



HAL
open science

School supply constraints in track choices: A French study using high school openings

Manon Garrouste, Meryam Zaiem

► To cite this version:

Manon Garrouste, Meryam Zaiem. School supply constraints in track choices: A French study using high school openings. *Economics of Education Review*, 2020, 78, pp.102041. 10.1016/j.econedurev.2020.102041 . hal-03129958

HAL Id: hal-03129958

<https://hal.science/hal-03129958>

Submitted on 14 Sep 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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

School supply constraints in track choices: A French study using high school openings

Manon GARROUSTE^{*}; Meryam ZAIEM[†]

Abstract

We study the effect of opening a new high school on individual schooling decisions at the end of lower secondary education. The working sample covers all ninth graders between 2007-2008 and 2012-2013 in France. The two-way fixed-effect estimation strategy uses variation in time and space to estimate the causal effect of an increase in school supply. Opening a new high school significantly increases the probability of pupils from neighboring middle schools continuing in higher secondary education. The effect is exclusively due to new high schools proposing 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 department of the French Ministry of Labor (DARES).

Introduction

Over the past decades, there has been escalating demand to liberalize schooling decisions in OECD countries, 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 vouchers issued to help pay for a different school outside the catchment area or for a private school, school choice reforms designed to give parents more freedom to decide where to school their child(ren) (Musset, 2012; OECD, 2019), and the development of alternative pedagogies methods like Montessori or Waldorf education (Sliwka, 2008). The essential rationale behind these policies is that individuals are constrained in their schooling decisions by catchment-area systems, financial constraints, information costs, and/or geographical constraints. However, little is known about how such constraints influence schooling decisions.

We do know that distance to school matters. Studies investigating the link between schooling supply and enrollment rate show that the probability of going to school increases significantly when new schools are built and when distance to school decreases (Burde and Linden, 2013; Duflo, 2001; Handa, 2002; Filmer, 2007). Papers focused on the link between school accessibility and pupils' performances find a negative impact of distance to school on academic achievement (Burde and Linden, 2013; Falch et al., 2013). A third set of papers highlight local school supply as a key factor explaining whether or not students pursue 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 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 school outside their catchment area. In the United States, although the context is very different, Hastings et al. (2009) found 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) found that attending a first-choice school increases college attainment. Studying a Tel-Aviv school choice program, Lavy (2010) showed that having a choice reduces drop-out rate and increases high school achievement, and has positive long-term 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).

Here we set out to assess how opening a new high school may alleviate constraints on pupils' schooling. We proceed with three questions. 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 does local school supply shape schooling decisions? To answer this, we analyze whether pupil allocations change when local schooling supply is increased by the opening of a new high school, based on analysis of pupil allocations in different tracks at the end of lower secondary education in France. Third, who are those pupils who are constrained by local school supply? To answer this, we look at heterogenous effects with respect to pupils' pre-opening results.

The main challenge to deal with is the complexity surrounding the relationship between school supply and schooling decisions. It is hard to isolate the impact of school supply on schooling decisions, as pupils are not randomly located relative to

schools. First, schools are not evenly distributed across the territory. In France, at the beginning of the 2013 school year, the national average was 8 high schools for every 10,000 pupils enrolled in secondary schooling. There were 13 per 10,000 pupils in the Paris district, but less than half as many in the neighboring Versailles school district. Second, households look at school supply in a prospective neighborhood 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 their school supply options) and their schooling decisions. For example, parents with strong preferences for academic achievement are both expected to locate to neighborhoods where there is abundant good-quality school supply and to have children with the best and longest academic outcomes. Thus schooling preferences are not exogenous to the quantity and quality of local school supply.

Our contribution to the literature is twofold. First, the literature on school supply and schooling decisions mainly focuses on primary education or on higher education, but little is known about schooling decisions at secondary level. We 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 ninth grade, when they can choose between a vocational track or a general track, or else elect to drop out of schooling. This choice has long-run consequences on both level of schooling attainment and labor market outcomes. Goux et al. (2017) showed that getting more low-achieving pupils to follow a vocational track after middle school leads to a significant 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. Here we use high school openings to highlight the constraint local school supply exerts on pupils' schooling decisions. We use exhaustive data on ninth grade pupils from 2007 to 2013. As we are able to precisely locate all middle and high schools, we are able to observe whether a high school was opened in the neighborhood of a given middle school in a given year. We also employ a generalized difference-in-differences estimation 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 pupil allocations at the end of middle school.

Our results show that opening a new high school significantly increases the probability of continuing 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-ninth-grade exam.

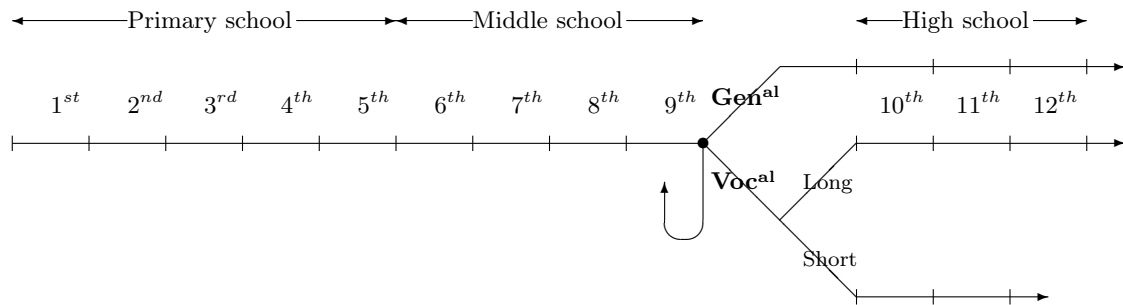
The paper is organized as follows. First we describe the institutional context of track choices at the end of ninth grade, and the administrative process of opening a new high school. We then describe the data and the estimation strategy, and go on to present key descriptive statistics. We then report the estimation results, discuss their implications and conclude.

1 Institutional context

1.1 Track choice at the end of ninth 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

Figure 1 – Illustration of the French school system



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 *Baccalauréat*). The general *Baccalauréat* gives access to higher education. The vocational track provides professional training. There are two types of vocational track: a two year track that prepares for a professional certificate and direct entry into the labor market as a skilled worker, and a three-year track that prepares for a vocational *Baccalauréat* giving 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 ninth grade if their results are not yet sufficient to continue in high school.

The track choice procedure starts in January of ninth 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 issuing temporary

¹Both vocational tracks can be completed through work-based training (apprenticeship).

allocation proposals. Before the end of the third and last term, families are asked to make a final choice. If their choice matches the school's recommendation, the pupil is officially allocated to this track. If the school and the family disagree on the allocation, then a meeting is organized with the school headmaster. If no agreement is reached at the end of the meeting, the family can resort to an appeal board. The decision of this board is final. However, whatever the decision, pupils are always free to choose to repeat ninth grade and go through the process again the next year.

After a decision about the track is made, pupils are allocated to high schools based on a catchment-area system. Each pupil has priority in the public general high school of the district in which they live. 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 director of the local education authority (*académie*). 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.

At the end of ninth grade, pupils take a national exam called "*Brevet*", which assesses the acquisition of general knowledge at the end of middle school. Taking the exam is not compulsory, but almost all ninth graders take it. The exam consists of a written exam in three core subjects (Mathematics, French, and History-Geography and Civic education) and a continuous assessment based on the grades obtained during the year of ninth grade. Pupils need to obtain more than 10 out of 20 on the final grade to pass the exam. The average success rate is more than 80%. Importantly, passing the exam is not a prerequisite for going to higher secondary

education and the result is independent from the end-of-middle-school track choice procedure. As such, it is a good measure of individual academic achievement that is not related to the opening of a new high school.

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 national curricula, delivering degrees, and recruiting and paying teachers. *Régions* are responsible for high schools, meaning that they are responsible for all allied building, maintenance and functioning involved. *Départements* are responsible for middle schools. Municipalities are responsible for primary schools.

Deciding to build a new high school is a long process that reflects shared responsibilities between these different entities. First, the regions have to regularly plan out their needs in terms of middle schools and high schools, based on the demographic picture in the region and the expected number of future pupils. Second, the region-level representative of the State has to approve the region's project. If a high school needs to be built, the regional assembly votes to enable the building project to go ahead. The whole building process (from selecting a contractor through execution) 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, provided they first notify the chief local education officer (*recteur d'académie*). However, in France, almost all private schools are publicly-

funded. This means that they follow the same national curriculum as public schools (except for religious education²) and prepare for the same national exams, their teachers are employed by the State, and the local authorities are in charge of their functioning, in the exact same way as for public schools. The main differences between public and private schools are religious instruction, the fact that private schools can ask families for fees (which are small compared to other countries as their staff in particular is paid by the State), and the fact that private schools are not subject to the catchment area system. About 20% of secondary education pupils are enrolled in a private school, and 98% of them go to a publicly-funded school.

2 Data

To analyze the effect of opening a new high school on pupils' school choices and academic achievement, we use exhaustive micro-level data, at both pupil and school level, sourced through the statistical department of the French Ministry of Education.

We use exhaustive annual individual datasets 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. Here we focus on the ninth-grade pupils enrolled in a middle school in France. Each of these pupils are observed in year t (the year of their ninth grade) and up to year $t+4$. The data provides the school and track of each pupil in each year. We know whether they are enrolled

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

in a private or 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 key socio-demographic characteristics: gender, age, origin, family background through the parents' occupations, and whether or not he/she receives a scholarship. Pupils' scores at the end-of-middle-school national exam (*Brevet*) are also observed. In addition, we know each year whether the pupil graduates. By the time of the analysis, the 2011 to 2013 cohorts could not be followed for four years, and so they could not be used for the regressions on long-term outcomes (graduation after ninth grade). Note that the data cover all schooling institutions except agricultural-vocation ones. Moreover, because there was a reform of the vocational track in 2007, there is a clear break in the data for pupils following this type of track. For that reason, estimations will be made on the cohorts enrolled in ninth grade from the 2007 school year only.

A second source of data comes from an exhaustive school-level panel dataset, which provides information on every French school. As each school's postal address is known, 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 known.

Working with these exhaustive datasets enables us to identify high schools that appear for the first time in the data, for every year. 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 pupils had been enrolled there in previous years. We also check that this year corresponds to the administrative date of first opening. A high school that

appears only one year in the dataset is not considered as an opening. Consequently, the last cohort of the data (2013) is excluded from the working sample because we cannot know whether the openings observed that year are permanent or not.

A pupil is then considered as “treated” if a high school opened in their middle school’s neighborhood in the year of their ninth grade. The treatment is thus defined at 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 was considered as treated. Second, we extended treated schools to the two schools geographically closest to each opening high school. Third, we used an alternative definition in which treated middle schools are those whose neighborhood contains an opening high school. The neighborhood of a middle school is defined as the circle of radius r centered on 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*, LGT), those offering vocational tracks (*lycées professionnels*, 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. We thus need to compute separate distances and treatments.

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

3 Descriptive Statistics

The main sample consists of more than 4.4 million ninth-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 ninth grade. The remaining 8% drop out of school or exit the data.³

Over the period, 63 new high schools opened in France (Table 1). They represented 1.6% of about 4,000 high schools, split 41 public schools and 22 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, which tended to open in municipalities counting about 160,000 inhabitants on average, whereas pre-existing high schools are sited in municipalities counting an average 180,000 inhabitants. According to Table 2, 4 new high schools were located in rural municipalities, representing 6.5% of new high schools, compared to 2% of pre-existing high schools in rural areas. 24 new high

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

schools opened in large cities with more than 200,000 but less than 2 million inhabitants, representing 39% of opening high schools, compared to 27% of pre-existing high schools in large cities. 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 can use across time variation to control for the specific nature of new high schools' neighborhoods. In other words, it means that we do not need high schools to appear randomly on the territory. We will just control for the fact that they appear in specific places by using a fixed-effect setting.

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, i.e. about 1.7% of all pupils, were enrolled in a new high school. 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 schools treated by median radius) and 30% on average (if we consider the schools treated as closest). Figure 3 presents the evolution in the proportion of pupils in treated middle schools who enroll at a new high school after ninth grade depending on the distance to the date when it opens.

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

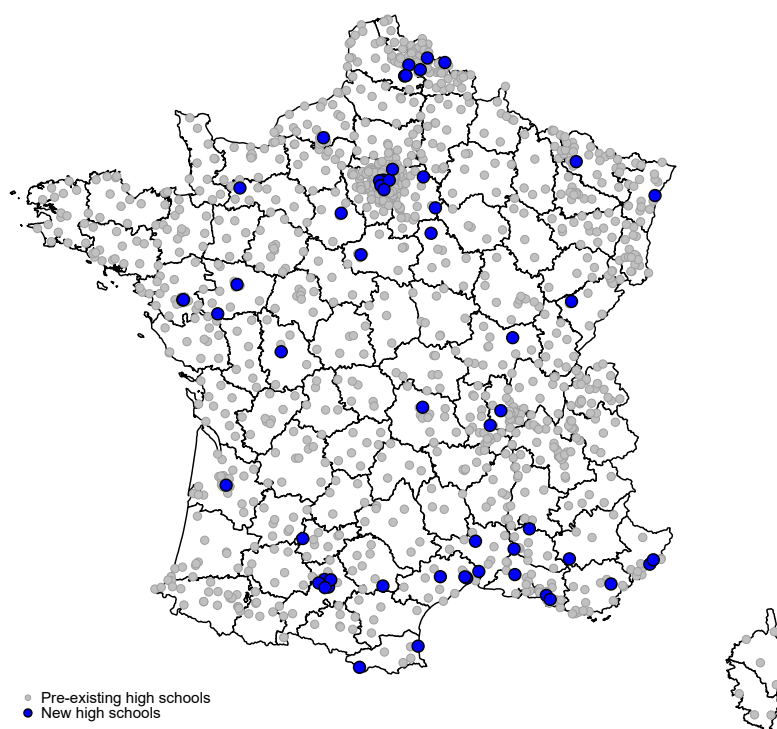
	Number of opening high schools					
	Total	Public	Private	LGT	LPO	LPR
2007	11	9	2	5	5	1
2008	10	8	2	4	3	3
2009	11	8	3	2	7	2
2010	10	5	5	6	2	2
2011	8	5	3	5	2	1
2012	13	6	7	6	5	2
Total	63	41	22	28	24	11
Mean over the period	11	7	4	5	4	2

Source: FAERE dataset, ninth-grade pupil cohorts from 2007-2008 to 2012-2013.

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

Reading note: 63 new high schools opened in France between 2007 and 2012.

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



Source: FAERE dataset, ninth-grade pupil 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 dataset, ninth-grade pupil cohorts from 2007-2008 to 2012-2013.

Note: Due to missing data, Mayotte is excluded from this table, which explains the smaller number of openings than in Table 1.

Reading note: 24 new high schools are located in municipalities with more than 200,000 but less than 2 million inhabitants. This represents 38.7% of new high schools, to be compared to 27.1% of pre-existing high schools located in this type of municipalities.

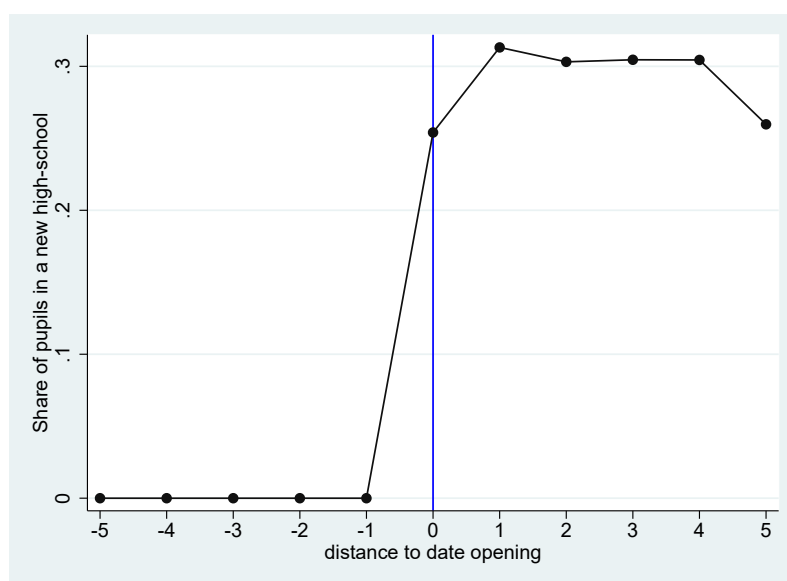
Table 3 – Treated middle schools in the sample by year and definition of treatment

	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 dataset, ninth-grade pupil cohorts from 2007-2008 to 2012-2013.

Note: In the first column, only the closest middle school to each opening high school is considered as treated. In the second column, the two closest middle schools to each opening high school are considered as treated. In the third column, the five closest middle schools to each opening high school are considered as treated. In the fourth column, every school whose radius contains an opening high school is considered as treated. Radius = weighted median distance to all high schools.

Figure 3 – Share of pupils from treated middle schools entering a new high school



Source: FAERE dataset, ninth-grade pupil cohorts from 2007-2008 to 2012-2013.

Note: A treated school is defined as the closest middle school to a new public high school.

Reading note: The black line presents the evolution in the proportion of pupils in treated middle schools who enroll at a new high school after ninth grade depending on distance to the date when it opens (located at 0).

4 Estimation strategy

We consider a model of repeated cross sections in which successive cohorts of ninth-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, or 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 , or 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 time-course change in 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, β , measures the effect of opening a new high school in the neighborhood on pupils' chosen track (and additional outcomes) at the end of ninth grade. Note that in this setting, β does not depend on t , meaning that we suppose that the treatment has the same effect regardless of the date it intervenes. We will relax this assumption later on.

When estimated by ordinary least squares in equation (1), the estimator for parameter β is equivalent to the generalized difference-in-differences estimator (Bertrand

⁴For the treatment definition based on the median 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 from the year of the second opening.

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 do not need 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 a common trend between treated and 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 as in non-treated ones. This hypothesis cannot be tested directly, but the observation of the evolution in track choices in both treated and control schools before the treatment is informative: if pupils' track choices in both groups followed a common trend before the treatment, then it would be credible to assume they would have continued to evolve in similar ways if the treatment had not occurred.

In our case, the period before (or after) treatment is not the same for all middle schools, since new high schools may open each year. This means that we cannot compare the treated groups against control groups before treatment. However, each year, we can compare middle schools entering treatment that year against 'to be treated' (control) middle schools. Note that here we consider the more conservative definition of the treatment here (only the closest middle school to a newly opened high school is treated). Figure 4 presents the evolution, for every possible year t of treatment, 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, with each group being characterized by the date of treatment. The model becomes:

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

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

In this kind of two-way fixed-effects setting, recent papers (Abraham and Sun, 2018; Athey and Imbens, 2018; Borusyak and Jaravel, 2017; de Chaisemartin and D’Haultfoeuille, 2020; 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 (2020) 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 (2020)’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 that were not treated

the previous year. The second assumption requires that, for each pair of consecutive time periods $t - 1$ and t , if one group switches from untreated to treated, then there is at least one group which remains untreated (and vice versa). In our case, it means that, for each year, if one middle school enters treatment, then at least one middle school 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. Each year, some middle schools are treated and others are not and will be treated later. In particular the assumption holds as long as treated middle schools in 2013 serve as untreated group for the middle schools treated in 2012. As another test, we use de Chaisemartin and D’Haultfoeuille (2020)’s estimator to estimate the average treatment effect of middle school-time cells whose treatment status changes between two consecutive time periods.

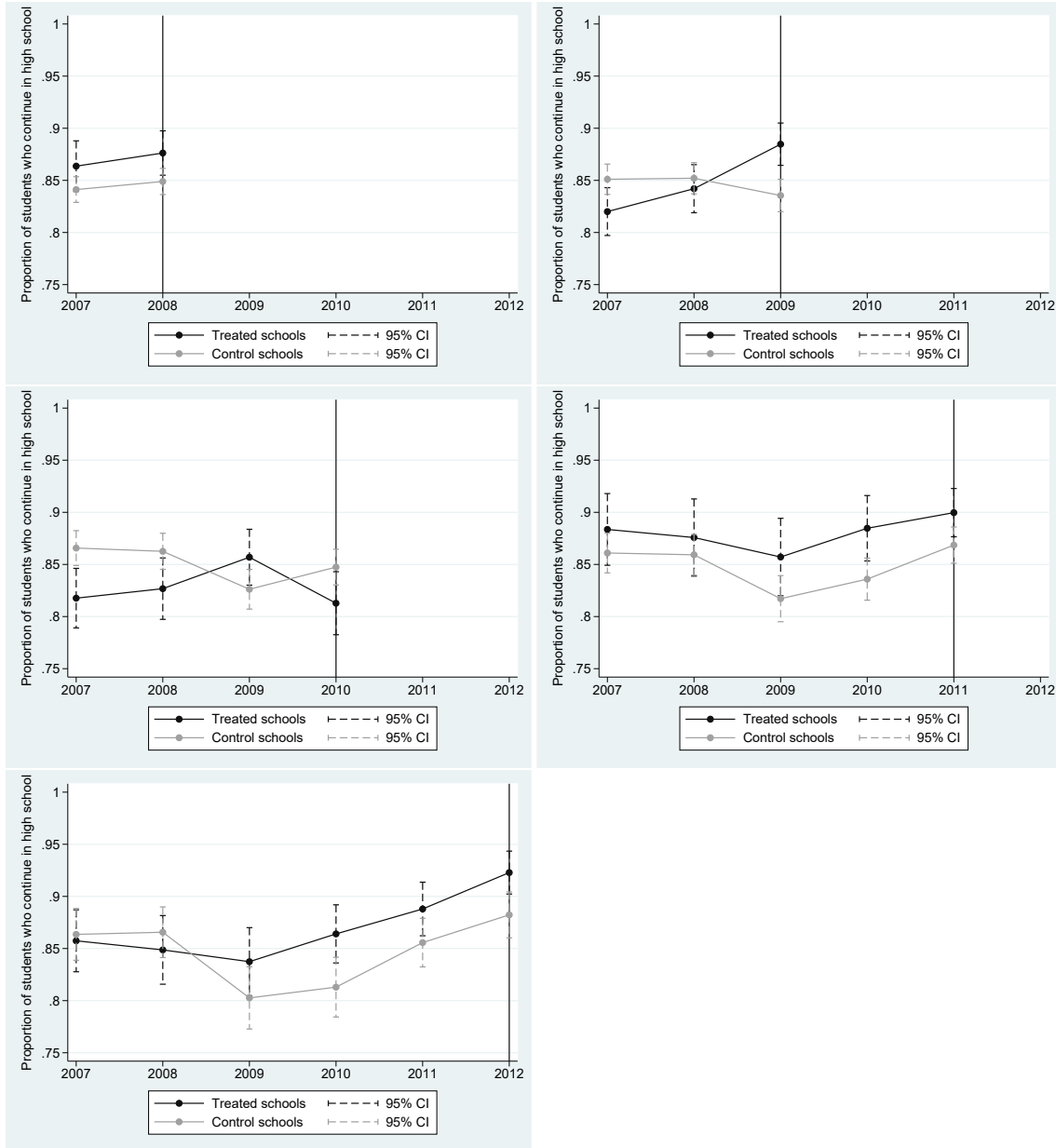
Another source of bias may come from treatment-induced changes in the composition of the neighborhood just before treatment. First, it could be that some parents anticipated the opening of a new high school and had their child change middle school just before the opening. If such children have unobserved 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 accommodate the new preferences. As a test, we can compare the composition of schools just before and after the date of opening. Figure 5 presents the compositional evolution of ninth-grade cohorts at treated schools before and after the treatment, with respect to observable characteristics in the data, namely proportion of boys, proportion of each parental

occupation, proportion of pupils born in France, and proportion of pupils receiving a scholarship. There is no significant discontinuity in the composition of ninth-grade cohorts around the date of treatment.⁵ One may worry that there is a significant change in the trend before and after the treatment for some socioeconomic characteristics (the proportion of pupils with unemployed parents for instance). However, as long as there is no significant discontinuity, adding pre- and post-opening time trends are enough to eliminate this concern.⁶

⁵See Section 5.2 for a formal test of a change in composition at treated middle schools around the year of the treatment.

⁶See Table A7 in the appendix for the estimation results including such time trends as additional controls. The results are similar to those of the following section.

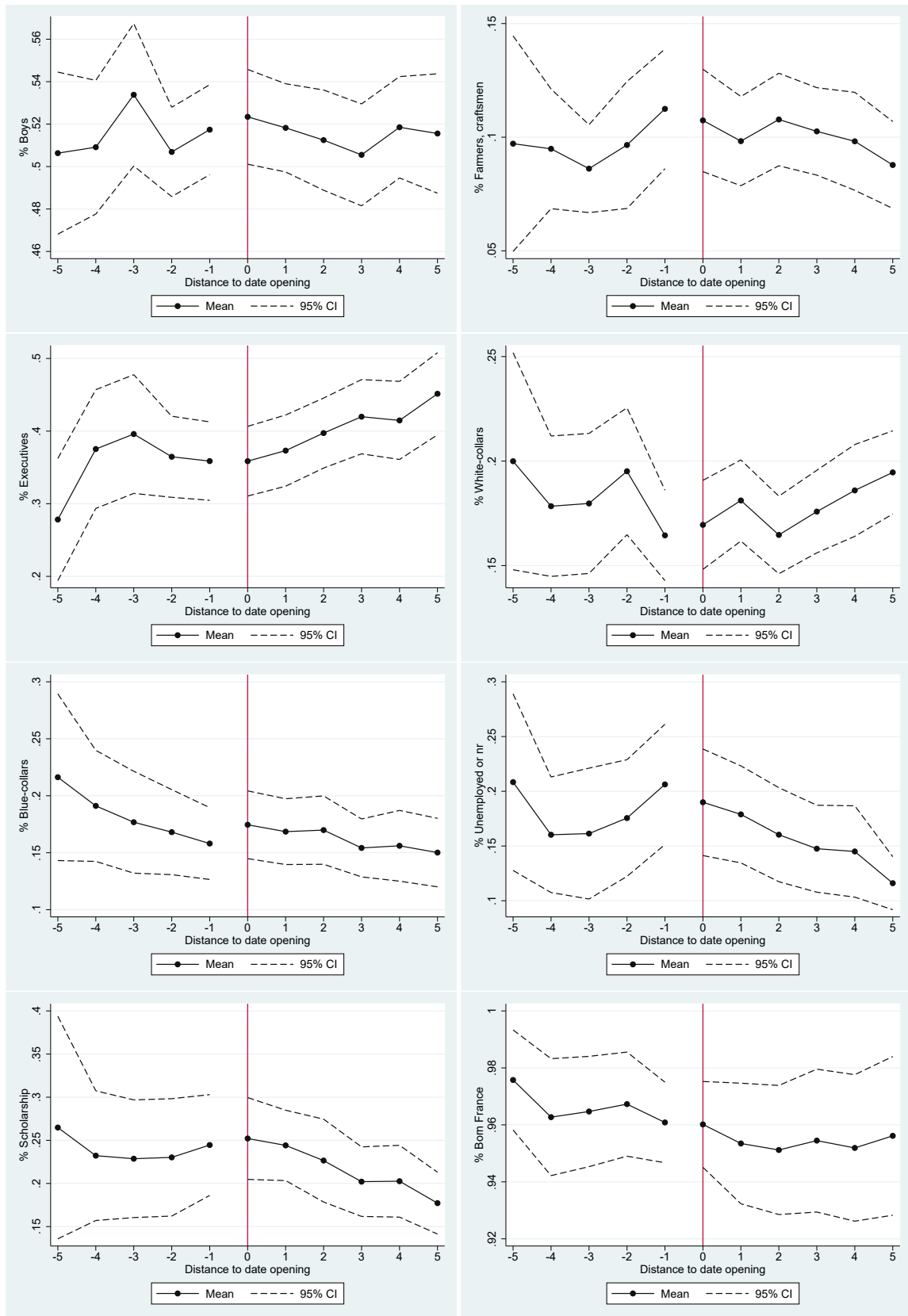
Figure 4 – Evolution of the proportion of pupils from treated and untreated middle schools who continue in high school



Source: FAERE dataset, ninth-grade pupil 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 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 dataset, ninth-grade pupil cohorts from 2007-2008 to 2012-2013.

Note: A treated school is defined as the closest middle school to a new public high school. The date when the new high school opened is located at 0. The first graph plots the proportion of boys every year in treated middle schools on the y-axis against 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, attending a general track, attending a vocational track; repeating ninth grade; dropping out; getting a diploma in the four years following ninth grade (*Brevet* excluded). The regression accounts for the following controls: gender, parents' occupation, scholarship status, 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 results only for the closest treated middle schools.

Note that we only consider the opening of public high schools here. As we saw in Part 1, the opening of a private school is a very specific process, and so the 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 (whatever the type) on pupils enrolled in the closest middle school. The first column shows that, on average, about 35% of ninth graders enrolled in a treated middle school go to a newly-opened public high school. According to the second column, the probability of continuing 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 2-percentage-point decrease in the probability of dropping out. As a comparison with the effect

of other covariates, a 2-percentage-point decrease is comparable to the difference in the probability of dropping out between pupils benefiting from a scholarship and non-scholarship pupils.

In the second panel, the treatment effect is differentiated according to type of high school. Interestingly, opening of a new public general high school (LGT) has no significant effect on the allocation of pupils from the closest middle school, whereas opening a new high school that offers both general and vocational tracks (LPO) significantly impacts pupils allocation after ninth grade. The individual probability of continuing 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 of repeating ninth grade significantly decreases by 1.5 percentage points in treated middle schools. Opening a public vocational high school (LPR) only has the effect of reducing the probability of dropping out by almost 4 percentage points. No type of high school opening appears to have a long term impact, as the effect on the probability of completing a degree after middle school is insignificant for all 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 of dropping out of school and increases their probability of continuing 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 constraint was alleviated. A second result is that this effect is driven by high schools providing vocational tracks, meaning that the pupils who are constrained are those who would like to go to a vocational high school but either 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 pupils affected 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 divided the ninth-grade pupils into two groups based on their scores at the end-of-middle-school *Brevet* exam. The first group is composed of the pupils who failed the exam.⁷ The second group is composed of pupils who passed the exam. 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 their probability of continuing in high school by 6.4 percentage points. The second panel of Table 5 further confirms that opening a high school that offers vocational tracks is the most effective driver of changing pupils' allocation and achievement. Moreover, the effect is driven by the pupils who failed the exam. In particular, opening a new LPO high school significantly increases the probability of pupils who failed the *Brevet* exam going on to continue in higher secondary education by 12 percentage points. It also significantly decreases their probability of repeating ninth grade by about 7 percentage points. In terms of magnitude, these effects are about three times larger than benefiting from a scholarship for the pupils who failed the exam. Pupils who passed the exam are not at all affected by the opening of a new high school, suggesting that they are not constrained by local school supply and would continue in higher secondary

⁷Let us remind that passing the *Brevet* exam is not a prerequisite for going to higher secondary education.

education whatever the catchment configuration.

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.

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

	In new HS	High school			Repetition	Dropout	Get a diploma Brevet excluded
		All tracks	General	Vocational			
<u>All types of high school</u>							
T(t >= ts) public	0.346*** (0.050)	0.032** (0.015)	0.012 (0.014)	0.020 (0.012)	-0.008 (0.008)	-0.024** (0.010)	0.011 (0.022)
Intercept	-0.008 (0.045)	0.823*** (0.015)	0.506*** (0.023)	0.318*** (0.024)	0.092*** (0.010)	0.084*** (0.011)	0.757*** (0.030)
<u>By type of high school</u>							
T(t >= ts) LGT	0.346*** (0.107)	-0.003 (0.021)	-0.018 (0.020)	0.014 (0.015)	0.005 (0.008)	-0.001 (0.019)	-0.018 (0.024)
T(t >= ts) LPO	0.408*** (0.057)	0.044** (0.017)	0.025 (0.017)	0.018 (0.016)	-0.015* (0.009)	-0.028** (0.011)	0.022 (0.029)
T(t >= ts) LPR	0.116** (0.054)	0.035 (0.032)	0.002 (0.016)	0.033 (0.041)	0.003 (0.019)	-0.038** (0.017)	0.019 (0.023)
Intercept	-0.008 (0.047)	0.829*** (0.016)	0.511*** (0.023)	0.319*** (0.023)	0.090*** (0.010)	0.080*** (0.012)	0.760*** (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 dataset, ninth-grade pupil cohorts from 2007-2008 to 2012-2013.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. A treated school is defined as the closest middle school to a new public high school. T(t >= ts) refers to the treatment dummy which equals one after a new high school opened. LGT=general high school, LPO=high school offering both general and vocational tracks, LPR=vocational high school. All estimations use year and middle-school fixed effects, and heterogenous time trends. Controls = Gender, origin, *Brevet* exam score, parents' occupation, scholarship status. Standard errors in parenthesis account for the autocorrelation of the residuals between observations on the same middle school.

Reading note: The individual probability of dropping out significantly decreases by 2.4 percentage points on average in treated middle schools after the opening of a new public high school.

Table 5 – Separate estimates of the effect of opening a new public high school by *Brevet* exam score

	In new HS	High school			Repetition	Dropout	Get a diploma Brevet excluded
		All tracks	General	Vocational			
<hr/>							
All types of high school							
<i>By Brevet score (ref.=Pass)</i>							
T(t >= ts)	0.388*** (0.058)	0.011 (0.013)	0.018 (0.017)	-0.007 (0.013)	0.000 (0.004)	-0.011 (0.011)	0.001 (0.025)
Fail × T(t >= ts)	-0.173** (0.065)	0.064* (0.038)	0.003 (0.022)	0.061 (0.039)	-0.039 (0.027)	-0.025 (0.026)	0.039 (0.067)
Intercept	0.136 (0.087)	0.966*** (0.014)	0.804*** (0.038)	0.162*** (0.038)	0.025** (0.011)	0.009 (0.008)	1.002*** (0.039)
<hr/>							
By type of high school							
<i>By Brevet score (ref.=Pass)</i>							
T(t >= ts) LGT	0.401*** (0.120)	-0.005 (0.025)	-0.018 (0.027)	0.013 (0.017)	0.007 (0.006)	-0.002 (0.022)	-0.012 (0.026)
T(t >= ts) LPO	0.458*** (0.066)	0.018 (0.014)	0.035* (0.020)	-0.017 (0.018)	-0.002 (0.004)	-0.016 (0.012)	0.001 (0.034)
T(t >= ts) LPR	0.062 (0.051)	0.006 (0.009)	0.001 (0.030)	0.005 (0.034)	-0.002 (0.011)	-0.004 (0.011)	0.028 (0.021)
Fail × T(t >= ts) LGT	-0.326*** (0.110)	-0.009 (0.047)	0.044 (0.045)	-0.053 (0.040)	0.010 (0.021)	-0.001 (0.034)	-0.047 (0.074)
Fail × T(t >= ts) LPO	-0.205*** (0.071)	0.121*** (0.039)	0.000 (0.022)	0.120*** (0.043)	-0.073** (0.030)	-0.048 (0.032)	0.079 (0.066)
Fail × T(t >= ts) LPR	0.153* (0.077)	-0.030 (0.069)	-0.025 (0.027)	-0.005 (0.062)	0.014 (0.043)	0.016 (0.032)	0.017 (0.102)
Intercept	0.120 (0.093)	0.976*** (0.017)	0.826*** (0.040)	0.150*** (0.041)	0.021* (0.011)	0.004 (0.013)	1.002*** (0.039)
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
Joint test of interaction terms (p-value) ^a	0.000	0.001	0.372	0.001	0.004	0.212	0.044
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 dataset, ninth-grade pupil cohorts from 2007-2008 to 2012-2013.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. A treated school is defined as the closest middle school to a new public high school. T(t >= ts) refers to the treatment dummy which equals one after a new high school opened. LGT=general high school, LPO=high school offering both general and vocational tracks, LPR=vocational high school. All estimations use year and middle-school fixed effects, and heterogenous time trends. Controls = Gender, origin, *Brevet* exam score, parents' occupation, scholarship status, and interactions with Fail. Standard errors in parenthesis account for the autocorrelation of the residuals between observations on the same middle school. ^a tests that Fail=1 × T(t >= ts) LGT, Fail=1 × T(t >= ts) LPO, and Fail=1 × T(t >= ts) LPR jointly are insignificant.

Reading note: The individual probability of repeating ninth grade significantly decreases by 7.3 percentage points on average in treated middle schools after the opening of a new LPO public high school for pupils who failed the end-of-middle-school *Brevet* national exam.

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, if the school composition had changed just before the treatment, we would be unable to separate the treatment effect from this modification in the treated population. To formally test this, we regressed equation (2) on the observable social characteristics, i.e. gender, parents' occupation, birthplace, and scholarship status. The results are reported in Table 6. We see no discontinuity in the social composition of treated schools in the year of the treatment.

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

	Students' characteristics			Parents' occupation				
	Boys	Born France	Scholarship	Farmers, craftsmen	Executives	White-collar	Blue-collar	Unemployed (or no response)
T(t >= ts) public	0.011 (0.012)	-0.006 (0.009)	0.004 (0.013)	-0.000 (0.009)	0.003 (0.010)	-0.006 (0.008)	-0.004 (0.012)	0.008 (0.007)
Intercept	0.506*** (0.008)	0.963*** (0.006)	0.283*** (0.011)	0.084*** (0.006)	0.310*** (0.009)	0.176*** (0.007)	0.213*** (0.007)	0.216*** (0.007)
Year fe.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Heterogenous trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.00	0.10	0.15	0.03	0.13	0.04	0.06	0.25
Nbr obs	34,129	34,129	34,129	34,129	34,129	34,129	34,129	34,129
Nbr clusters	47	47	47	47	47	47	47	47

Source: FAERE dataset, ninth-grade pupil cohorts from 2007-2008 to 2012-2013.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. A treated school is defined as the closest middle school to a new public high school. 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 on the same middle school.

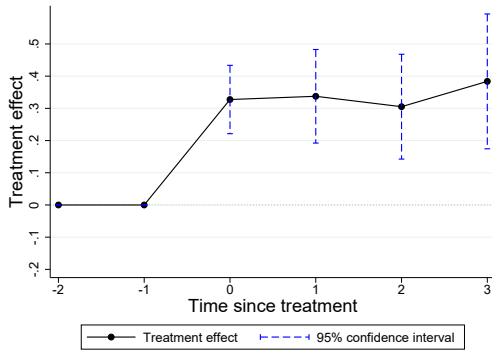
Second, to consolidate our findings, we used de Chaisemartin and D’Haultfoeuille (2020)’s estimator, which offers the huge advantage that it is valid when the treatment effect is not constant over time and across groups of treated units. The principle is to compute the mean difference in outcomes between $t - 1$ and t for middle schools which enter treatment in t and for middle schools which remain untreated between $t - 1$ and t . For each date, the average treatment effect is the difference in the mean evolution of the outcome between these two groups. The estimator is simply a weighted average of the average treatment effects at each date. A second advantage of this estimator is that it also allows us to estimate dynamic effects. Figure 6 presents the results. The graphs show the effect of opening a new high school at time t , as well as the effects at times $t + 1$, $t + 2$, and $t + 3$. The estimated effects are in line with our earlier results: opening a new public high school significantly increases the probability of pupils in the closest middle school continuing in higher secondary education and significantly decreases their probability of dropping out. The dynamics, though, tells us something new: the impact is significant only from one year after the new school opening. This may simply correspond to the transition period before the new high school project reaches its full capacity. Another new result is that the probability of graduating after middle school significantly increases. In the medium term, opening a new high school may thus increase the probability of graduating from high school.

Lastly, we tested for the validity of the results with respect to the definition of treated middle schools. Figure 7 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. Reassur-

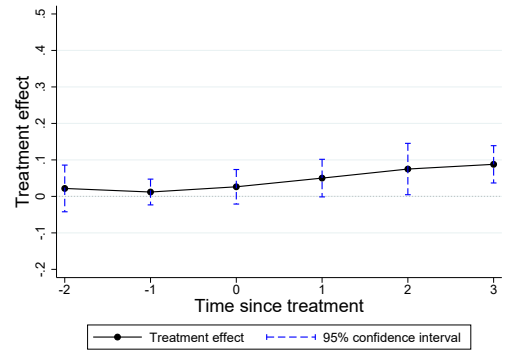
ingly, although less significant, the effects are qualitatively similar to our earlier results. As expected, average treatment effects get smaller with less conservative treatment definitions.

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

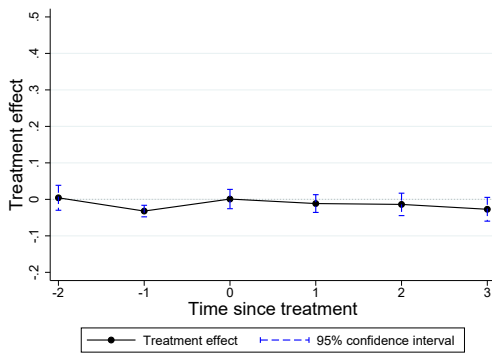
(a) Probability to go in new high school



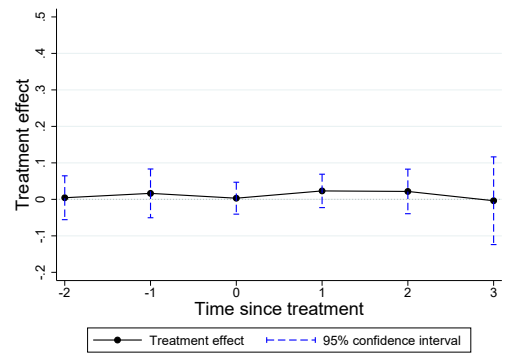
(b) Probability to continue in high school



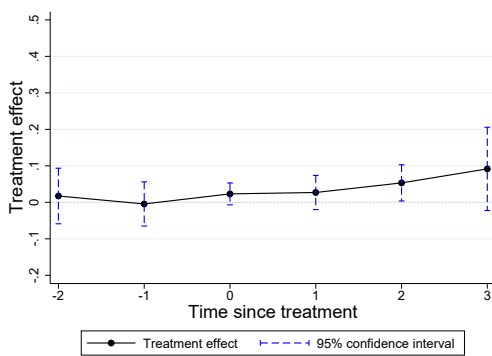
(c) Probability to repeat



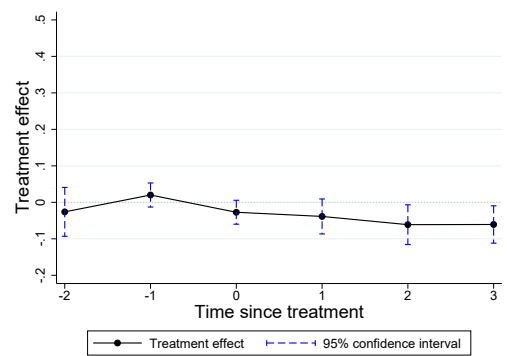
(d) Probability to go to general track



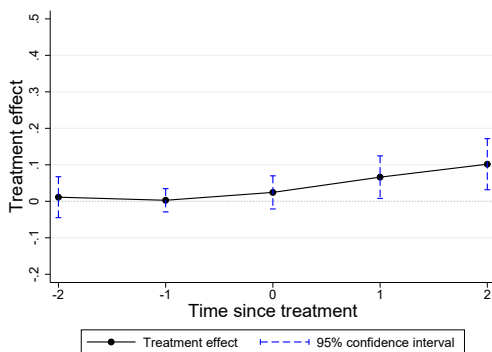
(e) Probability to go to vocational track



(f) Probability to drop out



(g) Probability to graduate

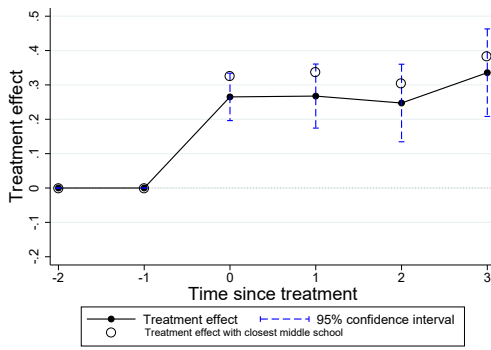


Source: FAERE dataset, ninth-grade pupil cohorts from 2007-2008 to 2010-2011.

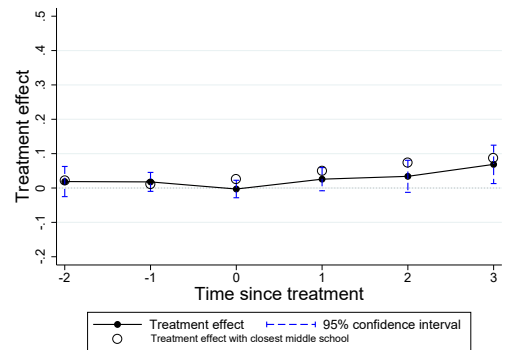
Note: A treated school is defined as the closest middle school to a new public high school. The plots present the average ATE on switching cells as proposed by de Chaisemartin and D'Haultfoeuille (2020), using the command described in de Chaisemartin et al. (2019). Confidence intervals are computed using block bootstrap with clusters at the middle school level.

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

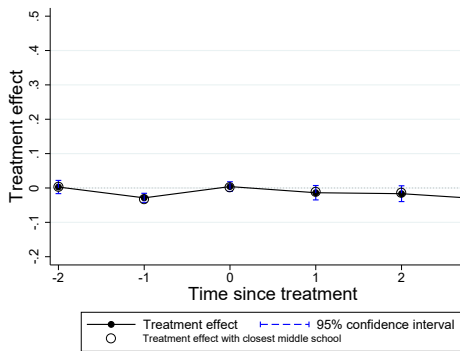
(a) Probability to go in new high school



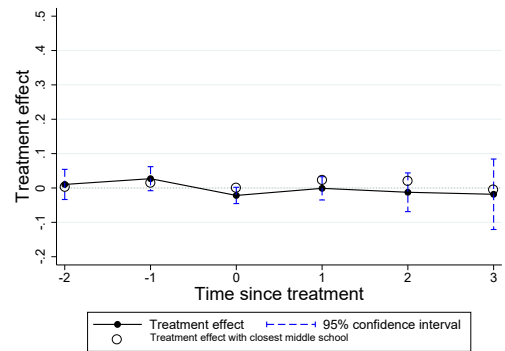
(b) Probability to continue in high school



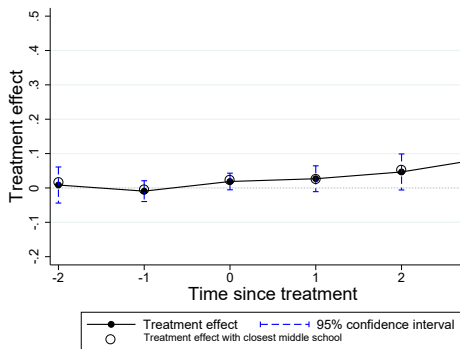
(c) Probability to repeat



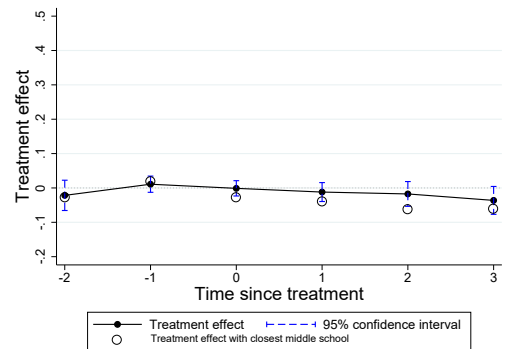
(d) Probability to go to general track



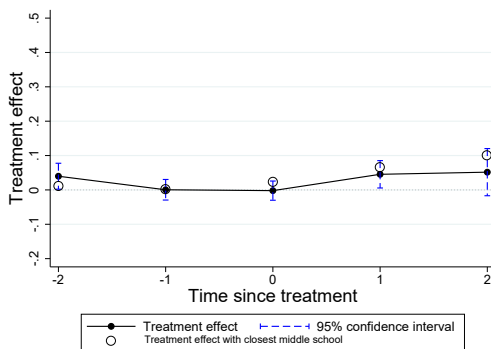
(e) Probability to go to vocational track



(f) Probability to drop out



(g) Probability to graduate

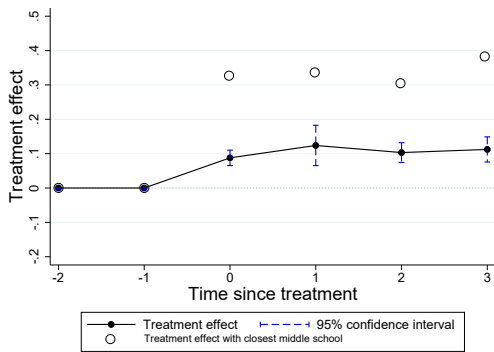


Source: FAERE dataset, ninth-grade pupil cohorts from 2007-2008 to 2010-2011.

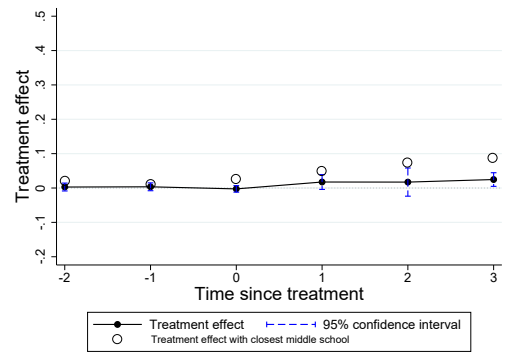
Note: A treated school is defined as being among the two closest middle schools to a new public high school. The plots present the average ATE on switching cells as proposed by de Chaisemartin and D'Haultfoeuille (2020), using the command described in de Chaisemartin et al. (2019). Confidence intervals are computed using block bootstrap with clusters at the middle school level. Hollow circles give the estimates for the closest-middle-school definition of treatment.

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

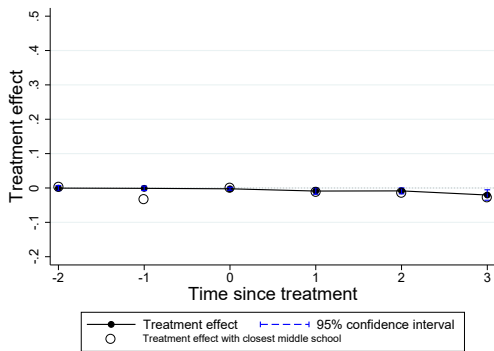
(a) Probability to go in new high school



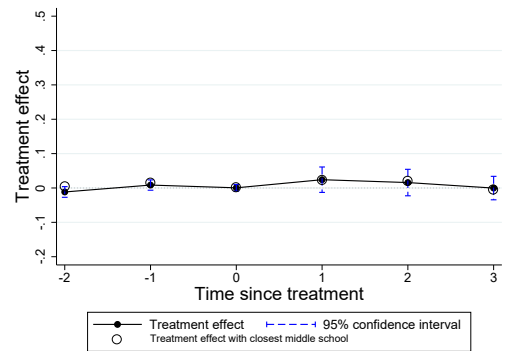
(b) Probability to continue in high school



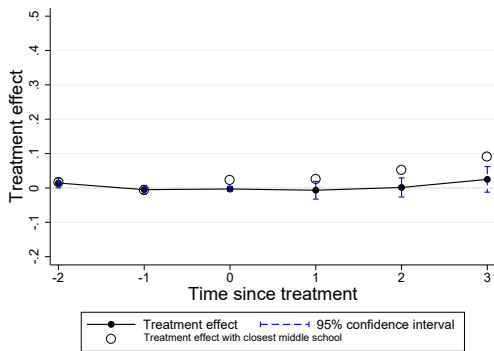
(c) Probability to repeat



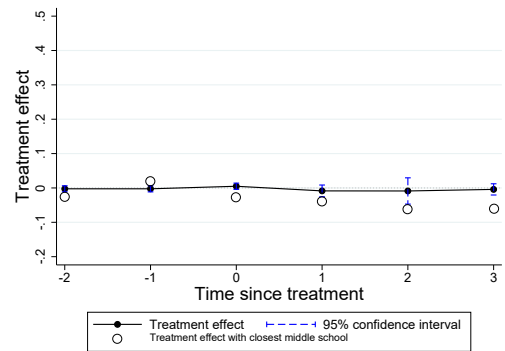
(d) Probability to go to general track



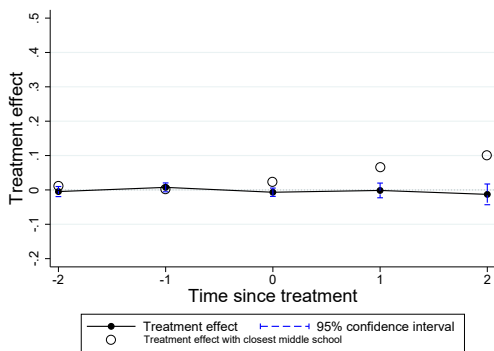
(e) Probability to go to vocational track



(f) Probability to drop out



(g) Probability to graduate



Source: FAERE dataset, ninth-grade pupil cohorts from 2007-2008 to 2010-2011.

Note: A treated school is defined as having a new public high school in its median distance radius. The plots present the average ATE on switching cells as proposed by de Chaisemartin and D'Haultfoeuille (2020), using the command described in de Chaisemartin et al. (2019). Confidence intervals are computed using block bootstrap with clusters at the middle school level. Hollow circles give the estimates for the closest-middle-school definition of treatment.

6 Discussion

Before we conclude, let us first discuss the potential mechanisms behind our results. Opening a new high school increases the local school supply in a given area. However, this could happen with or without an increase in demand. First, let us suppose that demand for high school is unchanged, meaning that the same number of pupils want to continue secondary education before and after a new high school is built, but some are constrained by local high schools capacities before. In this case, one could expect a decline in student academic quality if rationing due to supply constraints is made on the basis of academic performance. As explained in Section 1, in case of excess demand, allocation to general public high schools is not based on student past academic achievement but on a catchment area system, whereas allocation to vocational tracks is based on student academic records. Second, let us assume that demand for high school increases with the opening of a new high school, for instance because it reduces distance to high school for some students, who would not have pursued education otherwise. Then this may or may not come with a change in student academic quality, depending on the correlation between distance to high school and academic performance. As our results are both driven by the vocational tracks and by low-achieving students, they are compatible with a scenario in which supply constraint is relaxed but demand is unchanged, or demand increases but only from low-achieving pupils who want to go to a vocational track.

Then, let us briefly discuss the scope of our results and their validity outside the French context.

First, the results apply to a specific situation where pupils have to choose a track at the end of lower secondary education. As this type of selection exists

in all OECD countries, our results may shed light on the effect of local school supply in a variety of different educational systems. However the type and timing of track choice vary widely across countries, which may limit the generalization of our results. For instance, the age at which track choice occurs in France (about 15) is close to the OECD average (about 14, see OECD, 2016), but our results may not apply to countries where tracking intervenes very early in the educational process, like in Germany for instance.

Second, an important result is that the effects we found are only driven by the vocational track. The French vocational training is based on a twofold system in which pupils can prepare for the exact same diploma either through school-based training or through a work-based training. The opening of a new high school only increases the school-based training supply. In other words, our results say nothing about the effect of widening the offer of work-based training.

Third, our results are mainly driven by low-achieving pupils. One context-specific reason for this is that the vocational track is often considered a low-status choice in France. This is also true in many other countries, but our results may not necessarily apply in contexts where both high and low-achieving pupils may opt for vocational training if supply increases.

Let us now discuss the size of the effects. The magnitude of the effect seems economically significant, but it is not easy to compare against the existing literature. First, the effect of opening a new school varies widely across studies and countries. For instance, building a new school increases the primary education enrollment rate by 0.3 percentage points in Mozambique (Handa, 2002) but by 35 to 52 percentage points in Afghanistan (Burde and Linden, 2013). Furthermore, the expected

magnitude is of course not 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 on the effect of opening a new school on enrollment in upper secondary education. Dickerson and McIntosh (2013) report on a setting very similar to ours, but they look at the effect of distance to education institutions on post-compulsory secondary education, and not the effect of opening a new school.

Lastly, 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 outweighs the cost. There are other policies that may induce similar effects at lower 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 program⁸ shows similar effects on grade repetition and dropout reduction to our results. One could also compare opening a new high school with a policy consisting in reducing high school class size. According to Figure 3, about 30 pupils of a cohort of 100 from the closest middle school go to the new high school every year. On average, pupils from one middle school split up into four different high schools after ninth grade. If pupils were uniformly distributed across high schools, this would correspond to 7.5 pupils less in each high school every year. The variable cost (in terms of teachers salary and functioning of the high school) of reducing high school class size would be the same as that of opening a new high school. So the difference between the two interventions would be the fixed cost of building a new high school, which represents about 20 to 60 million

⁸The cost corresponds to organizing two meetings a year between the school headmaster and the parents.

euros as mentioned in Section 1. However, class size reduction in higher secondary education seem to have no impact on academic achievement (see Valdenaire, 2011, for the impact of class size reduction in French high schools). In any case, note that the rationales for opening a new high school stretch far beyond the objective of accommodating pupils' allocations at the end of middle school or improving individual academic achievement. 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 offers opportunities in terms of local employment and urban policies.

7 Conclusion

This paper analyzes the causal effect of a change in local school supply on pupils' track choices at the end of lower secondary education. We exploit data on high school openings to highlight the constraint school supply exerts on individual schooling choices. We use an exceptionally rich dataset in which we observe every single pupil enrolled in ninth grade in mainland France every year from 2004 to 2013. The data enables us to retrieve the key information on new high schools for each school year. A generalized difference-in-differences model 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 of continuing

in high school for pupils enrolled in the closest middle school. This increase comes with a decrease in the probability of dropping out. 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. For them, opening a vocational high school increases the probability of continuing in high school by about 12 percentage points. Following the results of Goux et al. (2017), our findings suggest that opening new high schools that offer vocational tracks may improve long-term achievement for at-the-margin pupils.

Acknowledgements

The authors would like to thank Son Thierry Ly for useful discussion and remarks on a preceding version of this paper. They gratefully acknowledge material support from the statistical service of the French Ministry of Education (DEPP). In particular, they thank Jean-Paul Caille, Agnès Brizard, Paul-Olivier Gasq, Sylvie Le Laidier, Cédric Afsa, and Caroline Simonis-Sueur for their help, as well as the participants to the DEPP’s workshop for their comments. For helpful remarks, they are grateful to Clément de Chaisemartin, Gabrielle Fack, Laurent Gobillon, Julien Grenet, Xavier d’Haultfoeuille, Nicolas Jacquemet, Francis Kramarz, Miren Lafourcade, Eric Maurin, Alain Trannoy, as well as two anonymous referees, seminar participants at CREST, University of Cergy-Pontoise, INED, University of Paris 1, Paris School of Economics, GATE Lyon-St Etienne, and Tinbergen Institute, as well as conference participants at the Education, Skills, and Labor Market Outcomes Workshop, International Workshop on Education Economics, 8th TEPP CNRS Winter School, Louis-André Gérard-Varet conference, Annual congress of the French

economic association, and Journées de microéconomie appliquée. Manon Garrouste acknowledges support from a Labex Phd grant, Labex iPOPs ANR10-LABX-0089, hosted by Institut national d'études démographiques (INED), in partnership with heSam Université. This article has been supported by the European center for humanities and social sciences (MESHS-Lille, France) and by the French Ministry of Higher education, research and innovation.

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 (2020). Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review (Forthcoming)*.
- 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? highschool 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.
- Musset, P. (2012). School Choice and Equity: Current Policies in OECD Countries and a Literature Review. OECD Education Working Papers 66, OECD.
- OECD (2016). PISA 2015 Results (Volume II). Policies and Practices for Successful Schools. OECD Publishing, Paris.
- OECD (2019). Balancing School Choice and Equity: An International Perspective Based on PISA. OECD Publishing, Paris.
- Sliwka, A. (2008). *The contribution of alternative education*, pp. 93–112. in OECD, *Innovating to learn, learning to innovate*, OECD Publishing, Paris.
- Spiess, C. and K. Wrohlich (2010). "Does distance determine who attends a university in Germany?". *Economics of Education Review* 29(3), 470–479.
- Valdenaire, M. (2011). *Essais en économie de l'éducation*. Ph. D. thesis, Paris, EHESS.
- Wondratschek, V., K. Edmark, and M. Frölich (2013). The Short- and Long-term Effects of School Choice on Student Outcomes. Evidence from a School Choice Reform in Sweden. *Annales d'Économie et de Statistique* (111-112), 4.

8 Appendix

Table A7 – DID estimates of the effect of opening a new public high school on track choice - Including before/after treatment time trends

	In new HS	High school			Repetition	Dropout	Get a diploma Brevet excluded
		All tracks	General	Vocational			
<u>All types of high school</u>							
T(t >= ts)	0.373*** (0.051)	0.036** (0.017)	0.007 (0.017)	0.029* (0.015)	-0.008 (0.011)	-0.027** (0.012)	-0.003 (0.019)
Intercept	-0.047 (0.043)	0.808*** (0.021)	0.529*** (0.028)	0.279*** (0.029)	0.099*** (0.013)	0.093*** (0.015)	0.749*** (0.033)
<u>By type of high school</u>							
T(t >= ts) LGT	0.383*** (0.102)	-0.004 (0.021)	-0.021 (0.022)	0.017 (0.017)	0.004 (0.010)	-0.001 (0.019)	-0.034 (0.025)
T(t >= ts) LPO	0.427*** (0.058)	0.047** (0.019)	0.014 (0.019)	0.034** (0.016)	-0.014 (0.011)	-0.033*** (0.012)	0.012 (0.026)
T(t >= ts) LPR	0.154** (0.062)	0.037 (0.035)	0.012 (0.015)	0.026 (0.038)	-0.001 (0.022)	-0.037* (0.019)	-0.000 (0.026)
Intercept	-0.052 (0.043)	0.814*** (0.020)	0.531*** (0.028)	0.283*** (0.028)	0.098*** (0.013)	0.088*** (0.015)	0.750*** (0.033)
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
Before/after time 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 dataset, ninth-grade pupil cohorts from 2007-2008 to 2012-2013.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. A treated school is defined as the closest middle school to a new public high school. All estimations use year and middle-school fixed effects, before/after treatment time trends, and heterogenous time trends. Controls = Gender, origin, *Brevet* exam score, parents' occupation, scholarship status. Standard errors in parenthesis account for the autocorrelation of the residuals between observations on the same middle school.