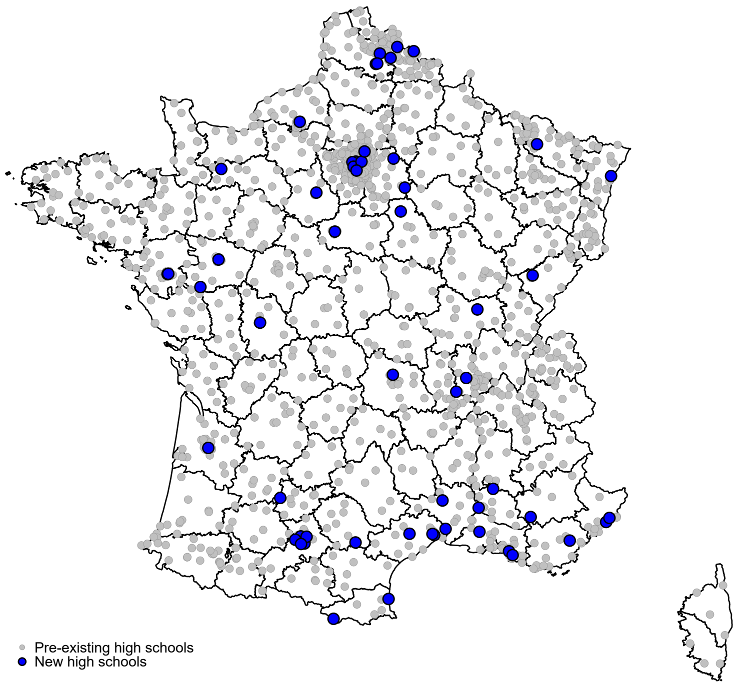
*Note:* LGT stands for general high schools, LPR for vocational high schools and LPO are high schools that provide both vocational and general tracks.

the territory.

On average, between 11 and 158 middle schools are treated each year, depending on the definition of treatment (Table 3).

In the 2007 to 2012 cohorts, about 62 000 pupils are enrolled in a new high school, that is, about 1.7% of pupils. Within treated middle schools, the share of pupils enrolling in a new high school the year when it opens varies between 8% on average, if we consider the median radius treated schools, and 30% on average, if we consider the closest treated schools (Figure 3).

Figure 2 – High school openings in mainland France (2007-2012)




*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2012-2013.

*Note:* Only mainland France and Corsica are shown on the map although the analysis also includes overseas departments.

Table 2 – Types of municipalities where high schools are located

| New high school | 0 | | 1 | |
|---|---|---|---|---|
| | Freq | % | Freq | % |
| Municipality size | | | | |
| Rural | 85 | 2.1 | 4 | 6.5 |
| $<$ 5 000 | 161 | 4.1 | 1 | 1.6 |
| $<$ 10 000 | 305 | 7.7 | 4 | 6.5 |
| $<$ 20 000 | 394 | 9.9 | 5 | 8.1 |
| $<$ 50 000 | 500 | 12.6 | 7 | 11.3 |
| $<$ 100 000 | 459 | 11.6 | 9 | 14.5 |
| $<$ 200 000 | 374 | 9.4 | 2 | 3.2 |
| $<$ 2 million | 1,074 | 27.1 | 24 | 38.7 |
| Paris | 615 | 15.5 | 6 | 9.7 |

*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2012-2013.

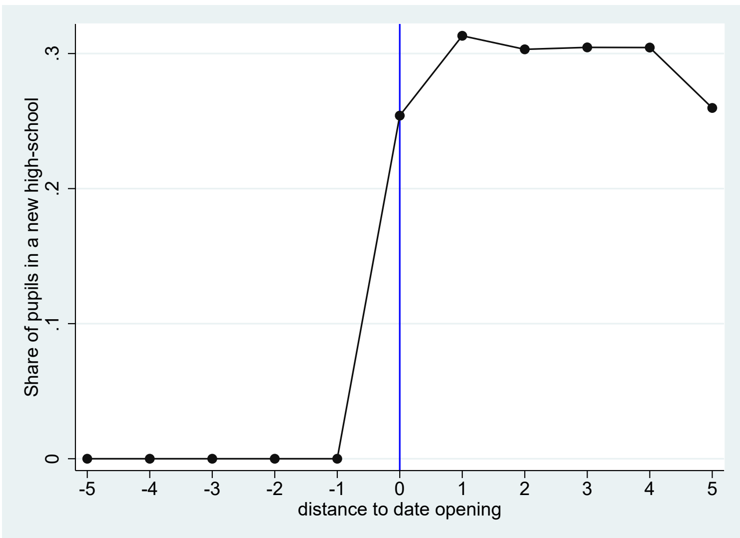
*Note:* Mayotte is excluded from this table which explain a smaller number of openings than in Table 1.

Table 3 – Treated middle schools in the sample by year and definition of treatment

| | Number of treated middle schools | | | |
|---|---|---|---|---|
| | 1st closest | 2 closest | 5 closest | Median |
| 2007 | 10 | 20 | 50 | 151 |
| 2008 | 10 | 20 | 49 | 171 |
| 2009 | 12 | 22 | 52 | 190 |
| 2010 | 10 | 20 | 47 | 106 |
| 2011 | 8 | 16 | 40 | 80 |
| 2012 | 13 | 26 | 58 | 252 |
| Total | 63 | 124 | 296 | 950 |
| Mean over the period | 11 | 21 | 49 | 158 |

*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2012-2013.

Figure 3 – Share of pupils entering a new high school




*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2012-2013.

# 4 Estimation strategy

We consider a model of repeated cross sections in which successive cohorts of 9th grade pupils are observed every year from 2007 to 2012 in $S$ middle schools. Let $Y_{ist}$ be the outcome for pupil $i$ enrolled in middle school $s$ in year $t$. $Y$ can be the track pupil $i$ is following in year $t+1$. We consider the following two-way fixed effects equation:

$$Y_{ist} = \alpha + \beta T_s \times \mathbb{1}\{t \geq t_s\} + \gamma' X_{it} + \sum_{t=2007}^{2011} \delta_t \mathbb{1}_t + \sum_{s=1}^{S-1} \mu_s \mathbb{1}_s + \varepsilon_{ist} \quad (1)$$

$T_s$ is the treatment variable with value 1 if middle school $s$ is treated and 0 otherwise. $\mathbb{1}\{t \geq t_s\}$ equals 1 for the years following the first year a new high school opened in the neighborhood of middle school $s$ and 0 otherwise[4]. $X_{it}$ is a vector of pupil $i$'s characteristics. The model includes year fixed effects, $\mathbb{1}_t$, that account for the evolution in time of track choices over the period 2007 to 2012. The middle school fixed effects control for the heterogeneity in ability and preferences across schools. The parameter of interest is $\beta$. It measures the effect of opening a new high school in the neighborhood on pupils' chosen track (and additional outcomes) at the end of 9th grade. Note that in this setting the parameters $\beta$ do not depend on $t$, meaning that we suppose the effect of the treatment to be the same whatever the date when it intervenes. We will relax this assumption later on.

When estimated by ordinary least squares in equation (1), the estimator for parameter $\beta$ is equivalent to the generalized difference in differences estimator (Bertrand et al., 2004; Hansen, 2007). It uses both the time and school dimensions and so accounts for potential selection into the treatment and for time trend. The middle school fixed effects control for the possibility that treated schools have unobserved characteristics correlated with high school openings. This means that high school openings need not to be exogenous events. The year fixed effects control, for instance, for the increase in the share of pupils following a general track over the period of observation.

The difference in differences estimator relies on the assumption of common trend between the treated and the control groups. This assumption means that, if no high school opening had occurred a given year, pupils' track choices would have evolved in the same way in treated middle schools and in non treated ones. This hypothesis cannot be tested directly, but the observation

[4]For the treatment definition based on the median of distance, some middle schools are treated twice over the observational period. In that case, we excluded observations from the year of the second opening, i.e. for these schools, $\mathbb{1}\{t \geq t_s\}$ equals 0 for the years before the first opening, 1 after the first opening, and missing starting form the year of the second opening.

of the evolution in track choices in both treated and control schools before the treatment is informative. Indeed, if pupils' track choices in both groups followed a common trend before the treatment, then assuming they would have continue to evolve in similar ways if the treatment had not occurred is a credible assumption.

In our case, the period before (or after) treatment is not the same for all middle school, since new high schools may open each year. Thus, we cannot compare the treated and control groups before treatment. However, each year, we can compare middle schools entering treatment that year with 'treated to be' (control) middle schools. Note that we consider the more conservative definition of the treatment here (only the closest middle school to a newly opened high school is treated). For every possible year $t$ of treatment, Figure 4 presents the evolution in the proportion of pupils who continue in higher secondary education until that date, both in the control and treatment groups. Overall, the graphs are inconclusive with wide confidence intervals. We thus consider a model including heterogeneous trends, i.e. a trend for each group of treated middle schools, each group being characterized by the date of treatment. The model becomes:

$$Y_{ist} = \alpha + \beta T_s \times \mathbb{1}\{t \geq t_s\} + \gamma' X_{it} + \sum_{t=2007}^{2011} \delta_t \mathbb{1}_t + \sum_{s=1}^{S-1} \mu_s \mathbb{1}_s + \sum_{g=2007}^{2011} \eta_g t + \varepsilon_{ist} \quad (2)$$

with $g$ representing a specific group of treated middle schools (those treated in 2007, 2008, 2009, 2010, or 2011).

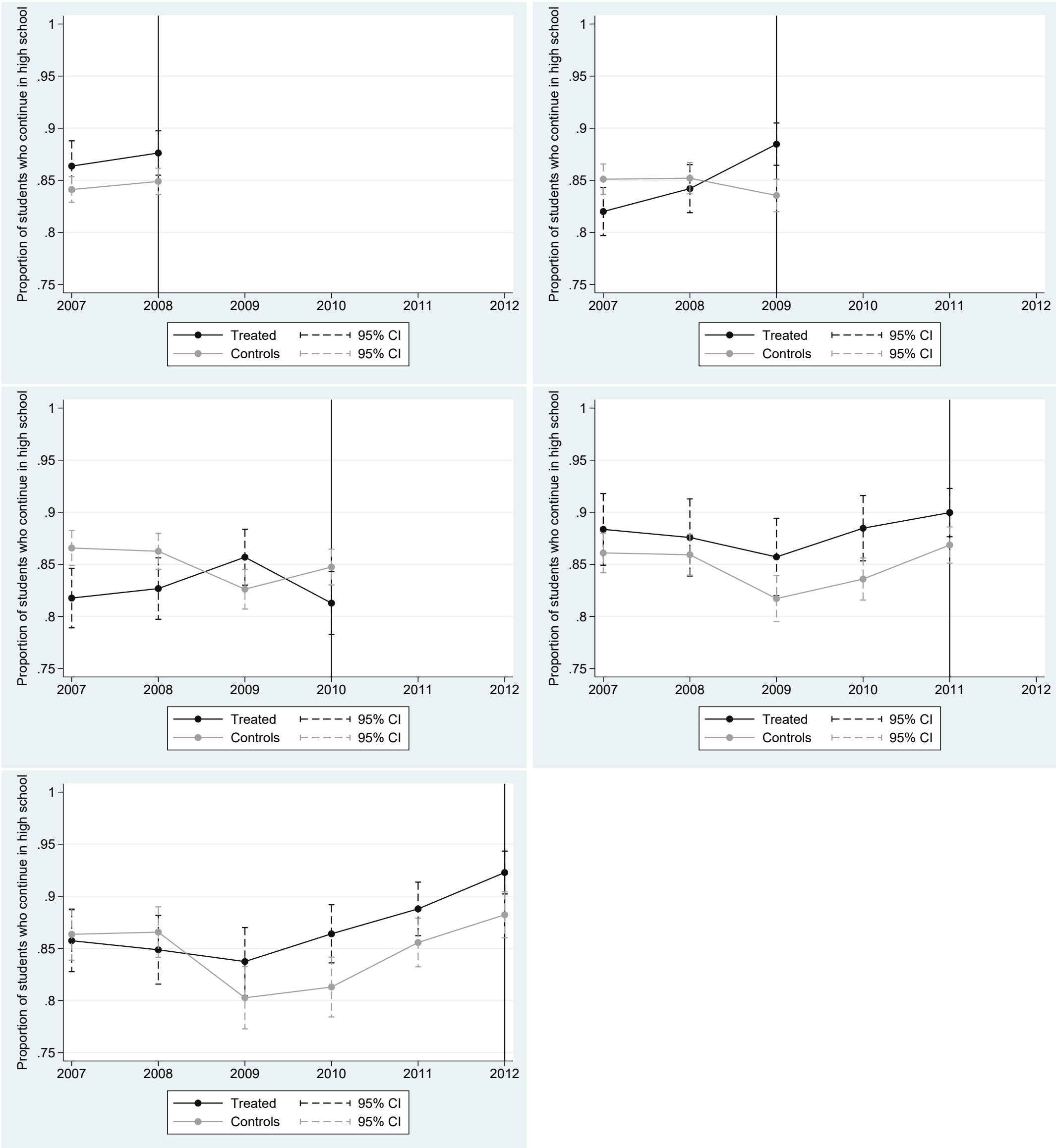
In such two-way fixed effects settings, recent papers (Abraham and Sun, 2018; Athey and Imbens, 2018; Borusyak and Jaravel, 2017; de Chaisemartin and D'Haultfoeuille, 2019; Goodman-Bacon, 2018) show that the estimated effect is a weighted average of treatment effects in each group and time period, with weights that depend on group size and treatment variance. In particular, they show that when the treatment effect is not constant over time and across groups, then the estimated effect is biased. To overcome this issue, de Chaisemartin and D'Haultfoeuille (2019) propose a new estimator corresponding to the average treatment effect of all group-time cells whose treatment status changes between two consecutive time periods.

de Chaisemartin and D'Haultfoeuille's estimator relies on two assumptions. The first one is a generalization of the traditional common trend assumption of difference in differences frameworks. It requires that the mean outcome of groups having the same treatment status in $t-1$ would have the same evolution between $t-1$ and $t$, in the absence of treatment. In our case, it means that, if no high school opening had occurred a given year, mean pupils' track choices would have evolved in the same way in treated and non treated middle schools which were not treated the preceding year. The second assumption requires that, for each year, if one middle school enters treatment, then there is at least one middle school which remains untreated. The first assumption is not testable, but it is weaker than the traditional common trend assumption. The second assumption is easy to check and will hold as long as there is a sufficient number of high schools opening every year. As another test, we will use de Chaisemartin and D'Haultfoeuille's estimator to estimate the average treatment effect of middle school-time cells whose treatment status changes between two consecutive time periods.

Another source of bias may be due to changes in the composition of the neighborhood just before treatment, that are due to treatment. First, it could be that some parents anticipated the opening of a new high school and had their child change middle school just before the opening. If such children have unobserved characteristics correlated to preferences over tracks, then we would observe a discontinuity in allocations just before the treatment (Ashenfelter *dip*) and the common trend assumption would not hold. Second, regions may anticipate a change in pupils' preferences and decide to open a new high school to satisfy the new preferences. As we have seen, the process of opening a new high school is a long one, so that the two situations discussed here are very unlikely. As a test, we can compare the composition of schools just before and after the date of opening. Figure 5 presents the evolution of treated schools 9th grade cohorts composition, before and after the treatment, with respect to observable characteristics in the data, namely the proportion of boys, the proportion of each parental occupation, the proportion of pupils born in France and the proportion of pupils benefiting from a scholarship. There is no significant discontinuity in the composition of 9th grade cohorts around the date of treatment.[5]

[5]For a formal test of a change in treated middle schools' composition the year of the treatment, see Section 5.2.

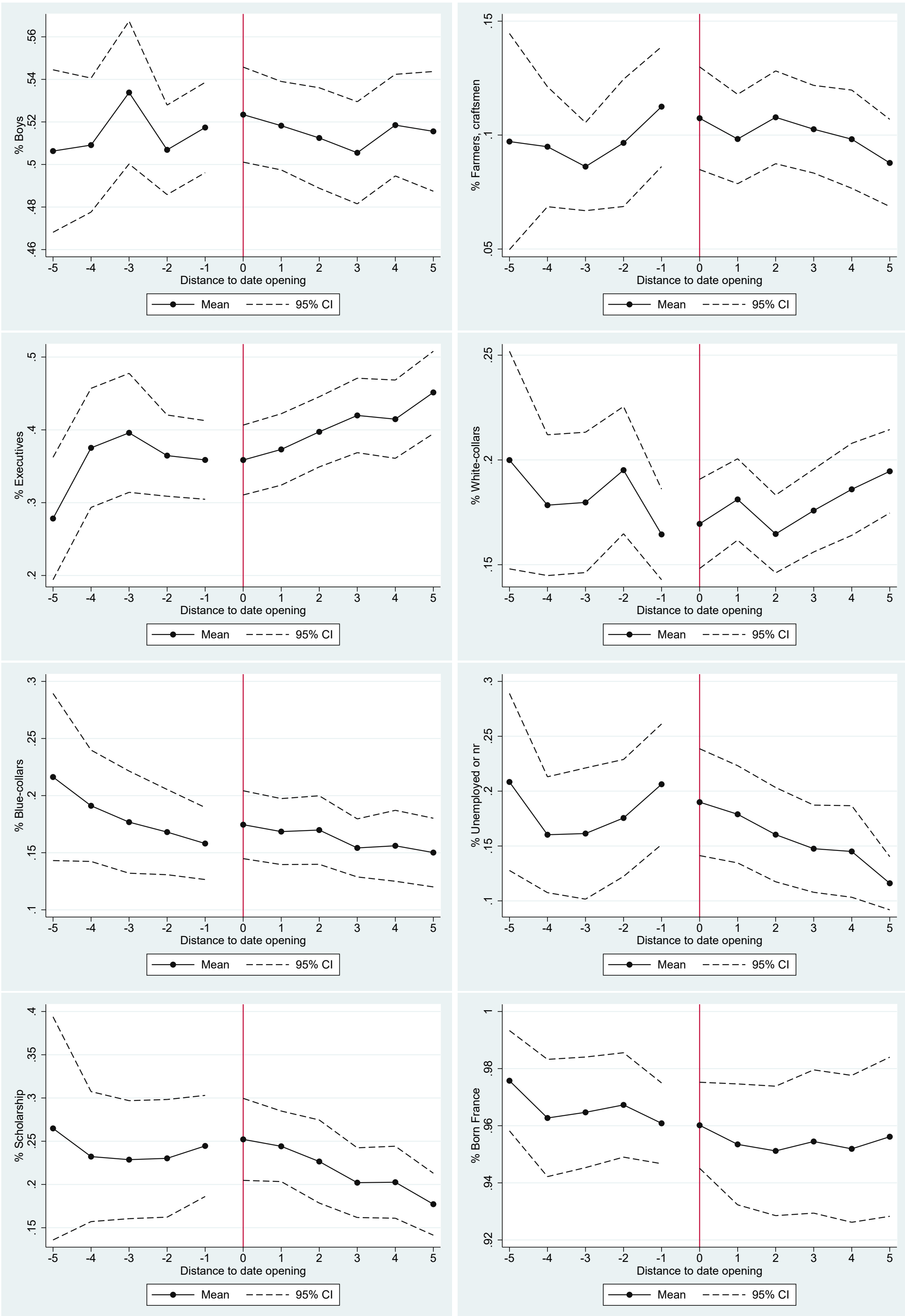
Figure 4 – Evolution of the proportion of pupils who continue in high school




*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2012-2013.

*Note:* Dashed lines represent 95% confidence intervals. Each graph plots the proportion of pupils who continue in higher secondary education in treated middle schools (in black) and in non treated middle schools (in gray) for each possible date of treatment. A treated school is defined as the closest middle school to a new public high school.

Figure 5 – Evolution of the composition of treated schools before and after an opening




*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2012-2013.

*Note:* The first graph gives, on the y-axis, the proportion of boys every year in treated middle schools, with respect to the distance to treatment on the x-axis.

# 5 Results

## 5.1 Main Results

Table 4 presents the estimates of the two-way fixed effects model presented in equation (2) on eight outcomes: going to a newly opened high school, going to high school, either by attending a general track, or by attending a vocational track; repeating 9th grade; dropping out; and getting a degree in the four year following 9th grade (*Brevet* excluded). The regression accounts for the following controls: sex, parents' occupation, scholarship status and achievement at the Brevet exam. $T(t \geq t_s)$ represents the treatment dummy, and the corresponding estimated coefficient measures the average effect of opening a new high school in treated middle schools' neighborhoods. For the moment, we present the results only for the closest treated middle schools.

Note that we only consider the opening of public high schools here. Because, as explained in Part 1, the opening of a private school is a very specific process, identifying assumptions are less likely to hold in that case.

The top panel of Table 4 presents the effect of opening a new public high school on pupils enrolled in the closest middle school, whatever the type of the new high school. The first column shows that, on average, about 35% of 9th graders enrolled in a treated middle school go to a newly opened public high school. According to the second column, the probability to continue in higher secondary education significantly increases by about 3 percentage points on average in treated middle schools, going from 82% of a cohort to more than 85%. This effect comes with a significant decrease in the probability to drop out by about 2 percentage points.

In the second panel, the treatment effect is differentiated according to the type of high school. Interestingly, the opening of a new public general high school (LGT) has no significant effect on the allocation of pupils of the closest middle school. Opening a high school providing both general and vocational tracks (LPO) however significantly impacts pupils allocation after 9th grade. The individual probability to continue in higher secondary education significantly increases by more than 4 percentage points on average, the probability of dropping out significantly decreases by almost 3 percentage points, and the probability to repeat 9th grade significantly decreases by 1.5 percentage points in treated middle schools. Opening a public vocational high school (LPR) has only little effect, by reducing the probability to drop out by almost 4 percentage points. No type of high school opening seems to have a long term impact as the effect on the probability to complete a degree after middle school is insignificant for the three types of high school.

To sum up, our results first show that opening a new high school reduces the probability for pupils in the closest middle school to drop out of school and to continue in higher secondary education. This suggests that individuals are constrained by local school supply, and that they would continue in higher secondary education if this constraints was alleviated. A second result is that the effect is driven by high schools providing vocational tracks, meaning that the pupils who are constrained are those who would go to a vocational high school, but repeat or drop out instead. Alleviating a supply constraint by opening a new vocational high school allows these pupils to continue in high school.

The affected pupils may thus be those pupils who do not perform well enough to access the general track but may continue in a vocational track if offered a place that matches their preferences. To test for this assumption, we divide 9th grade pupils into three groups, depending on their scores at the end-of-middle-school *Brevet* exam. The first group is composed of those pupils who failed the exam.[6] The second group is composed of pupils who passed the exam without honors (that is, they obtained less than 12 over 20). Pupils from the third group passed with honors (they got at least 12 over 20). Table 5 presents the heterogeneous effects of opening a new high school with respect to pupils' test scores. The top panel of the table shows that opening a new public high school has no significant effect on pupils who passed the *Brevet* exam. For pupils who failed, however, it significantly increases the probability to continue in high school by about 7 percentage points. The second panel of Table 5 further confirms that opening a high school that offers vocational tracks is the most effective in changing students allocation and achievement. Moreover, the effect is mainly driven by the pupils who failed the exam. In particular, opening a new LPO high school significantly increases the probability to continue in higher secondary education by 13 percentage points for pupils who failed the *Brevet* exam. Their probability to drop out significantly decreases by 6 percentage points and their probability to repeat 9th grade significantly decreases by about 7 percentage points. Interestingly, for those pupils, opening a new vocational high school may increase the probability to get a diploma later on. More precisely, for pupils who failed the exam, opening a new vocational high school increases the probability to graduate from high school by almost 10 percentage points, although the effect is hardly significant. Pupils who passed the exam are not affected at all by the opening of a new LGT or LPO high school, suggesting that they are not constrained by local school supply and would continue in higher secondary education whatever happens. Note, however, that pupils who pass the exam with honors may continue

[6] Note that passing the *Brevet* exam is not a prerequisite for going to higher secondary education.

in a vocational track, when a new vocational high school opens in the neighborhood of their middle schools. Given the very low number of new public vocational high school opening each year, we should interpret this results with caution, but it could be that some well performing pupils consider going to a vocational track instead of a general track if they are given the opportunity.

To sum up, the results suggest that the effect of opening a new high school is driven by low-achieving pupils, who are at-the-margin of getting the end-of-middle-school exam. For these pupils, it may even increase the probability to graduate from high school.

Table 4 – DID estimates of the effect of opening a new high school on track choice - Main specification

| | In new HS | High school | | | Repetition | Dropout | Get a diploma |
|---|---|---|---|---|---|---|---|
| | | All tracks | General | Vocational | | | Brevet excluded |
| **All types of high school** | | | | | | | |
| T(t >= ts) public | 0.346*** | 0.032** | 0.012 | 0.020 | -0.008 | -0.024** | 0.011 |
| | (0.050) | (0.015) | (0.014) | (0.012) | (0.008) | (0.010) | (0.022) |
| Intercept | -0.008 | 0.823*** | 0.506*** | 0.318*** | 0.092*** | 0.084*** | 0.757*** |
| | (0.045) | (0.015) | (0.023) | (0.024) | (0.010) | (0.011) | (0.030) |
| **By type of high school** | | | | | | | |
| T(t >= ts) LGT | 0.346*** | -0.003 | -0.018 | 0.014 | 0.005 | -0.001 | -0.018 |
| | (0.107) | (0.021) | (0.020) | (0.015) | (0.008) | (0.019) | (0.024) |
| T(t >= ts) LPO | 0.408*** | 0.044** | 0.025 | 0.018 | -0.015* | -0.028** | 0.022 |
| | (0.057) | (0.017) | (0.017) | (0.016) | (0.009) | (0.011) | (0.029) |
| T(t >= ts) LPR | 0.116** | 0.035 | 0.002 | 0.033 | 0.003 | -0.038** | 0.019 |
| | (0.054) | (0.032) | (0.016) | (0.041) | (0.019) | (0.017) | (0.023) |
| Intercept | -0.008 | 0.829*** | 0.511*** | 0.319*** | 0.090*** | 0.080*** | 0.760*** |
| | (0.047) | (0.016) | (0.023) | (0.023) | (0.010) | (0.012) | (0.030) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Heterogenous trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Nbr obs | 31,213 | 34,129 | 34,129 | 34,129 | 34,129 | 34,129 | 21,557 |
| Nbr clusters | 47 | 47 | 47 | 47 | 47 | 47 | 45 |

*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2012-2013.

*Note:* *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects, and heterogenous time trends. Controls = Gender, origin, brevet exam score, parent's occupation, scholarship status. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 5 – Separate estimates of the effect of opening by *Brevet* exam score

| | In new HS | High school | | | Repetition | Dropout | Get a diploma |
|---|---|---|---|---|---|---|---|
| | | All tracks | General | Vocational | | | Brevet excluded |
| All types of high school | | | | | | | |
| *By brevet score (ref. = Pass with honors)* | | | | | | | |
| T(t >= ts) | 0.422*** | 0.002 | -0.007 | 0.010 | -0.000 | -0.002 | -0.011 |
| | (0.070) | (0.014) | (0.020) | (0.012) | (0.002) | (0.013) | (0.028) |
| T(t >= ts) × Pass without honors | -0.082* | 0.013 | 0.034 | -0.021 | 0.005 | -0.018 | -0.007 |
| | (0.044) | (0.022) | (0.031) | (0.025) | (0.010) | (0.018) | (0.043) |
| T(t >= ts) × Fail | -0.207** | 0.072* | 0.028 | 0.044 | -0.038 | -0.034 | 0.052 |
| | (0.081) | (0.041) | (0.023) | (0.038) | (0.030) | (0.028) | (0.061) |
| By type of high school | | | | | | | |
| *By brevet score (ref. = Pass with honors)* | | | | | | | |
| T(t >= ts) LGT | 0.484*** | -0.002 | -0.027 | 0.025 | 0.001 | 0.001 | -0.007 |
| | (0.142) | (0.021) | (0.030) | (0.017) | (0.002) | (0.021) | (0.028) |
| T(t >= ts) LGT × Pass without honors | -0.222*** | -0.008 | 0.016 | -0.024 | 0.016 | -0.008 | -0.026 |
| | (0.072) | (0.032) | (0.037) | (0.026) | (0.016) | (0.031) | (0.059) |
| T(t >= ts) LGT × Fail | -0.408*** | -0.012 | 0.053 | -0.065 | 0.016 | -0.004 | -0.053 |
| | (0.136) | (0.045) | (0.043) | (0.041) | (0.022) | (0.036) | (0.066) |
| T(t >= ts) LPO | 0.475*** | 0.004 | 0.008 | -0.003 | -0.001 | -0.003 | -0.019 |
| | (0.081) | (0.015) | (0.020) | (0.014) | (0.002) | (0.014) | (0.035) |
| T(t >= ts) LPO × Pass without honors | -0.046 | 0.023 | 0.038 | -0.015 | 0.003 | -0.026 | 0.003 |
| | (0.058) | (0.021) | (0.038) | (0.030) | (0.011) | (0.018) | (0.049) |
| T(t >= ts) LPO × Fail | -0.222** | 0.134*** | 0.028 | 0.106** | -0.073** | -0.061* | 0.099* |
| | (0.092) | (0.040) | (0.024) | (0.040) | (0.031) | (0.034) | (0.057) |
| T(t >= ts) LPR | -0.010 | 0.002 | -0.051** | 0.053*** | 0.000 | -0.002 | 0.010 |
| | (0.080) | (0.016) | (0.024) | (0.019) | (0.001) | (0.015) | (0.016) |
| T(t >= ts) LPR × Pass without honors | 0.134 | 0.004 | 0.059* | -0.056 | 0.000 | -0.004 | -0.010 |
| | (0.081) | (0.034) | (0.035) | (0.049) | (0.023) | (0.016) | (0.033) |
| T(t >= ts) LPR × Fail | 0.225* | -0.026 | 0.027 | -0.053 | 0.012 | 0.014 | 0.034 |
| | (0.117) | (0.084) | (0.025) | (0.084) | (0.054) | (0.036) | (0.091) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Heterogenous trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year fe. | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Nbr obs | 29,779 | 32,018 | 32,018 | 32,018 | 32,018 | 32,018 | 20,067 |
| Nbr clusters | 46 | 46 | 46 | 46 | 46 | 46 | 44 |

*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2012-2013.
*Note:* *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects, and heterogenous time trends. Controls = Gender, origin, brevet exam score, parent's occupation, scholarship status. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

## 5.2 Robustness

The common trend assumption requires that, in the absence of treatment, treated schools would not have evolved differently from control schools. Although this hypothesis is impossible to test directly, we can check the robustness of our results to some changes in the specification of the model.

First, we tested for a change in the social composition of treated schools at the exact date of the opening of a new high school. As explained earlier, we need the school composition to have not change just before the treatment, so that we would not be able to separate the treatment effect from a modification of the treated population. To formally test this, we regressed equation (2) on the observable social characteristics: sex; parents' occupation; birthplace and scholarship status. Table 6 in the appendix gives the results. We see no discontinuity in the social composition of treated schools the year of the treatment.

Second, we ran Placebo regressions to test whether the estimated effects are not due to chance. Each school year, we randomly drew as many high schools as new high schools from the sample, and assumed that these already existing high schools were new high schools. If the effects we find are genuine, there should not be any significant effect of these Placebo high school openings. To test for this, we ran the exact same difference in differences regression, assuming that middle schools located in the neighborhood of Placebo new high schools are treated. Table 7 in the appendix gives the results. There is no significant effect of Placebo treatment on track choice.

Third, to consolidate our findings, we use de Chaisemartin and D'Haultfoeuille's estimator. The huge advantage is that it is valid when the treatment effect is not constant over time and across groups of treated units. It also allows us to estimate dynamic effects. Figure 6 in the appendix presents the average ATE on switching middle school-time cells over the period. The graphs show the effect of opening a new high school at time $t$, as well as the effects at times $t+1$ and $t+2$. The estimated effects are in line with our preceding results: opening a new public high school significantly increases the probability for pupils in the closest middle school to continue in higher secondary education and significantly decreases the probability to drop out. The dynamics though tells us something new: the impact is significant only from one year after the opening. Another new result is that the probability to graduate after middle school significantly increases.

Lastly, we test for the validity of the results with respect to the definition of treated middle schools. Figure 7 in the appendix gives the results when the two closest middle schools are considered as treated. Figure 8 presents the results when the median distance from middle schools to high schools is used to define treatment. Reassuringly, the effects are qualitatively similar. Only the effects are less significant. As expected, the less conservative the definition of treatment, the smaller the average treatment effects.

# 6 Conclusion

This paper aims at analyzing the causal effect of a change in local school supply on pupils' track choices at the end of lower secondary education. We take advantage of high school openings to highlight the constraint school supply exerts on individual schooling choices. We use an exceptionally rich data set in which we observe every pupil enrolled in 9th grade in mainland France every year from 2004 to 2013. From the data, we recover the information about new high schools each school year. A model of generalized difference in differences makes use of the variation in time and location of opening high schools to identify the causal effect of a change in local school supply on the allocation of pupils at the end of middle school.

We show that pupils are constrained by the local school supply as opening a new high school increases the proportion of pupils who continue in upper secondary education. The effect is driven by the opening of vocational high schools that induces an increase of around 4 percentage points in the probability to continue in high school for pupils enrolled in the closest middle school. This increase comes with a decrease in the probability to dropout. These results hold when the assumption of constant effect over time and across groups of treated units is relaxed. The results are driven by low-achieving pupils. Following the results of Goux et al. (2017), our findings suggest that opening new high schools that offer vocational tracks may improve pupils' long-term achievement for at-the-margin pupils.

The magnitude of the effect seems economically significant but is not easy to compare to the existing literature. First, the effect of opening a new school varies a lot across studies and countries. For instance, building a new school increases the primary education enrollment rate by 0.3 percentage point in Mozambique (Handa, 2002) and by 35 to 52 percentage points in Afghanistan (Burde and Linden, 2013). Furthermore, the expected magnitude is of course not to be the same in developing and in developed countries. Second, we do not expect to find the same magnitude in primary and in secondary education. Third, to our knowledge, there is no pre-existing study of the effect of opening a new school on enrollment in upper secondary education. Dickerson and McIntosh (2013) setting is very similar to ours, although they look at the effect of distance to education institutions on post-compulsory secondary education, and not that of the opening of a new school.

Because opening a new high school is expensive and takes time, it is worth asking whether the gain in terms of reducing dropout and increasing graduation balances the cost. In particular, some policies may induce similar effects with smaller costs. For instance, the policy studied by Goux et al. (2017) consists in organizing meetings with pupils and parents to help them build realistic educational projects. This very affordable program shows similar effects on grade repetition and dropout reduction than our results. However, one need to keep in mind that opening a new high school largely exceeds the objective of pupils' allocation at the end of middle school. First, the scope of a new high school in terms of catchment area is potentially very large (especially for vocational high schools, which may attract pupils from other regions). Second, opening a new high school also represents opportunities in terms of local employment and urban policies.

# References


Abraham, S. and L. Sun (2018). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Working paper.

Athey, S. and G. W. Imbens (2018). Design-based analysis in difference-in-differences settings with staggered adoption. Working Paper 24963, National Bureau of Economic Research.

Barrow, L. (2002). School choice through relocation: evidence from the washington, dc area. *Journal of Public Economics 86*(2), 155–189.

Bayer, P., F. Ferreira, and R. McMillan (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy 115*(4), 588–638.

Bertrand, M., E. Duflo, and S. Mullainathan (2004, Feb). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Borusyak, K. and X. Jaravel (2017). Revisiting event study designs. Working paper.

Burde, D. and L. Linden (2013). "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools". *American Economic Journal: Applied Economics 5*(3), 27–40.

Chumacero, R. A., D. Gómez, and R. D. Paredes (2011). I would walk 500 miles (if it paid): Vouchers and school choice in chile. *Economics of Education Review 30*(5), 1103–1114.

de Chaisemartin, C. and X. D'Haultfoeuille (2019, May). Two-way fixed effects estimators with heterogeneous treatment effects. Working Paper 25904, National Bureau of Economic Research.

de Chaisemartin, C., X. D'Haultfoeuille, and Y. Guyonvarch (2019). DID_MULTIPLEGT: Stata module to estimate sharp Difference-in-Difference designs with multiple groups and periods.

Deming, D. J., J. S. Hastings, T. J. Kane, and D. O. Staiger (2014). School choice, school quality, and postsecondary attainment. *American Economic Review 104*(3), 991–1013.

Dickerson, A. and S. McIntosh (2013). "The Impact of Distance to Nearest Education Institution on the Post-compulsory Education Participation Decision". *Urban Studies 50*(4), 742–758.

Duflo, E. (2001, Sep). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review 91*(4), 795–813.

Epple, D. N. and R. Romano (2003, January). *Neighborhood Schools, Choice, and the Distribution of Educational Benefits*, pp. 227–286. University of Chicago Press.

Fack, G. and J. Grenet (2010). When do better schools raise housing prices? Evidence from Paris public and private schools. *Journal of Public Economics 94*(1), 59–77.

Fack, G. and J. Grenet (2012). Rapport d'évaluation de l'assouplissement de la carte scolaire. Technical report, Centre pour la recherche économique et ses applications.

Falch, T., P. Lujala, and B. Strøm (2013). Geographical constraints and educational attainment. *Regional Science and Urban Economics 43*(1), 164–176.

Filmer, D. (2007). "If you build it, will they come? School availability and school enrolment in 21 poor countries". *Journal of Development Studies 43*(5), 901–928.

Frenette, M. (2009). Do universities benefit local youth? evidence from the creation of new universities. *Economics of Education Review 28*(3), 318–328.

Gibbons, S. and A. Vignoles (2012). Geography, choice and participation in higher education in england. *Regional Science and Urban Economics 42*(1-2), 98 – 113.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Working Paper 25018, National Bureau of Economic Research.

Goux, D., M. Gurgand, and E. Maurin (2017). Adjusting your dreams? high-school plans and dropout behaviour. *Economic Journal 127*(602), 1025–1046.

Griffith, A. and D. Rothstein (2009). "Can't get there from here: The decision to apply to a selective college". *Economics of Education Review 28*(5), 620–628.

Handa, S. (2002). "Raising primary school enrolment in developing countries: The relative importance of supply and demand". *Journal of Development Economics 69*(1), 103–128.

Hansen, C. B. (2007, Oct). Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects. *Journal of Econometrics 140*(2), 670–694.

Hastings, J., T. Kane, and D. Staiger (2009). Heterogeneous preferences and the efficacy of public school choice. *unpublished working paper*.

Lavy, V. (2010). Effects of free choice among public schools. *The Review of Economic Studies 77*(3), 1164–1191.

Lavy, V. (2015, January). Long run effects of free school choice: College attainment, employment, earnings, and social outcomes at adulthood. Working Paper 20843, National Bureau of Economic Research.

Spiess, C. and K. Wrohlich (2010). "Does distance determine who attends a university in Germany?". *Economics of Education Review 29*(3), 470–479.

Wondratschek, V., K. Edmark, and M. Frölich (2013). The Short- and Long-term Effects of Scool Choice on Student Outcomes. Evidence from a School Choice Reform in Sweden. *Annales d'Economie et de Statistique* (111-112), 4.

# 7 Appendix

Table 6 – DID estimates of the effect of opening a new high school on school social composition - Closest middle school

| | Boys | Born France | Scholarship | Farmers, craftsmen | Executives | White-collar | Blue-collar | Unemployed or nr |
|---|---|---|---|---|---|---|---|---|
| T(t >= ts) public | -0.018 | -0.001 | 0.004 | 0.007 | -0.010 | 0.003 | 0.008 | -0.008 |
| | (0.011) | (0.003) | (0.010) | (0.008) | (0.011) | (0.010) | (0.010) | (0.007) |
| Intercept | 0.512*** | 0.975*** | 0.240*** | 0.095*** | 0.342*** | 0.194*** | 0.216*** | 0.153*** |
| | (0.011) | (0.005) | (0.009) | (0.007) | (0.010) | (0.009) | (0.009) | (0.005) |
| Year fe. | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Heterogenous trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| R2 | 0.01 | 0.04 | 0.13 | 0.05 | 0.10 | 0.02 | 0.06 | 0.11 |
| Nbr obs | 34,712 | 34,712 | 34,712 | 34,712 | 34,712 | 34,712 | 34,712 | 34,712 |
| Nbr clusters | 48 | 48 | 48 | 48 | 48 | 48 | 48 | 48 |

*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2010-2011.

*Note:* *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year, middle school fixed effects, and heterogenous time trends. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

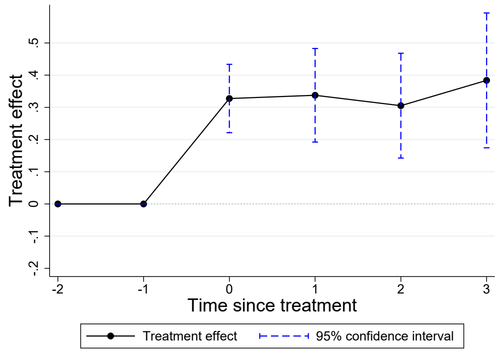
Table 7 – DID estimates of the effect of opening a Placebo new high school on track choice

| | In new HS | High school | | | Repetition | Dropout | Get a diploma |
|---|---|---|---|---|---|---|---|
| | | All tracks | General | Vocational | | | Brevet excluded |
| All types of high school | | | | | | | |
| T(t >= ts) public | -0.010 | 0.005 | -0.003 | 0.008 | 0.003 | -0.009 | 0.010 |
| | (0.014) | (0.008) | (0.010) | (0.009) | (0.005) | (0.007) | (0.012) |
| Intercept | 0.283*** | 0.865*** | 0.557*** | 0.308*** | 0.061*** | 0.074*** | 0.767*** |
| | (0.026) | (0.026) | (0.020) | (0.029) | (0.015) | (0.018) | (0.019) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Heterogenous trends | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Nbr obs | 32,165 | 34,712 | 34,712 | 34,712 | 34,712 | 34,712 | 23,190 |
| Nbr clusters | 48 | 48 | 48 | 48 | 48 | 48 | 48 |

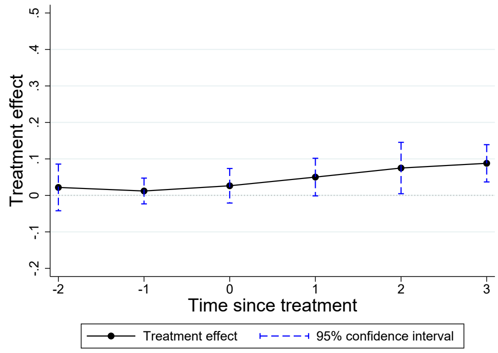
*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2010-2011.
*Note:* *** p-value<0.001, ** p-value<0.05, * p-value<0.1. Each year, as many high schools as new high schools were randomly drawn from the sample. Coefficients represent the DID estimates of the fake opening of these high schools. All estimations use year, middle school fixed effects, and heterogenous time trends. Controls = Gender, origin, brevet exam score, parent's occupation, scholarship status. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Figure 6 – Average ATE of opening a new high school on switching cells

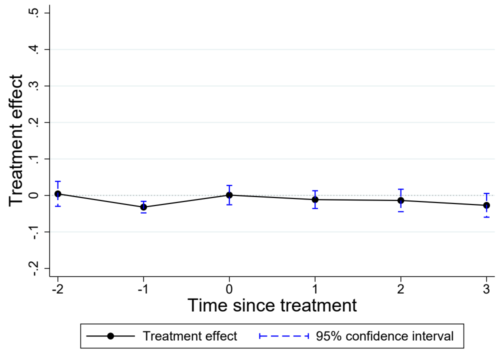
(a) Proba to go in new high school



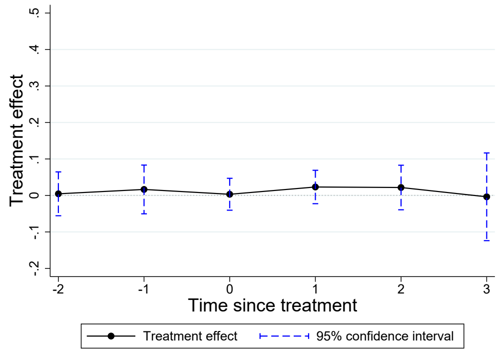

(b) Proba to continue in high school



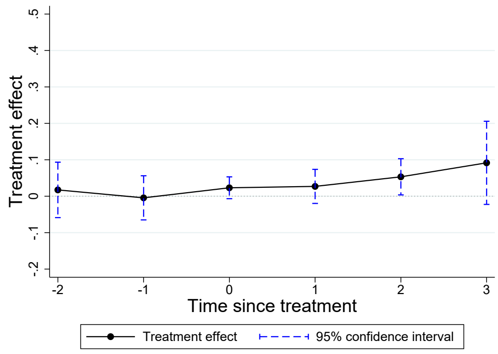

(c) Proba to repeat



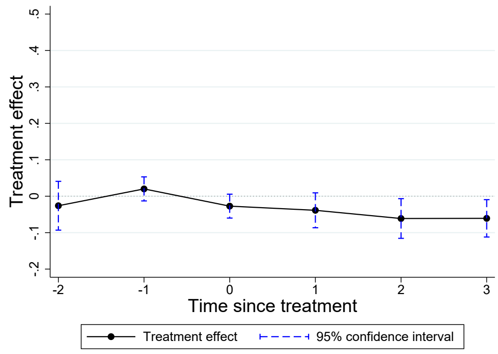

(d) Proba go to general track



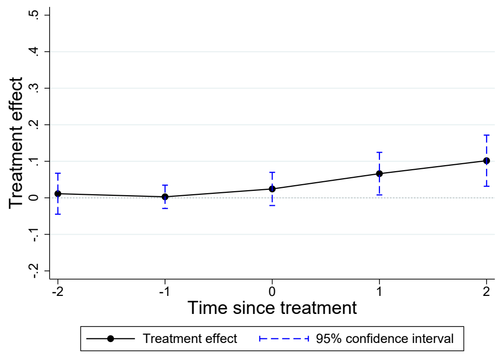

(e) Proba to go to vocational track




(f) Proba to drop out




(g) Proba to graduate

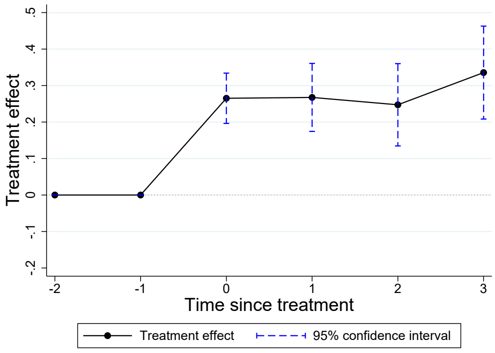



*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2010-2011.

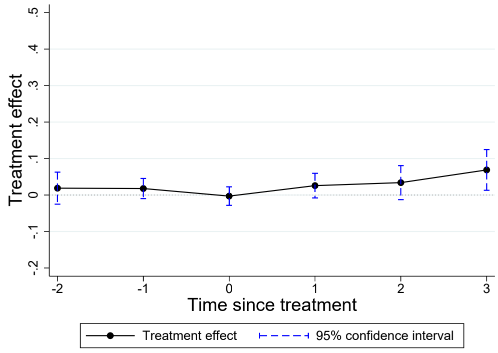
*Note:* The plots present the average ATE on switching cells as proposed by de Chaisemartin and D'Haultfoeuille (2019), using the command described in de Chaisemartin et al. (2019)

Figure 7 – Average ATE of opening a new high school on switching cells - Two closest middle schools

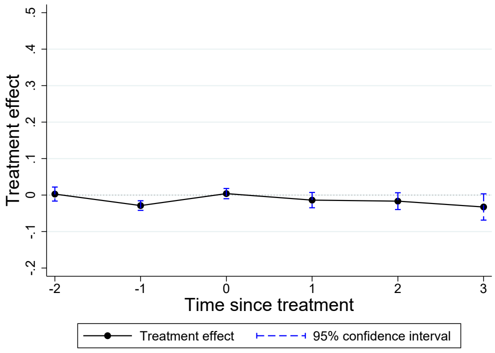
(a) Proba to go in new high school



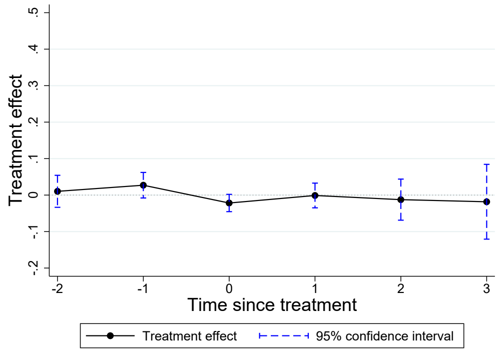

(b) Proba to continue in high school



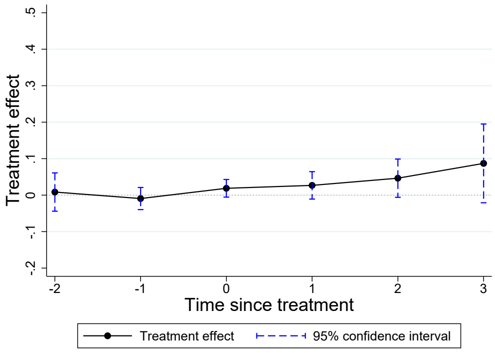

(c) Proba to repeat



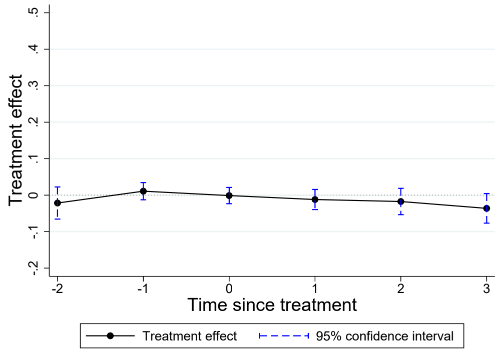

(d) Proba go to general track



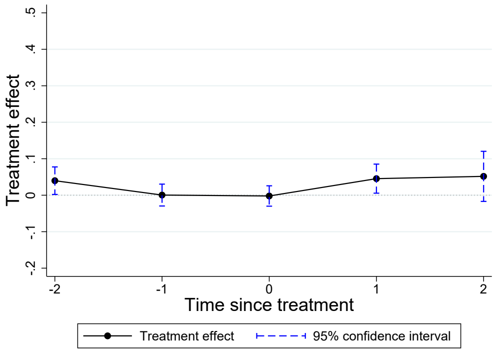

(e) Proba to go to vocational track




(f) Proba to drop out




(g) Proba to graduate

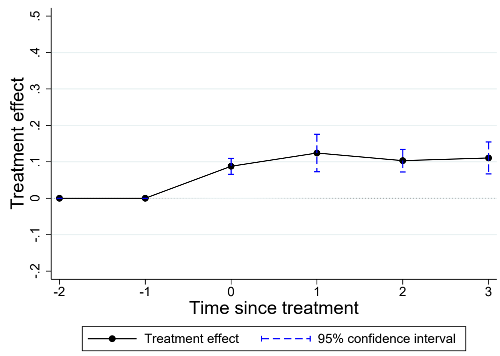



*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2010-2011.

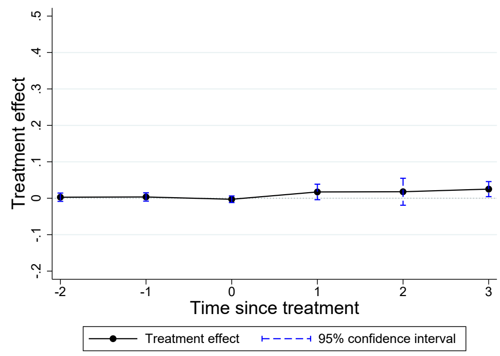
*Note:* The plots present the average ATE on switching cells as proposed by de Chaisemartin and D'Haultfoeuille (2019), using the command described in de Chaisemartin et al. (2019)

Figure 8 – Average ATE of opening a new high school on switching cells - Middle schools in median distance radius

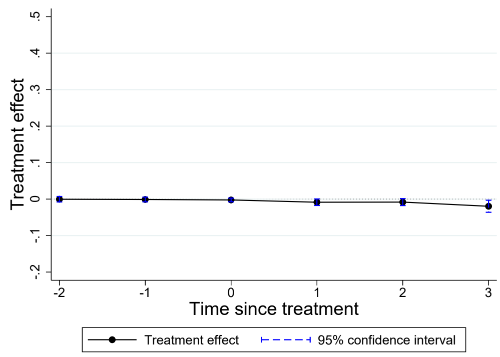
(a) Proba to go in new high school



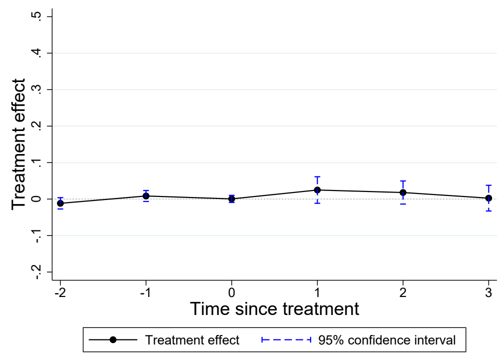

(b) Proba to continue in high school



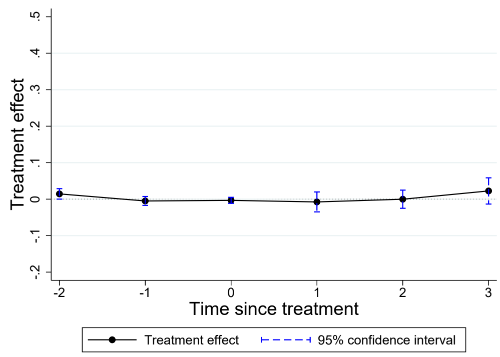

(c) Proba to repeat



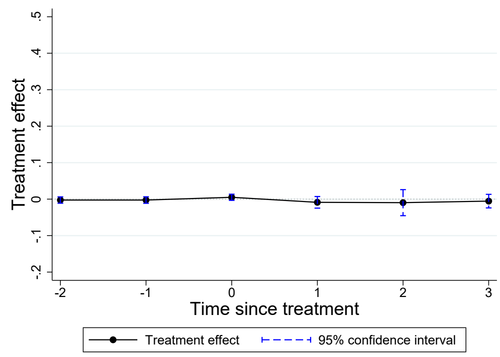

(d) Proba go to general track



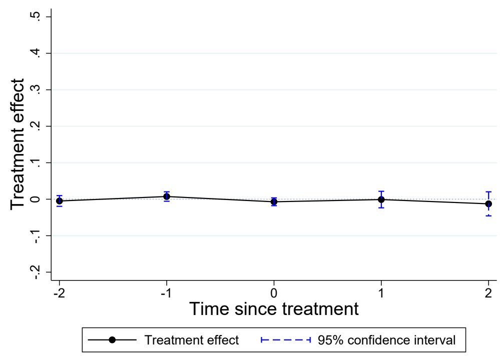

(e) Proba to go to vocational track




(f) Proba to drop out




(g) Proba to graduate




*Source:* FAERE data set, 9th grade pupils cohorts from 2007-2008 to 2010-2011.

*Note:* The plots present the average ATE on switching cells as proposed by de Chaisemartin and D'Haultfoeuille (2019), using the command described in de Chaisemartin et al. (2019)