*Nicolas Pistolesi**

Enrolling at university and the social influence of peers

Abstract

This article studies peer effects on the decision to enroll at university. To determine the social influence of peers, we use a measure encompassing the two major dimensions of social influence in the classroom: the ability and capacity of peers to exchange information about study options. This paper uses French administrative data on the universe of first year applicants to a single university over seven consecutive cohorts. We exploit idiosyncratic variations in the proportion of peers advised to change their educational choice. We find that our variable of interest has a small but negative and significant effect on the individual decision to attend university and observe stronger peer effects among groups of students of similar gender or socio-economic background. We also find a weaker impact of the proportion of peers advised to change their educational choice on the individuals of higher level of academic ability.

Current version: November 02, 2021
Keywords: peer effects, higher education, university enrollment
JEL codes: I21, I23, J24
Corresponding author: Nicolas Pistolesi
nicolas.pistolesi@tse-fr.eu

Universite Toulouse 1 Capitole Toulouse (Toulouse School of Economics), France

https://doi.org/10.2478/izajole-2022-0001

1 Introduction

At the completion of high school, senior students must make important decisions on their post-secondary education. These choices have long lasting effects, determining job opportunities and level of wages throughout their working career. However, high school students can sometimes make poor educational decisions (Bettinger et al., 2012; Hoxby and Avery, 2012; Oreopoulos and Dunn, 2013; Wiswall and Zafar, 2015; Hastings et al., 2016). The literature on the determinants of schooling choices has mostly concentrated on the role of expected earnings, the influence of perceived ability, or gender specific preferences.[1] Classmates may be an alternative and important determinant of the schooling choice, since students in the same class spend a large amount of time together and may share similar concerns regarding their schooling opportunities. Peers can have an impact on an individual's schooling decisions for two main reasons: first, the academic level of classmates can have a direct effect on the level of education of their peers'; and second, classmates can be a valuable source of information about future studying options. It is therefore important to study the influence of high school peers on an individual's future education choices.

In this article, we study whether or not high school students are influenced by their classmates in their decision to study at university. In the period under study, when French students apply for higher education, they receive feedback from the university on their application. This feedback is intended to assist them in selecting the institute for their post-secondary education. We use the feedback received by the high school classmates to measure the peer effects on enrollment. This measure encompasses the two channels of social influence previously mentioned. First, it is directly related to the students' academic level, as universities use high school grades to advise their applicants. Second, the feedback received represents relevant information for students to use to update their individual strategy when choosing higher education. Overall, with this measure we can assess the joint impact of both channels of social influence, among high school classmates on their decision to enroll at university.

We exploit a policy called *Orientation Active* (OA) that was implemented in France from 2009 to 2017. It provides individual-specific advice at the time of application on the choice of post-secondary education. The policy has been designed to assist high school students in choosing a field of study best suited to their abilities. Whatever the type of feedback received by students, they are permitted to register in the field of their choice. The feedback to students is determined relative to their previous schooling records and each institution is entitled to choose how to advise them. In previous work (Hestermann and Pistolesi, 2017; Pistolesi, 2017), we show the causal effect of the feedback on students' decisions to enroll at university. In this paper, we investigate whether the feedback received by high school peers has any causal impact on the individual enrollment decision. Lower-achieving applicants receive negative feedback, called "C feedback" henceforth. We investigate if the proportion of peers receiving C feedback has any impact on their individual decision to register at the same institution.

Using French administrative data on access to university, we observe the applications and the enrollment decisions of the universe of high school students applying to a given university for seven consecutive cohorts, representing around 20,000 individual applications. Our data

1 See Altonji et al. (2016) for a review of the literature on the determinants of higher education choices. Berger (1988) studies the role of expected earnings, Arcidiacono (2004) studies perceived ability, and Zafar (2013) examines gender specific preferences among others.

provide individual characteristics for all applicants to that university. It also indicates their school and class identifiers and the total number of students in each class. However, the data does not include the students that do not apply to this university. As a consequence, the group of peers is partially observed: among the peers of individuals applying to this university, we observe only those applying to the same institution. To investigate the consequences of this important data characteristic, we test whether our results change between the classes that are fully observed and those that are not. As detailed below, the results remain unchanged between these two groups. Finally, our data is representative of the population of high school students applying to university. From these data, we observe that students in a class with a larger share of peers receiving C feedback are less likely to register in the same university.

Assessing whether or not this relationship is causal is a challenging issue: indeed, the endogenous sorting of peers into high schools is particularly challenging in this setting. We use the idiosyncratic variations in the share of peers receiving C feedback across different classes, within the same high school major and during the same academic year. In France, high schools offer three different types of majors, with each major having several classes. In a high school and within a major, there is no tracking by ability between the different classes; however, there is a clear selection of students between high school majors. Using major-by-high-school-by-year fixed effects, we correct the non-random allocation of students between peer groups and control for unobserved factors that are correlated with the choice of high school, the choice of high school major, and the decision to enroll at university. Therefore, our analysis is valid only if there is a random allocation of students within major, high school, and year. We provide an empirical test and a discussion of that assumption.

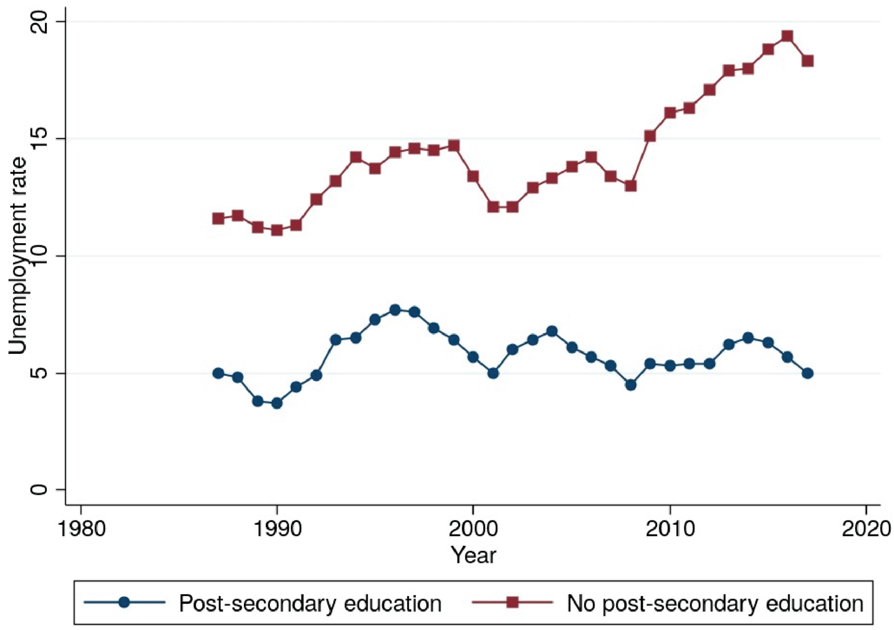
Our paper stands at the crossroads of two bodies of literature. The first studies ability peer effects. There is a large literature on this topic surveyed in Epple and Romano (2011) and Sacerdote (2011). Hoxby (2000) measures peer quality by fixed students' characteristics and exploits within-school variation in these characteristics across cohorts. Fruehwirth (2013) studies the effect of a policy that aims to assist low-achieving students to improve their effort input. Lavy et al. (2012) consider the variation of low-ability students within high school and across cohorts, and find that the proportion of academically weaker students has a negative effect on the performance of average students.[2] The second body of literature measures peer effects on higher education choices. Sacerdote (2001) uses the random allocation of students into dorms to study peer effects on various individual outcomes; however, he does not find any peer effects in the choice of college major. Arcidiacono and Nicholson (2005) consider the influence of peers in their choice of specialty among medical students and reach a similar conclusion. De Giorgi et al. (2010) collect data from Bocconi University. They use the random allocation of students between peer groups to identify peer effects and conclude that classmates have a significant influence on the choice of major. Fletcher (2012) uses data from Texas to study the magnitude of peer correlation in university preferences and choices. He finds a positive and significant effect among classmates in their choices. Overall, the literature is inconclusive regarding peer effects on the choice of major. Finally, recent papers consider the effect of peer characteristics on the choice of major, such as gender composition (Anelli and Peri, 2017; Zölitz and Feld, 2018;

2 Other related articles on ability peer effects are those of Hanushek et al. (2003), Zimmerman (2003), Withmore (2005), Foster (2006), Stinebrickner and Stinebrickner (2006), Vigdor and Nechybba (2007), Kang (2007), Graham (2008), Ammermueller and Pischke (2009), Carell et al. (2009), Carell and Hoekstra (2010), Lavy and Schlosser (2011), Lavy et al. (2012), Arcidiacono et al. (2012), Burke and Sass (2013), Sojourner (2013), and Patacchini et al. (2017), among others.

Brenøe and Zölitz, 2020) or the share of foreign peers (Anelli et al., 2017). They find a positive and significant effect of peers' characteristics on the individual's choice of major.

Our study complements the previously mentioned works in at least three important dimensions. First, we investigate the influence of peers on the decision to enroll at the same university. As discussed, the transition from secondary to higher education is a key determinant of labor market outcomes. In France, there is a large difference in unemployment rates between individuals with and without postsecondary education. Figure 1 displays the unemployment gap in France between these two groups from 1986 to 2017. It has held at 8.5 percentage points (5.7 *vs.* 14.3) on average and has increased over the last few years to 13.5 percentage points in 2017. Students who do not register at university can turn to vocational studies.[3] These are delivered in different institutions, have a shorter duration, and offer specializations in the technical fields. Vocational studies provide very different labor market prospects than a university education. Second, we use a unique administrative dataset containing applications and enrollment decisions for 20,000 university applicants. The data include high school class identifiers linking individual observations from the same class. Third, compared to the previously mentioned studies, our sample is composed of high school students of heterogeneous schooling abilities. They are representative of the high school students' population.[4] As noted by Stinebrickner and

Figure 1 Unemployment gap by education in France, 1987–2017.



Notes: The figure represents the unemployment rate in France for individuals with and without postsecondary education. *Source*: Observatoire des inégalités https://www.inegalites.fr/Le-taux-de-chomageselon-le-diplome-et-l-age

3 Another type of institution in France that provides higher education is called the *Grandes Ecoles*. These institutions are elite business or engineering schools that admit a very small and exclusive cohort. As acknowledged below, we are not able to distinguish between whether a student does not register at university because she is accepted at a *Grandes Ecoles* or because she chooses a vocational track. Nonetheless, these options represent very different choices to university enrollment.

4 Sacerdote (2001) uses data from Dartmouth College, an elite institution in the U.S., Arcidiacono and Nicholson (2005) examine students in medical schools, and De Giorgi et al. (2010) collect data from Bocconi University, a top European institution.

Stinebrickner (2006), students of lower academic ability or from lower socio-economic backgrounds are more likely to be influenced by their classmates in their choices of post-secondary education. In our sample, individuals represent the full range of academic ability among high school students.

Our results show that there are negative peer effects in enrollment choices. The size of the estimated coefficient is small but negative and significant. The estimated effect is larger within groups of similar observable characteristics, such as gender, or by social origin. The difference between the groups is not always significant. We find that individuals in the top quartile of the ability distribution are less impacted by their peers than the rest of the students. We test whether the results vary with the number of observed peers or whether the full group of peers is observed; however, we find no differences in the results between the groups.

The paper is organized as follows. In Section 2, we present the institutional background and our data. In Section 3, we detail the empirical strategy. Section 4 provides the results, and Section 5 concludes.

2 Background and Data

2.1 Institutional context

In France, from the second year of high school, students must choose a subject major. Each major gives a different weight to the variety of topics on offer for the *Baccalauréat*, the high school exit exam held at the end of the third year. High schools propose three possible subject majors: science (S), languages and literature (L), and economics and social sciences (ES). Within each major, students are divided into different classes, and remain with the same classmates throughout the academic year. There is no tracking by ability between the different classes of the same major within a high school.

Each year in April, senior high school students must select their higher education path using a website called Application Post Bac (APB).[5] On that platform, they can list up to 24 institutions and fields of study in higher education that they are considering for the following academic year. In July of each year, if the student has graduated from *Baccalauréat*, she is allowed to enroll in university and into the subject field of choice from those listed on the APB selection process.

In 2009, a new policy, called Orientation Active (OA), was enacted with the objective of helping students to select their higher education pathway. For any choice the student has selected on APB, the university must deliver individual feedback on the student's likelihood of graduating from that field of study. This feedback is determined by the student's previous academic records, and is one of the three possible responses: first, a positive response, (henceforth called "A"), stating that the student is encouraged to register in that institution, and is likely to graduate, given their past academic records; and, second, a "B" or neutral response states that the student should succeed but will have to work conscientiously in order to graduate; and third, a negative response (henceforth called "C"), which states that based on the student's academic records, it would be advisable to amend the choice of postsecondary education.

5 The detailed procedure to gain access to higher education in France is described in Appendix A.I. In 2018, the web site has changed to *Parcoursup*, and the procedure has changed slightly. Our data run from 2009 to 2015.

Feedback is sent out in May prior to the students graduating from high school. Whatever the feedback received, students are able to register in their choice of university and in their field of choice. In previous work (Hestermann and Pistolesi, 2017; Pistolesi, 2017), we demonstrate the impact of this policy on the students' choices of higher education: the students receiving C feedback are less likely to register and this relationship is causal. In this article, we investigate whether the share of classmates receiving C feedback from their local university has an impact on the individual probability of enrolling in the same university.

2.2 APB data

We use administrative records collected by the APB website on all applicants at a large French university for seven consecutive cohorts from 2009 to 2015. The data contains individual characteristics, such as a unique student identifier, gender, age, place of residence, nationality, and whether or not she received a needs-based scholarship.[6] The data also includes individual schooling records from the last 2 years: numerical grades by topic, high school major, and class identifier. This information is provided by the school. Moreover, the data indicates the name of the high school attended, a unique national high school identifier, and whether it is a public or private institution. The data also lists the feedback each student has received from the university listed on APB. We match these data with the university's internal records in order to determine those students enrolled at a particular institution in the following academic year as a first year undergraduate.

The university that provided the data used in this paper is a large research public institution. Around 25,000 individuals study in this university in the fields of economics, law, management, and computer sciences. At the graduate level, it attracts students from all over the country and from other European and non-European countries. It offers professional and research Master degrees in each of these fields, as well as PhD programs. At the undergraduate level, most students are French and come from within the local district (as is the case for most other French universities).[7] This university is not in the Paris area and there are no other universities offering these fields of study in the same region. In France, there are 13 regions, and there are 5,800,000 inhabitants in this particular region; therefore, it faces no competition in attracting local students. In France, there is limited competition among universities to attract undergraduate students, since, at that level, most of the programs are similar. Moreover, families are used to sending their children to study at the closest university from their place of residence at the undergraduate level. At the graduate level, however, competition between universities is more intense with each university attempting to differentiate their offerings within each of their different graduate programs. The university providing the data used in this paper is clearly identified at the graduate level, as a research intensive university.

The data sample is restricted to high school students living in the same region as the university providing the data. Those living abroad face very different costs and opportunities. We drop observations from vocational high schools.[8] As a further restriction, we concentrate on students from high school majors with at least two different classes, as some of our

6 A proxy for coming from a low socio-economic background.

7 As shown in the descriptive statistics below 92% of the applicants in a first-year undergraduate programme of that university are French.

8 These students turn to vocational studies. These are short programs having a duration of 2–3 years. Moreover, vocational studies take place in institutions other than universities.

specifications are based on within high school major variation in the proportion of classmates gaining access to university. The data only includes university applicants, and does not include high school students who do not apply to this university. Therefore, our sample is not representative of the population of high school students, since unobserved factors can affect the probability of applying to university and the share of peers receiving C feedback. The sample is representative of the population of students from a general high school applying to university; however, the difference between the two is limited.

One can wonder whether observing only the university applicants affects the internal and external validity of our study. First, as regards external validity (as explained above) the vast majority of students in France choose to apply to the university within their region of residence. By law, when applying to their local university, students are prioritized for enrollment. Moreover, universities are the single type of higher education institutions that are non-selective.[9] Therefore, high school students have a strong incentive to list their local university among their 24 choices on APB in order to ensure their admittance to higher education. Second, as regards internal validity, in Section 4.2, we test whether the estimation results change between classes in which everyone applies to university and classes in which there are more class members than university applicants.[10] The test does not reject the null hypothesis of similar peer effects between the classes that are fully observed and the classes that are partially observed.

Table 1 provides descriptive statistics for our sample, which includes 19,181 individual observations. Panel A shows individual observable characteristics: for example male students represent a smaller share (41%) of university applicants in our data; a large majority of students (92%) are French; and in high school, 14% of the students come from families with a low socio-economic background. Several measures of academic level in high school are available. First, the mean numerical score over the different topics can range from 0 (worst) to 20 (best). We observe a large variance in achievement, with low-achievers scoring 3.25, while high achievers score 19.33. Second, 32% are potentially high-achievers.[11] In our data, 44% of the candidates come from a high school major in ES, 37% from science, and 18% from language and literature. When applying to university, 27% receive C feedback from this institution. Finally, 50% register at this university.

In panel B, we compute the mean value of various class characteristics, and study their distribution across classes. We observe a median class size of 28 students. The composition of the class varies largely between the different classes: the share of male students varies from 5% to 70% of the class and the share of students from low socio-economic backgrounds varies from 0% to 50% of the class. Importantly, the mean academic level of the class is heterogeneous, since it ranges from 10.10 to 14.17. The last row indicates the distribution, of the share of peers that register in this university. In the classes in the lower part of the class distribution (P25), 42% of the students register at this university, while for classes in the higher part of the class distribution (P75), this proportion increases to 87% of the class. Overall, this panel shows that

9 Universities have capacity constraints and some fields can be oversubscribed, such as medicine, sports, or psychology. None of these programmes are offered in the university providing the data for this study. If capacity constraints are reached, by law, students who live within the same region are favored over students from a different region, making it highly unlikely that the student is not admitted.

10 As explained below, even if we do not observe the individual characteristics of non-applicants, the data indicates the total number of students in the class. Therefore, we know—for each peer group—how many students do not apply to university.

11 These students have chosen to take elective courses that require additional work, such as Latin or Greek.

Table 1 Descriptive statistics

Panel A: Individual characteristics	Mean	St. Dev.	Min.	Max.	Obs.
Age	18.04	1.03	16.69	23.18	19,181
Male	0.41	0.49			19,181
French nationality	0.92	0.25			19,181
Mean grade in h-school	11.81	2.06	3.25	19.33	19,181
Held back in h-school	0.07	0.26			19,181
Low social background	0.14	0.35			19,181
High achievers	0.32	0.46			19,181
Major in h-school: economics	0.44	0.49			19,181
Major in h-school: sciences	0.37	0.48			19,181
Major in h-school: literature	0.18	0.38			19,181
Public h-school	0.81	0.38			19,181
C feedback at univ. application	0.27	0.30			19,181
Enrolled at university	0.50	0.50			19,181

Panel B: High School Class Mean Characteristics	Quantile				
	Min	P25	P50	P75	Max
Mean class age	18.05	18.16	18.38	18.53	19.33
Class share of males	0.05	0.33	0.42	0.50	0.70
Class share of French	0.75	0.92	0.96	1.00	1.00
Class size	22	24	28	31	35
Mean class grade	10.10	11.25	11.79	12.29	14.17
Class share held back	0.00	0.00	0.05	0.09	0.38
Class share of low soc. back.	0.00	0.05	0.12	0.18	0.50
Class share high achievers	0.00	0.09	0.26	0.43	0.92
Class share enrolled at university	0.08	0.42	0.57	0.87	1.00
Class share observed	0.62	0.77	0.88	1.00	1.00

Notes: Panel A indicates the mean characteristics for the full sample, the standard deviation the minimum, maximum and number of observations of the students applying at university. Panel B reports the distribution of the class mean over the different classes.

the composition of the classes changes considerably from one class to another. Peer groups are mostly heterogeneous in this sample.

Table 2 provides descriptive statistics on the type of feedback received and the individual probability to register. Panel A displays the individual probability of attending this university by the type of feedback received: 51% of students receiving A or B feedback are likely to register,[12] which is 4 percentage points more than those receiving C feedback. The difference is statistically significant at the 5% level. Panel B presents the individual probability of enrolling at university, relative to the share of classmates receiving C feedback. When 20% of the peers receive C feedback, 69% of the class register in the same university; on the other hand, when 80% of the class receive the same negative recommendation, only 35% of students attend this institution in the next following year. There is a clear drop in enrollment with the share of peers

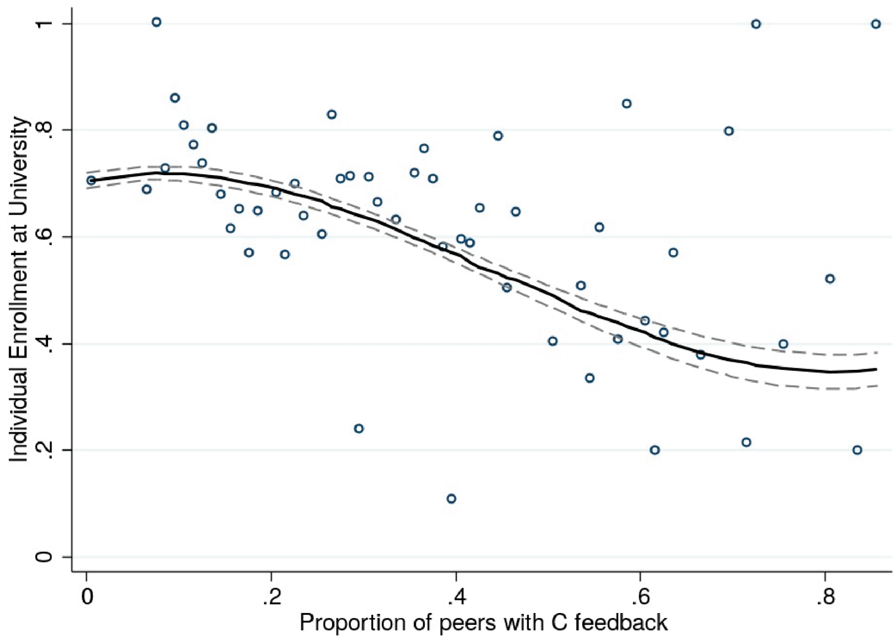
12 There is no statistically significant difference in enrolment between the group of students receiving A feedback and those receiving B feedback.

Table 2 OA policy and enrollment

Panel A: individual feedback				
	A or B Feedback		**C Feedback**	
Enrollment	0.51		0.47	
	(0.01)		(0.01)	
Panel B: peers' feedback				
Share of peers with C feedback	0.2	0.4	0.6	0.8
Enrollment	0.69	0.57	0.42	0.35
	(0.01)	(0.01)	(0.01)	(0.01)

Notes: The table indicates the proportion of students enrolling at university relative to the feedback received. Panel A relates individual enrollment and individual positive or negative feedback. Panel B relates individual enrollment and the share of peers receiving C feedback. Standard errors are in parentheses.

Figure 2 University Enrollment and Share of Peers with C Feedback.



Notes: The figure represents the relation between individual enrollment at university and the proportion of peers with C feedback. Dots are data averages over small intervals. The black continuous line represents the predicted value from a regression of the enrollment dummy variable on a polynomial function of order three of the share of peers with C feedback. The dashed grey lines represent 95% standard errors. The regression coefficients used to compute the predicted values are displayed in Table A3. The predicted value is computed from individual data and not data averages.

receiving C feedback. In the rest of the paper, we investigate whether that relationship can be given a causal interpretation conditional on major-by-high-school-by-year fixed effects.

Figure 2 graphically depicts the relationship between the share of students enrolling in this university and the proportion of classmates receiving C feedback.[13] It also presents the decreasing pattern of enrollment when the proportion of peers receiving C feedback increases.

13 To draw this figure, we regress the enrolment dummy variable on a polynomial function of order three for the share of peers receiving C feedback and we compute the predicted value. The regression results are displayed in the Appendix Table A3.

We observe a continuous drop in enrollment when >20% of the classmates receive C feedback. The next section details our empirical strategy.

3 Empirical Strategy

Endogenous sorting of peers into high schools and into high school majors is the biggest challenge faced when measuring peer effects with these data. In most countries, the self-selection of students into schools, within academic tracks and classes, is a widely observed. As discussed, there are strong significant differences in France in terms of academic ability by high school and by high-school major. To deal with this issue, we use major-by-high-school-by-year fixed effects, comparing students between classes within the same major and the same high school during the same academic year. In the next section, we present in more detail the econometric specification used in this paper. We discuss the validity of the empirical strategy in Section 3.2.

3.1 Identification of peer effects

Using repeated cross-sectional data, we estimate the following reduced-form equation, in order to explore how the feedback received by classmates in high school affect an individual's choices for post-secondary education:[14]

$$y_{smgti} = \beta_1 \overline{C}_{smgt(-i)} + \boldsymbol{X}'_{smgti}\beta_2 + \overline{\boldsymbol{X}}'_{smgt(-i)}\beta_3 + \theta_{smt} + \varepsilon_{smgti} \quad (1)$$

where sm represents school-major[15], g represents groups of peers measured by classes in high school, t is time and i denotes individuals; y_{smgti} is a dummy variable taking the value 1 if individual i in high school major sm in class g at time t enrolls in the university and 0 otherwise; $\overline{C}_{smgt(-i)}$ represents the proportion of classmates, excluding individual i, from class g at time t receiving C feedback from their university application; $\boldsymbol{X}_{smgti}$ is a vector of individual's covariates such as age, gender, nationality, and socio-economic background; $\overline{\boldsymbol{X}}_{smgt(-i)}$ is a vector of the same characteristics of the peer group, excluding individual i; θ_{smt} is a high school major-time fixed effect and ϵ_{smgti} is an error term; the coefficient of interest is β_1, which captures the effect of having a higher proportion of high school classmates receiving C feedback on the decision to enroll at the same institution.

Estimating the enrollment equation by OLS regression without the major-by-high-school-by-year fixed effects is unlikely to deliver credible peer effect estimates, since the composition of the peer group is not random. Using major-by-high-school-by-year fixed effects, we focus on the random assignment of students to classes within high school major-year and use the variation between classes of the same major, school, and year. We discuss the validity of this assumption in the next section.

14 This model is called a *linear in means model* in the literature. The theoretical properties under which we can identify the parameters have been studied in Angrist (2014) and Ushchev and Zenou (2020).

15 As explained in Section 2.1, a school-major is a collection of different classes, usually between two and four classes, within a high school following the same curriculae. In most high schools three types of majors are proposed: science, language and literature, and economics and social sciences.

3.2 Discussion of the strategy

First, we discuss the self-selection of students into peer groups.[16] To correct for the endogeneity of the formation of groups, we introduce major-by-high-school-by-year fixed effects.[17] As explained above, high schools in France offer different types of majors. Within a school and between the different majors, there is a potential selection of students, since many high-achieving students select the science major in high school, even if they do not plan to study in a scientific field at university. Using major-by-high-school-by-year fixed effects and multiple years, we compare different classes within the same major in the same school and during the same academic year. Using this approach we correct for the non-random allocation of students between the different high school majors and therefore between the different high schools, which are the drivers of self-selection at the high school level in France (Ly et al., 2014). Moreover, this empirical strategy controls for unobserved factors that are correlated with high school major and the decision to register at university.

We now present evidence of the validity of the fixed effect identification strategy. First, in Appendix AII, we check that the peers' mean variables have enough variation within high school majors, schools, and years. Table A1 presents the decomposition within and between high school major by school and by year components and shows that, on average, the within component of the peers' mean variables represents around 45% of the variance. Second, in Table 3, we test whether the change in the proportion of peers receiving C feedback within a high school major, a school, and a year is correlated with the changes in students' background characteristics. Figure 3 presents the results graphically. The first column of Table 3 displays the OLS regression coefficients. In this column, we see that several observable characteristics are correlated with the share of peers receiving C feedback: males, individuals with lower academic records, those who have been academically held back in the past, and those with a larger share of peers receiving C feedback. In column (2), we introduce major-by-high-school-by-year fixed effects and find that the coefficients are closer to zero. Only two variables are still significant: the dummy for males and the one for French nationality. In column (3), we add peer mean characteristics and find that none of the coefficients are now significant. From the results of column (3), we conclude that the variation in the share of peers receiving C feedback is exogeneous after taking out major-by-high-school-by-year and peer mean characteristics. Finally, to complement this analysis, in Appendix AII.2, we test whether or not the assignment of students between classes within a high school major, a school, and a year is randomly determined. The empirical test presented here validates this assumption.

Our major-by-high-school-by-year fixed effects strategy controls for the most obvious potential confounding factor: the endogenous sorting of students across majors and across schools based on their socio-economic background. It is still possible that unobservable factors varying across classes and within high school majors that are correlated with the proportion of peers receiving C feedback may generate some spurious correlation within the group of peers. In this respect, the major potential threat to our strategy comes from the potential difference in

16 Studies using random assignment of peers include Sacerdote (2001), Zimmerman (2003), Foster (2006) Stinebrickner and Stinebrickner (2006), Kang (2007), Carrell, Fullerton and West (2009), Duflo et al. (2011), Sojourner (2013), and Zölitz and Feld (2018) among others.

17 Studies using fixed effects include Hoxby (2000), Hanushek et al. (2003), Betts and Zau (2004), Lavy and Scholsser (2011), Lavy et al. (2012), Patacchini et al. (2017), Anelli and Peri (2017), and Brenøe and Zölitz (2020), among others.

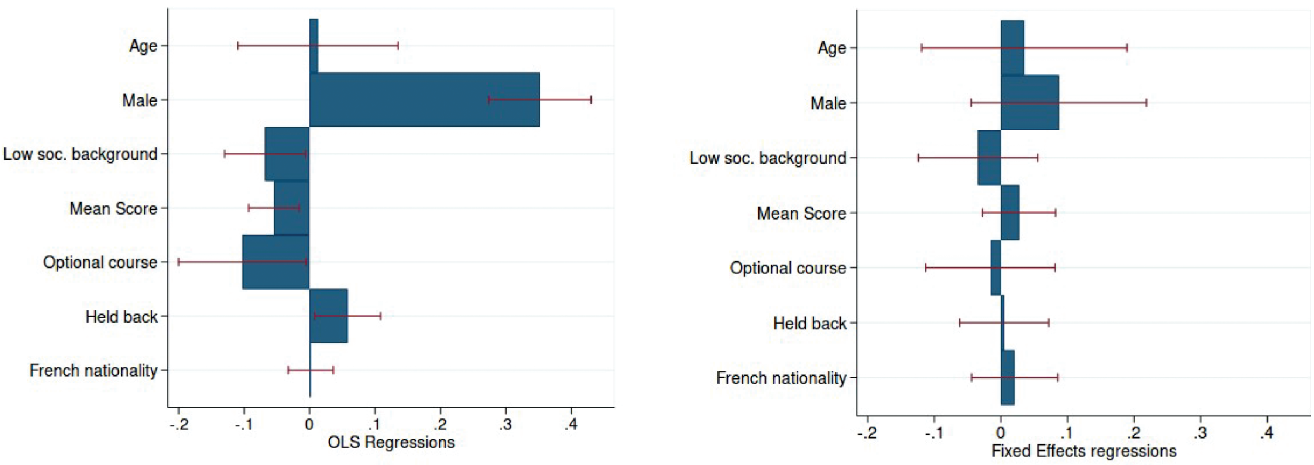
Table 3 Balancing tests for the proportion of peers with C feedback

	(1)	(2)	(3)
Dependent variable			
Age	0.013	−0.060	0.035
	(0.062)	(0.069)	(0.079)
Male	0.352***	0.157***	0.087
	(0.040)	(0.052)	(0.067)
Low soc. background	−0.068***	−0.040	−0.034
	(0.032)	(0.033)	(0.046)
Mean grade in h-school	−0.054***	0.034	0.027
	(0.020)	(0.025)	(0.028)
Optional course in h-school	−0.103***	−0.060	−0.016
	(0.050)	(0.054)	(0.050)
Held back in h-school	0.058***	0.004	0.005
	(0.026)	(0.028)	(0.034)
French nationality	0.002	0.041*	0.021
	(0.018)	(0.025)	(0.033)
Major-by-high-school-by-year fixed effects		✓	✓
Peers characteristics			✓
Observations	19,181	19,181	19,181

Notes: The table indicates OLS and major-by-high school-by-year fixed effects estimates from separate regressions of the relevant dependent variable on the proportion of peers with C feedback. Column (3) uses the following mean peers' characteristics as control variables: age, gender, a dummy for low social background, mean numerical grade, and dummies for optional classes, having been academically held back, French nationality, and public high school. Robust standard errors clustered at the high school major level are reported in parentheses.

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Figure 3 Balancing tests for the proportion of peers with C feedback.



Notes: The figure represents the estimated coefficient from the separated regressions of each individual's observable characteristics on the share of peers with C feedback. The left panel displays OLS regressions, while the right panel presents fixed-effect regressions using major-by-high-school-by-year fixed effects and controlling by the peers mean characteristics; horizontal red lines are the standard errors.

teacher quality. A class can have a particularly effective teacher, which reduces the share of the class receiving C feedback. Admittedly, our strategy does not completely prevent that threat. However, high school students are allowed to apply to university whatever their score during the high school academic year. The single requirement in applying to university is to succeed at the *Baccalauréat* exam, which is set at the national level and is assessed anonymously by various teachers, with the results unknown at the time that universities provide feedback. Nonetheless, our analysis relies on the assumption that potential class unobserved shocks are not correlated with the share of peers receiving C feedback. In the empirical analysis below, we present the results with major-by-high-school-by-year fixed effects.

4 Results

In Section 4.1, we detail the results of the regression estimations. In Section 4.2, we investigate whether the results are heterogeneous over different sub-samples. In Section 4.3, we present the falsification tests.

4.1 Peer effect on enrollment

Table 4 reports the results of regressions of individual enrollment at university on the proportion of high school classmates receiving C feedback from their application in the same institution during the same year. The dependent variable is a dummy taking the value 1 if the individual registers and 0 otherwise. These regressions are linear probability models. In any of these regressions, standard errors are clustered at the major-by-high-school-by-year level.[18] The table displays different specifications using the proportion of peers receiving C feedback

Table 4 Regressions of individual enrollment on peers with C feedback from university application

	Dep. Var: Probability to enroll at university					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Prop. of peers with C feedback	−0.153***	−0.094**	−0.095**	−0.055	−0.065	−0.068*
	(0.048)	(0.037)	(0.037)	(0.040)	(0.041)	(0.041)
Individual characteristics		✓	✓		✓	✓
Peers characteristics			✓			✓
Major-by-high-school-by-year fixed effects				✓	✓	✓
Observations	19,181	19,181	19,181	19,181	19,181	19,181
R^2	0.001	0.067	0.073	0.001	0.019	0.023
Mean dep. var.	0.507	0.507	0.507	0.507	0.507	0.507

Notes: The table indicates the regression of individual enrollment on the proportion of peers with C feedback. The table reports regressions using the share of peers with C feedback as the main explanatory variable. Individual and peer mean characteristics used as control are: age, gender, a dummy for low social background, mean numerical grade, optional classes, having been academically held back, French nationality, public high school, and year dummy variables. Robust standard errors clustered at school-major-year level are reported in parentheses.

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

18 We have attempted to cluster standard errors at the school-major-year-class level, as advised by Abadie et al. (2017), with very similar results.

as the main variable of interest. The first column shows a negative correlation of −0.153 of this variable with the probability to enroll at university. The second column adds students' individual characteristics: age, gender, whether the individual is French, whether she attends optional courses at high school (such as Latin or Greek), whether she receives a high school scholarship (a proxy for coming from a low socio-economic background), and a dummy for whether she has been academically held back in the past. It also includes an indicator variable if the high school is public and 6-year dummy variables. The coefficient drops slightly in absolute value to −0.094. In the third column, we add the mean of the peers' same observables. The estimated coefficient remains significant and similar at −0.095. In columns (4)–(6), we repeat the same exercise adding major-by-high-school-by-year fixed effects. The peer effect estimates are identified by comparing different classes within the same major, within the same school, and during the same year. In column (6), the coefficient drops in absolute value to −0.068 and is significant at the 10% level. The drop in the peer coefficient when using major-by-high-school-by-year fixed effects can be interpreted as a consequence of the non-random allocation of students between the different high school majors.

From the regression in column (6), we document that an increase of 10 percentage points of the share of peers receiving C feedback is associated with a smaller probability to register at university of around 0.68 percentage points, indicating that there is a negative peer effect of being in a group of academically weaker students on the choice of gaining access to university. While the effect may seem modest, the average probability to register to university is 0.5, and therefore we need alternative ways to measure its importance. Alternatively, we compute standardized coefficients: the effect of a one standard deviation change in the share of peers receiving C feedback. These regressions indicate that an increase by a one standard deviation of the share of peers receiving C feedback results in a decrease in individual registration of around −0.01 in standard deviation unit, compared to −0.02 with the OLS coefficients with the fixed effects.[19] Admittedly, the effect is small from a policy perspective.

It is difficult to compare our estimates to the literature, since we are not aware of any study of the effect of academically low achieving peers on the choice to enroll at university, as previously mentioned. Most of the literature on peer effects in higher education estimates peer effects in the choice of field of study, conditional on enrollment. In the next section, we test whether the previous estimates are homogeneous over the population.

4.2 **Heterogeneous effects**

Heterogeneous effects of the proportion of peers receiving C feedback are explored in Table 5. In this table, we report the results of OLS regressions of individual enrollment on the share of peers receiving C feedback for different subsamples. In panel A, we study whether there are heterogeneous effects by gender. Columns (1) and (2) show the results for male students, columns (3) and (4) display results for female students, and column (5) presents results for male and female students together. We provide mean peer effects' estimates as well as estimates distinguishing female and male peers.

19 The standard deviation of the share of peer receiving C feedback is 0.1059, and the standard deviation of the individual registration rate is 0.499. Therefore, the standardized coefficient stands at $-0.068*0.1059/0.499 = -0.014$.

Table 5 Heterogeneous effects: Regressions by gender and by social background

	Panel A : Dep. Var: Prob. to Enroll				
	Male		**Female**		**All**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Prop. of peers with C feedback	−0.122*		−0.034		
	(0.070)		(0.063)		
Prop. of male peers with C feedback		−0.110**		−0.013	
		(0.054)		(0.035)	
Prop. of female peers with C feedback		−0.032		−0.035	
		(0.035)		(0.045)	
Prop. of same gender peers with C feedback					−0.044*
					(0.024)
Prop. of other gender peers with C feedback					−0.028
					(0.024)
Individual Characteristics	✓	✓	✓	✓	✓
Peers Characteristics	✓	✓	✓	✓	✓
Major-by-high-school-by-year fixed effects	✓	✓	✓	✓	✓
Observations	7,891	7,891	11,290	11,290	19,181
R^2	0.026	0.026	0.022	0.022	0.022
Mean dep. var.	0.509	0.509	0.506	0.506	0.507
	Panel B: Dep. Var.: Prob. to Enroll				
	Low social background		**High social background**		**All**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Prop of peers with C feedback	−0.105*		−0.027		
	(0.062)		(0.050)		
Prop of low soc. back. peers with C feedback		−0.053*	0.021		
		(0.029)	(0.021)		
Prop of high soc. back. peers with C feedback		−0.053	−0.019		
		(0.044)	(0.039)		
Prop of same soc. back. peers with C feedback				−0.034	
				(0.022)	
Prop of other soc. back. peers with C feedback				0.002	
				(0.018)	
Individual characteristics	✓	✓	✓	✓	✓
Peers characteristics	✓	✓	✓	✓	✓
Major-by-high-school-by-year fixed effects	✓	✓	✓	✓	✓
Observations	5,068	5,068	14,113	14,113	19,181
R^2	0.027	0.027	0.018	0.018	0.022
Mean dep. var.	0.535	0.535	0.497	0.497	0.507

Notes: The table reports the regressions of individual enrollment on the share of peers with a C feedback by gender and social background. Panel (A) restricts the sample by gender, while panel B splits the sample by social background. Individual and peers' mean characteristics used as control are: age, gender (panel B only), scholarship in h-school (panel A only), mean numerical grade, optional classes, having been academically held back, and public high school. Major-by-high-school-by-year fixed effects are included as controls. Robust standard errors clustered at school-major-year level are reported in parentheses.

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

From these regressions, we see that mean peer effects are larger for males than for females (−0.122 *vs.* −0.034); the estimated coefficient for males is significant, while the one for females is not. The difference between the two is not statistically significant.[20] In columns (2) and (4), we split the peers into two groups: first, the male classmates and second, the female classmates, and we compute the proportion receiving C feedback for each of these two groups. We see that for the sample of males, the estimated peer effect is larger with male peers than with female peers (−0.110 *vs.* −0.032). In column (4), when we study female students, we obtain a symmetric result: peer effects by female peers are larger than peer effects by male peers, even if the coefficients are not statistically significant. In column (5), we pool the sample of males and the sample of females and define two groups of peers; the first group includes peers of the same gender as the individual, and the second group includes peers of a different gender. Only the same gender variable, estimated at −0.044, is significant at 10%, while the different gender variable is not. These results are indicative that within groups of students of the same gender, peer effects within the group are more important than those between the groups. A vast literature on ability peer effects shows that a large proportion of females in a class improves overall achievement levels (Hoxby, 2000; Whitmore, 2005; Lavy and Schlosser, 2011). Other studies conclude that a large proportion of male students negatively affects classroom outcomes (Figlio, 2007; Carrell and Hoekstra, 2010). Zölitz and Feld (2018) study gender peer effects and find that a larger proportion of female students affects the choices of education of males and females differently. Anelli and Peri (2017) find a positive effect on males but not on females. Brenøe and Zölitz (2020) study gender peer effects on the choice on STEM participation. They find a greater number of female peers widens the gender gap. While our results cannot directly be compared to these studies, since we consider a different outcome and do not measure the direct effect of the gender composition, we still find that males and females have asymmetric effects on their fellow classmates. Moreover, in addition to most of the papers outlined in this literature, we provide the estimated cross effects between the two gender groups.

In panel B, we repeat the same exercise, but split the sample by social background.[21] First, the effect of peers is relatively more important for students from families with a low socio-economic background (−0.105 *vs.* −0.027); however, the standard errors are too large for the difference to be significant[22]. Second, in columns (2) and (4) we consider the cross effects of these two groups of peers. Only the proportion of lower social background individuals receiving C feedback is significant for the students of the low socio-economic group. These results are somewhat parallel to those of Angrist and Lang (2004), who find that the presence of minority students has little effect on non-minority students' academic outcomes.

In Table 6, we study peer effects for different groups of students defined by ability level[23]. This is defined by the within-classroom percentile in the score distribution, ranging from 0 for the student with the lowest numerical score in her class to 1 for the student with the highest

20 The share of males receiving C feedback is slightly larger than the share of females receiving C feedback (0.30 *vs.* 0.23).

21 A low socio-economic background is measured by receiving a needs-based scholarship in high school. A privileged or high socio-economic background represents those who do not receive such a scholarship.

22 The share of students receiving C feedback is very close between these two groups (0.27 and 0.26).

23 We use the average numerical score over the different topics in the senior year of high school in order to measure ability. These raw scores are not standardized across schools. The raw scores are difficult to interpret since they depend on the individual school's grading policy. We build a measure of ability relative to the group of peers, computing each individual's rank within the score distribution of her class. Then, we translate these ranks into within-classroom percentiles.

Table 6 Heterogeneous effects: Regressions by ability level

	Dep Var: Prob. to Enroll at University				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Prop. of peers with C feedback	−0.068*	−0.071*	−0.071*	−0.071*	−0.092**
	(0.041)	(0.041)	(0.043)	(0.041)	(0.043)
Top 50% in ability distribution		0.010	0.010		
		(0.009)	(0.009)		
Top 50% × Prop. of peers with C feedback			0.001		
			(0.121)		
Bottom 25% in ability distribution				−0.081***	−0.090***
				(0.009)	(0.011)
Top 25% in ability distribution				−0.079***	−0.082***
				(0.009)	(0.009)
Bottom 25% × Prop. of peers with C feedback					0.023
					(0.014)
Top 25% × Prop. of peers with C feedback					0.048*
					(0.026)
Individual characteristics	✓	✓	✓	✓	✓
Peers characteristics	✓	✓	✓	✓	✓
Major-by-high-school-by-year fixed effects	✓	✓	✓	✓	✓
N	19,181	19,181	19,181	19,181	19,181
R^2	0.023	0.023	0.023	0.030	0.030
Mean dep. var.	0.507	0.507	0.507	0.507	0.507
P-value F-test					0.033

Notes: The table indicates the results of regressions of individual enrollment on the proportion of peers with C feedback. Top 50%: dummy variable taking the value 1 if the students' score belongs to the top 50% of the score distribution in her class and 0 otherwise. P-value F-test: test of the null hypothesis that the two interaction terms are equal to zero. Individual and peers mean characteristics used as control are: age, gender, scholarship, optional classes, having been academically held back, French nationality, and public high school. Major-by-high-school-by-year fixed effects are included as controls. Robust standard errors clustered at school-major-year level are reported in parentheses.

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

score in her peer group. In column (2), we add a dummy variable taking the value 1 if the individual's score is above the 50 percentile of the ability distribution. The results do not change. In column (3), we add the interaction between the dummy and the proportion of peers receiving C feedback to study whether the effect of the peer variable changes between individuals above or below the median of the ability distribution. The coefficient is not significant and close to zero.

Finally in column (5), we interact the dummy variables for the first and last quartile of the ability distribution with the proportion of peers receiving C feedback. The coefficient of the interaction is positive and significant for the high academic ability students. For this group, the negative effect of having academically low performing peers represents two-thirds of the effect estimated for the overall sample: these results, therefore, provide some insight into how social interactions impact individual enrollment. A larger fraction of academically low performing peers is less detrimental to the students that are in the higher quartile of the ability distribution. A potential explanation could be that those receiving C feedback interact less with their academically high achieving peers.

Table 7 Heterogeneous effects: Regressions by number of observed peers

	Dep Var: Prob. to Enroll at University			
	(1)	**(2)**	**(3)**	**(4)**
Prop. of peers with C feedback	−0.068*	−0.080	−0.061	−0.075*
	(0.041)	(0.051)	(0.042)	(0.043)
Nb of obs. peers above the median		−0.002		
		(0.013)		
Prop. of obs. peers with C feedback × Above median		0.044		
		(0.081)		
Total class observed			−0.021	
			(0.022)	
Prop. of obs. peers with C feedback × Total Class			−0.057	
			(0.188)	
Prop. obs. peers >80%				−0.014
				(0.016)
Prop. of obs. peers with C feedback × Observed peers >80%				0.087
				(0.051)
Individual characteristics	✓	✓	✓	✓
Peers characteristics	✓	✓	✓	✓
Major-by-high-school fixed effects	✓	✓	✓	✓
N	19,181	19,181	19,181	19,181
R^2	0.023	0.023	0.023	0.023
Mean dep. var.	0.507	0.507	0.507	0.507

Notes: The table indicates the results of regressions of individual enrollment on the share of peers receiving C feedback. Nb of observed peers above median: dummy variable taking the value 1 if the number of observed peers in the class is above the median number of observed peers in the data and 0 otherwise. Total Class observed: Dummy variable taking the value 1 if the number of observed peers is equal to the total number of peers in the class. Prop. obs. peers >80%: More than 80% of the class peers are observed. Individual and peers mean characteristics used as control are: age, gender, scholarship, optional classes, having been academically held back, French nationality, and public high school. Major-by-high-school-by-year fixed effects are included as controls. Robust standard errors clustered at school-major-year level are reported in parentheses.

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

In Table 7, we investigate whether the peer effects' estimates change with the number of observed peers. The first column presents our main result from Table 4 for comparison. In column (2), we add a dummy variable taking the value 1 if the number of observed peers in the class of individual *i* is above the median number of observed peers in the data and 0 otherwise, and we interact this dummy variable with the proportion of peers receiving C feedback. The interaction coefficient measures if there are differences in the peer effect estimate between the high and low number of observed peers: results show that it is positive but not significant. From this regression, we conclude that there are no differences in the peer effect estimate, irrespective of whether the group of observed peers is small or large.

So far, the peer groups have been defined in relation to university applicants only. In this setting, the group of peers is partially observed when some students in the class do not apply to this university; however, it is fully observed when every student in the class does so. To test if the peer effects estimate changes between these two groups, we use the information provided in the data for each applicant on the total number of students in the class. We define a dummy variable taking the value 1 if the number of observed peers plus one is equal to the total number of students in the class[24] and 0 otherwise, and we interact the dummy with the proportion of peers receiving C feedback. In column (3), the estimated coefficient of the interaction term is negative at −0.057 and not significant. We conclude that there are no differences in the effect of the proportion of peers receiving C feedback between the groups of peers that are fully observed and the groups of peers that are partially observed.

4.3 Falsification tests

In Table 8, we perform some falsification tests, changing the proportion of students that register at university in cohort t by the proportion of enrolled peers from the same school and the same major in younger cohorts: the individuals of the cohort $(t + 1)$ in columns (1) and (2), and those of the cohort $(t + 2)$ in columns (3) and (4).

The results of these four regressions point to no clear relationship between the adjacent cohorts and the university enrollment choices of the current cohort. None of the four estimated coefficients is significant at the usual levels. We interpret the estimates from these regressions as evidence that the results from Tables 4–7 do not come from a correlation between the share of students enrolling at university and time-varying high school major unobserved characteristics.

Table 8 Placebo regressions

Dep. Var: Probability to Enroll at University				
	Peer mean computed in			
	year t+1 (1)	**year t+1 (2)**	**year t+2 (3)**	**year t+2 (4)**
Prop. of peers with C feedback	−0.029	−0.048	−0.054	−0.056
	(0.048)	(0.051)	(0.049)	(0.053)
Individual characteristics		✓		✓
Peers characteristics		✓		✓
Major-by-high-school-by-year fixed effects	✓	✓	✓	✓
	14,402	14,402	13,544	13,544
Mean dep. var.	0.508	0.508	0.505	0.505

Notes: The table presents the results of regressions of individual enrollment on the proportion of peers with C feedback for different samples of peers. Columns (1) and (2) use the share of peers of students in the same high school and the same high school major as individual *i* one year after individual *i*. Columns (3) and (4) repeat the same exercise, using the peers from year t+2. Individual and peers mean characteristics used as control are: age, gender, scholarship, mean numerical grade, optional classes, having been academically held back, French nationality, and public high school. Major-by-high-school-by-year fixed effects are included as controls. Robust standard errors clustered at school-major-year level are reported in parentheses.

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

24 We observed the full group of peers for 31% of observations.

5 Conclusion

In this paper, we empirically measure the extent of peer effects on high school students' choices to attend university. Focusing on the impact of the share of peers advised to change their educational choice, we can aggregate the two major channels of social influence in the classroom: first, the level of academic achievement of the peers; and second, the possibility of them sharing relevant information about future study options. We use a unique and rich dataset on the universe of applicants to a large university over seven consecutive cohorts. These data allow us to study students' enrollment choices and compare them to the share of their classmates that do not receive a positive reply to their application to the same university. In this respect, we develop an empirical strategy using major-by-high-school-by-year fixed effects to deal with the endogenous sorting of students into peer groups.

Our results suggest that having peers with a lower academic record has a slight but negative impact on the decision to enroll at university. An increase in the share of classmates receiving C feedback when they apply to university decreases the likelihood that an individual registers at the same institution by between 0.01 and 0.02 in standard deviation units, which is small but significant. The effect is larger within groups of individuals sharing similar observable characteristics, such as gender or socio-economic background, though the difference is not always significant. Moreover, we find that individuals in the top quartile of the academic ability distribution are less impacted by their weaker peers in their choices than those in the lower part of the ability distribution. Different mechanisms may be at play here: students may learn about their ability as in Stinebrickner and Stinebrickner (2012); they may also learn about the level of difficulty of the university program, as in Altonji et al. (2012); and furthermore, the university application feedback policy may alter the students' beliefs about their own ability (Wiswall and Zafar, 2015). It would be interesting to distinguish between the relative importances of each of these potential mechanisms; however, the available data limits this ability. These results show that when designing a policy to improve access to higher education, an individual's peer group is a factor in the selection decision of future academic or vocational choices.

Several caveats are necessary so as not to over-interpret the results of our study. First, as already discussed, the data provide information on university applicants only. Among the group of class peers, we do not observe students who decide not to apply to university. In France, 70% of high school graduates start studying at university (Inan, 2017). Therefore, the difference between university applicants and their group of peers is slight: we cannot claim that our sample is representative of the full population of high school students, but it does remain representative of the population of university applicants. Second, our empirical analysis compares classes of students in the same high school major in the same high school during the same year. It does not prevent from shocks that would be class-specific and related to the unobserved determinants of university enrollment. While we are not able to see where these class-specific shocks would come from, the results presented here rely on the assumption that common shocks do not prevail.

A large empirical literature has already documented that educational attainment is partly determined by the academic records of a student's peers. Our results complement these studies, showing that an individual's choices upon entering higher education are also subject to the academic achievement level of their peers. Our paper also complements the literature showing peer effects in the choice of subject major in higher education. Since most of the literature is focused on the intensive margin, while we focus on the extensive margin, future studies could extend this work by investigating alternative schooling choices, such as the choice between vocational or academic tracks. In this situation, the choices have long-term consequences and more research is needed to assess whether or not the results presented here hold in these alternative contexts.

Declaration

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have nocompeting interests.

Funding

I acknowledge funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010.

Authors' contributions

It is a single author paper.

Acknowledgements

I am grateful to Gigi Foster, Pauline Grosjean, Yves Zenou and audiences at TSE, UNSW, and Monash University for comments and suggestions. I appreciated comments from the editor and three anonymous referees.

References

Abadie, A.; S. Athey; G. W. Imbens; J. Wooldridge (2017): When Should you Adjust Standard Errors for Clustering? National Bureau of Economic Research.

Altonji, J. G; P. Arcidiacono; A. Maurel (2016): The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects, in: Hanushek, E.; S. Machin; L. Woessmann (eds.) Handbook of the Economics of Education, Amsterdam: Elsevier Science.

Altonji, J. G.; E. Blom; C. Meghir (2012): Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers. Annual Review of Economics 4(1), 185-223.

Ammermueller, A.; J. S. Pischke (2009): Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study. Journal of Labor Economics 27(3), 315-348.

Anelli, M.; G. Peri (2017): The Effects of High School Peers' Gender on College Major, College Performance and Income. The Economic Journal 129(618), 553-602.

Anelli, M.; K. Shih; K. Williams (2017): Foreign Peer Effects and Stem Major Choice. CESifo Working Paper Series.

Angrist, J. D. (2014): The Perils of Peer Effects. Labour Economics 30, 98-108.

Angrist, J. D.; K. Lang (2004): Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. American Economic Review 94(5), 1614-1634.

Arcidiacono, P. (2004): Ability Sorting and the Returns to College Major. Journal of Econometrics 121(1), 343-375.

Arcidiacono, P.; G. Foster; N. Goodpaster; J. Kinsler (2012): Estimating Spillovers Using Panel Data, with an Application to the Classroom. Quantitative Economics 3(3), 421-470.

Arcidiacono, P.; S. Nicholson (2005): Peer Effects in Medical School. Journal of Public Economics 89(2–3), 327-350.

Berger, M. C. (1988): Predicted Future Earnings and Choice of College Major. ILR Review 41(3), 418-429.

Bettinger, E. P.; T. Long; P. Oreopoulos; L. Sanbonmatsu (2012): The Role of Application Assistance and Information in College Decisions. Quarterly Journal of Economics 127(3), 1205-1242.

Betts, J. R.; Zau, A. (2004): Peer groups and academic achievement: Panel evidence from administrative data, Unpublished manuscript.

Brenøe, A. A.; U. Zölitz (2020): Exposure to More Female Peers Widens the Gender Gap in Stem Participation. Journal of Labor Economics 38(4), 1009-1054.

Burke, M. A.; T. R. Sass (2013): Classroom Peer Effects and Student Achievement. Journal of Labor Economics 31(1), 51-82.

Carrell, S. E.; R. L. Fullerton; J. E. West (2009): Does Your Cohort Matter? Measuring Peer Effects in College Achievement. Journal of Labor Economics 27(3), 439-464.

Carrell, S. E.; M. L. Hoekstra (2010): Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. American Economic Journal 2(1), 211-228.

De Giorgi, G.; M. Pellizzari; S. Redaelli (2010): Identification of Social Interactions Through Partially Overlapping Peer Groups. American Economic Journal 2(2), 241-275.

Duflo, E.; P. Dupas; M. Kremer (2011): Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. American Economic Review 101(5), 1739-1774.

Epple, D.; R. E. Romano (2011): Peer Effects in Education: A Survey of the Theory and Evidence. Handbook of social economics 1, 1053-1163.

Feld, J.; U. Zölitz (2017): Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects. Journal of Labor Economics 35(2), 387-428.

Figlio, D. N. (2007): Boys Named Sue: Disruptive Children and their Peers. Education Finance and Policy 2(4), 376-394.

Fletcher, J. M. (2012): Similarity in Peer College Preferences: New Evidence from Texas. Social Science Research 41(2), 321-330.

Foster, G. (2006): It's not your Peers, and it's not your Friends: Some Progress Toward Understanding the Educational Peer Effect Mechanism. Journal of Public Economics 90(8–9), 1455-1475.

Fruehwirth, J. C. (2013): Identifying Peer Achievement Spillovers: Implications for Desegregation and the Achievement Gap. Quantitative Economics 4(1), 85-124.

Graham, B. S. (2008): Identifying Social Interactions through Conditional Variance Restrictions. Econometrica 76(3), 643-660.

Hanushek, E. A.; J. F. Kain; J. M. Markman; S. G. Rivkin (2003): Does Peer Ability Affect Student Achievement? Journal of Applied Econometrics 18(5), 527-544.

Hastings, J. S.; C. A. Neilson; A. Ramirez; S. D. Zimmerman (2016): (Un) informed College and Major Choice: Evidence from Linked Survey and Administrative Data. Economics of Education Review 51, 136-151.

Hestermann, N.; N. Pistolesi (2017): Does the Provision of Information on their Skills Affect Students' Enrollment Choices? Economie et Prevision (2), 177-193.

Hoxby, C. M (2000): Peer Effects in the Classroom: Learning from Gender and Race Variation. NBER Working Paper 7867.

Hoxby, C. M.; C. Avery (2012): The Missing "One-offs": The Hidden Supply of High-Achieving, Low Income Students. National Bureau of Economic Research 18586.

Inan, C. (2017): Vœux d'Orientation et Propositions Exprim´es Sur Le Portail Admission Post-bac (apb). In: Etat de l'Enseignement sup´erieur et de la Recherche´ en France - 49 indicateurs. Paris: Minist`ere de l'Education nationale, de´ l'Enseignement sup´erieur et de la Recherche.

Kang, C. (2007): Classroom Peer Effects and Academic Achievement: Quasirandomization Evidence from South Korea. Journal of Urban Economics 61(3), 458-495.

Lavy, V.; M. D. Paserman; A. Schlosser (2012): Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. The Economic Journal 122(559), 208-237.

Lavy, V.; A. Schlosser (2011): Mechanisms and Impacts of Gender Peer Effects at School. American Economic Journal 3(2), 1-33.

Lavy, V.; O. Silva; F. Weinhardt (2012): The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools. Journal of Labor Economics 30(2), 367-414.

Ly, S. T.; E. Maurin; A. Riegert (2014): La mixit´e Sociale et Scolaire en Ile-de-France: Le rˆole des ´Etablissements. Institut des Politiques Publiques, 4.

MSER (2012): L'Etat de l'Enseignement Sup´erieur et de la Recherche en France. Minist`ere de l'Enseignement Sup´erieur et de la Recherche, n 6.

Oreopoulos, P.; R. Dunn (2013): Information and College Access: Evidence from a Randomized Field Experiment. Scandinavian Journal of Economics 1(01), 3-26.

Patacchini, E.; E. Rainone; Y. Zenou (2017): Heterogeneous Peer Effects in Education. Journal of Economic Behavior & Organization 134, 190-227.

Pistolesi, N. (2017): Advising Students on their Field of Study: Evidence from a French University Reform. Labour Economics 44, 106-121.

Sacerdote, B. (2001): Peer Effects with Random Assignment: Results for Dartmouth Roommates. The Quarterly Journal of Economics 116(2), 681-704.

Sacerdote, B. (2011): Peer Effects in Education: How might they Work, how Big are they and how much do we know thus far? Handbook of the Economics of Education 3, 249-277.

Sojourner, A. (2013): Identification of Peer Effects with Missing Peer Data: Evidence from Project Star. The Economic Journal 123(569), 574-605.

Stinebrickner, R.; T. R. Stinebrickner (2006): What can be Learned About Peer Effects Using College Roommates? Evidence from New Survey Data and Students from Disadvantaged Backgrounds. Journal of Public Economics 90(8–9), 1435-1454.

Stinebrickner, T.; R. Stinebrickner (2012): Learning About Academic Ability and the College Dropout Decision. Journal of Labor Economics 30(4), 707-748.

Ushchev, P.; Y. Zenou (2020): Social Norms in Networks. Journal of Economic Theory 185, 104969.

Vigdor, J.; T. Nechyba (2007): Peer Effects in North Carolina Public Schools. Schools and the Equal Opportunity Problem 73-101.

Whitmore, D. (2005): Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment. American Economic Review 95(2), 199-203.

Wiswall, M.; B. Zafar (2015): Determinants of College Major Choice: Identification Using an Information Experiment. The Review of Economic Studies 82(2), 791-824.

Zafar, B. (2013): College Major Choice and the Gender Gap. Journal of Human Resources 48(3), 545-595.

Zimmerman, D. J. (2003): Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. Review of Economics and Statistics 85(1), 9-23.

Zölitz, U.; J. Feld (2018): The Effect of Peer Gender on Major Choice. University of Zurich, Department of Economics, Working Paper (270).

Appendix

AI OA policy

In France, in order to get access to post-secondary education, students need to graduate from high school through a national exam, called *Baccalauréat*. Every year, a large number of students enter universities, since >86% of high school students succeed in this exam. Universities are not selective and they set very low tuitions fees. Once high school students have graduated from *Baccalauréat*, they are free to choose their field of study and their university. While this system does not restrict an individual's possibilities for education, choosing a field of study and an institution is a demanding task. In France, >1,800 different undergraduate programs are proposed at university level (MSER (2012)). A reform called Orientation Active (OA), was implemented in 2009 in order to guide students into the various higher education tracks. The OA policy requests universities to advise high school students on their choice of field of study.

This policy is implemented through a website called APB. During the spring term, high school students can list up to 24 universities and fields of study they consider registering at in the following academic year. Before choosing their postsecondary education, students receive feedback from each university on each choice they have listed. The recommendations are designed to assess each student's abilities to graduate in this field. They can be of three types: positive, neutral, or negative, and are based on each individual's previous academic record and produced by university officers. On the APB website, students have to rank their different choices. In July, once the students have graduated from high school, they are entitled to register in the field and university of their choice. The recommendations are in no case binding. The students receiving negative feedback can take this into account and change their field of study, or they can choose to ignore the feedback and register anyway into their preferred field and university.

AII Validity of the Identification Assumption

AII.1 Decomposition of the variance in peer group-level means

The decomposition of the variance of the class mean characteristics in a within and between high school majors components is given by the following formula (Ammermueller and Pischke, 2009):

$$\frac{1}{G}\sum_{f=1}^{F}\sum_{g=1}^{G_f}\left(x_{G_f}-\overline{x}\right)^2=\frac{1}{G}\sum_{f=1}^{F}\sum_{g=1}^{G_f}\left(x_{G_f}-\overline{x}_f\right)^2+\frac{1}{G}\sum_{f=1}^{F}G_f\left(\overline{x}_f-\overline{x}\right)^2 \tag{1}$$

with G denoting the total number of classes, f the high school major index, G_f the total number of classes in high school major f, x_{Gf} the mean peer characteristics in class G in high school major f, $\overline{x}$ the overall mean, and x_f the mean in high school major f. We measure the share of the variance of the mean peer characteristics within and between high school majors. The first

Table A1 Decomposition of the variance in peer group-level means

	Total peer mean Variance	Within h-school Majors (%)	Between h-school Majors (%)
Enrollment	0.09	46.48	53.52
C feedback	0.02	62.66	37.34
Age	0.76	8.82	91.18
Male	0.06	65.89	34.11
Low soc. background	0.03	60.21	39.79
Mean score in h-school	1.26	55.60	44.40
Elective class in h-school	0.09	47.06	52.94
Held-back in h-school	0.02	61.31	38.69
French nationality	0.04	17.86	82.14

Notes: The table indicates the decomposition of the variance of the class mean within and between high school majors. Column (1) provides the total variance, column (2) shows the within high school major share of the variance, and column (3) displays the between high school major share of the variance.

term on the right hand side of equation AII.1 measures the within-major component, while the second term on the right represents the between high school major one. Our analysis using high school major fixed effects is valid only if a sufficiently large share of this variance takes place between classes and within high school majors.

Table A1 provides the results of this decomposition. For most mean peers' characteristics, the variance is fairly divided in the within and between high school major components. The between component is around 55% of the total peer-means variance, while the within component represents around 45%; however, the variance in mean age is mostly at the between component level. Overall, using high school major fixed effects preserves an important share of the variance in the peers' mean characteristics.

AII.2 Randomization Checks

Our fixed effect approach assumes that the assignment of students between classes within a high school major is randomly determined. To assess the validity of this assumption, we test whether class dummies jointly predict students' pre-treatment characteristics. The pre-treatment characteristics we consider are age, gender, low social background, attending optional courses, having been academically held back, and French nationality. For each high school major in our sample, we run a regression of pre-treatment characteristics on class dummies and we F-test for joint significance of the class dummies. Thus, for each pre-treatment characteristic, we run 1,100 regressions.[25] As discussed in Feld and Zölitz (2017), under conditional random assignment, the P-values of the F-tests of these regressions should be uniformly distributed with a mean of 0.5.[26] Moreover, if students are as randomly assigned to classes within high school majors, the F-test should reject the null hypothesis of no relationship between class assignment and students' pre-treatment characteristics at the 10% level in 10% of the cases and at the 5% level in 5% of the cases.

25 The total number of high-school major combination in our data is 1,100.

26 See Murdoch et al. (2008) for more details.

The results of this test are displayed in Table A2. In this table, we see that for most of the characteristics, the proportion of rejection is close to the level of the P-value from 10% to 1%. Moreover, for each characteristic, the mean P-values are close to 0.5. Overall, these results point to a random allocation of students to classes within high school majors.

Table A2 Randomization check of high school class assignment

	(1)	(2)	(3)	(4)
	Percent Significant at the			Mean P-value
	10% level	5% level	1% level	
Age	0.1	0.06	0.02	0.42
Male	0.12	0.06	0.02	0.47
Low social background	0.08	0.04	0.01	0.48
Optional courses in h-school	0.11	0.05	0.02	0.45
Held back in h-school	0.06	0.04	0.01	0.47
French Nationality	0.07	0.03	0.01	0.44

Notes: The table is based on separate OLS regressions with age, gender, low social background, optional courses, having been academically held back, and French nationality as dependent variables. The explanatory variables are a set of class dummies. We run one regression for each high school major. Columns (1)–(3) show the proportion of F-tests rejected the null hypothesis at the 10%, 5%, and 1%, respectively. Column (4) displays the mean P-value over the different regressions.

Table A3 Regression of individual enrollment on the share of peers with a C feedback

Dep. Var: Probability to Enroll at University	
	(1)
Prop. of peers with C feedback	0.350
	(0.104)
Prop.^{2}	−2.513
	(0.325)
Prop.^{3}	1.892
	(0.235)
Observations	19,181
R-squared	0.024

Notes: The table indicates the regression of individual enrollment on the share of classmates receiving C feedback from their university application. The regression is used to compute the predicted value depicted in Figure 2.