Persistent Classmates: How Familiarity with Peers Protects from Disruptive School Transitions*

Son Thierry Ly† Arnaud Riegert‡

Version 2, April 2014

Abstract

By understanding peer effects within classrooms, guideful recommendations could be drawn to optimize their composition and improve school achievement at no cost. To overcome the issue of endogenous selection across classes, we exploit rare natural experiment settings in which freshmen are credibly randomly allocated to classes within high schools. By examining the effect of several classrooms characteristics on students' achievement in freshman year and high school graduation exams, we find an important, positive effect of assignment with more *persistent classmates*, *i.e.* classmates who were already in the freshman's class before high school. We provide strong evidence that this result derives from the benefit of *familiarity* with peers, rather than from some unobserved ability characteristics of these classmates. In particular, we show that the estimates are driven mainly by low-achieving students likely to experience a difficult transition to high school. These students at risk of underachievement perform way better when assigned to their low-achieving former classmates in particular. The magnitude of the estimates suggests that grouping low-achieving freshmen who know each other could decrease their current retention rate by around 20 percent, and raise their graduation rate by the same amount.

JEL codes: I21, I28, Z13

Keywords: Friendships, Social Networks, High schools, Class composition, Peer effects.

---

*We thank Éric Maurin, Julie Berry Cullen, Luc Behaghel, Gordon Dahl, Thomas Piketty, Gwenaël Roudaut, Camille Terrier and Margaux Vinez for their helpful comments and suggestions on earlier drafts of this paper. We thank participants at IWAEE conference (Catanzaro 2013), PSE Applied Lunch seminar and CREST internal seminar. We are also grateful to the statistical services at the French Ministry for Education (DEPP) and in particular Cédric Afsa who facilitated our access to the datasets. This research was supported by a grant from the CEPREMAP research center.

†Paris School of Economics / École normale supérieure. son.thierry.ly@ens.fr

‡Paris School of Economics / Insee / EHESS. arnaud.riegert@ens.fr

Introduction

If peer effects exist, people are unlikely to internalize them perfectly when choosing their environment, *e.g.* a neighborhood or a school. Thus, addressing the way individuals are allocated across contexts could be Pareto-efficient. This potential for welfare-improving policies has sustained the strong interest of economists in peer effects, despite the huge empirical challenges raised by endogenous sorting. For example, a great deal of papers investigate the role of neighborhood (Goux & Maurin, 2007; Kling *et al.*, 2007) or school composition (Hoxby, 2000; Angrist & Lang, 2004; Cullen *et al.*, 2006; Lavy & Schlosser, 2011) on students' outcomes. Surprisingly, the literature is much less extensive on the estimation of peer effects within classrooms, although most students' interactions are likely to occur at this level. It is also unfortunate in a policy perspective, because school administrators have much more latitude in setting-up classrooms than policy-makers have in influencing neighborhood or school choice. The main studies on the subject are based on experimental data either in primary schools in developing countries (Duflo *et al.*, 2011), or in college in developed countries (Carrell *et al.*, 2011). Evidence based on observational data is rather poor, since they require both rich data at the classroom level and convincing natural experiments that are rarely available.

In this paper, we exploit the institutional features of students' allocation across classrooms in the first year of high school in France. By the time they assign freshmen to their classes (which are common to almost all subjects), principals do not know them and only rely on the set of formal characteristics they observe in their registration files. Occasionally, they will encounter the case of two (or more) freshmen whose files are very similar (we call them *similar-file* or *SF students* throughout the paper). For instance, they have the same middle school of origin, gender, age, socioeconomic status (SES), list of optional courses and approximately the same scores in 9th grade. Therefore, if for any reason they decide to separate them, we argue that the choice of assigning student 1 to class X and student 2 to class Y or the other way around will be random. In other words, for freshmen coming from the same middle school to the same high school, class assignment is random conditional to a set of observable characteristics.

Using a unique administrative dataset, we are able to observe most characteristics observed by principals. Most of all, we have access to the scores obtained by students at the national anonymous exam at the end of middle school, *while principals do not*. Therefore, we are able to show that principals do not use the small remaining information we do not observe (*e.g.* students' names from which ethnicity could be inferred) to assign separated similar-file students to their classes, or at least not in correlation to potential achievement. Within groups of separated SF students, statistical correlations between achievement and classroom characteristics are thus not attributable to differences in unobserved individual characteristics.

We use this strategy to compare the estimated effect of several dimension of the classroom environment. Common measures of peer characteristics are considered, such as peer ability, gender or socio-economic status (SES). But surprisingly, the most robust effect we find comes from the number of *persistent classmates* (PC) a student gets, *i.e.* classmates who were already in the freshman's class in the last year of middle school. Not only does the number of PC reduce significantly the risk of retention in freshman year, but by contrast with the other measures of peer characteristics, the effect also persists in the long run and is associated with differences in graduation rates at the end of high school.

The second part of this paper is aimed to understand why the presence of these persistent classmates generate positive spillovers. Although the number of PC could capture some omitted classroom characteristics linked for instance to peer ability, our investigations suggest that students benefit from having more persistent classmates only because of a *familiarity* effect, *i.e.* because they know each other well. Three findings drive us toward this conclusion. Firstly, the estimates are extremely robust to the inclusion or not of controls for the other classmates' characteristics (ability, gender and SES). Secondly, we find that the PC effect is very heterogenous and mainly driven by low-achieving, low-SES students, especially when they are suddenly more exposed to high-SES students. This is consistent with our interpretion, as being surrounded by familiar faces should matter more when the transition to high school is highly disruptive. Lastly, these students at risk of underachievement in high school are much less impacted by their high- than their low-achieving persistent classmates, which could hardly be explained by differences in their unobserved ability. Robustness checks to our main results are provided.

The main academic contribution of this study is to shed light on the ongoing debate over the complexity of peer effects. While some recent studies offer an insight in the role of social networks during school transitions (Lavy & Sand, 2012, see *e.g.*), we exploit natural experiments that provide a stronger identification of the impact of classmates' characteristics, including their social links. *Ex ante*, there was no theoretical reason to expect classmates persistency to have positive consequences on students as they arrive in new schools[1]. This paper settles this

[1] On the one hand, former classmates may be friends, and keeping friends may have a positive effect on well-being and achievement (see Lavy & Sand, 2012, in the case of the transition from primary to middle school). But former classmates may also be simple peers to whom it is easier to talk during the first weeks, to

question empirically, by assessing for the first time the positive role of basic familiarity with peers, whatever the precise relationships between these former classmates. Following Foster (2006), our results therefore call into question the common idea that agents should be more influenced by their friends than by other peers (see also Halliday & Kwak, 2012).

The focus of transition to high school is also particularly relevant, in view of the issues at stake. In many countries, formal tracking is implemented in high school, so that conjectural low achievement in the first year may end up in mismatched enrollment in low-ability tracks. In addition, the beginning of high school is often simultaneous with the end of compulsory schooling, meaning that underachievement may lead to drop-out at that stage compared to previous ones. As exposed in Figure I, a high drop-out rate appears in France in the first year of high school (10th grade), as well as a very high retention rate, resulting from the decline in achievement and students' decisions to avoid low-ability tracks. At the same time, most high school freshmen experience a dramatic social disruption that goes much beyond the loss of friends, as they often arrive in an environment in which they know (almost) nobody. In the case of France, Figure II reports that the average freshman starts high school in a classroom with only 5 percent of persistent classmates (1.7 student) compared to the 30 percent (7.5 student) they keep each year during middle school (more in section 1.3.2).

In this perspective, this work provides guidelines on classroom composition. By assigning low-ability freshmen who already know each other to the same class, principals could substantially increase their achievement in high school. According to our analysis, these students could see their risk of retention in freshman year reduced by 6.7 percentage points, and their graduation rate increased by the same amount. While very costly policies are usually targeting this population of students at risk, our recommendation on classroom composition could improve their performance by around 20 percent for no cost. Besides, moving these students across classes regarding their former networks is not a zero-sum game, in contrast to their ability or gender. Although a high-ability student or a girl might benefit to everyone[2], grouping freshmen coming from the same class should not affect freshmen from other classes.

sit next to in the classroom or to ask for help, thus making it easier to adapt to higher academic expectations and lower supervision from teachers. Even without bonds of friendship, familiarity within the classroom could therefore reduce anxiety, prevent social isolation and foster students' sense of belonging in their new school and classroom. On the other hand, former classmates could prevent students from socializing with new peers, or favor bad behavior in the classroom if disruptive students stay together. Former classmates may also be enemies rather than friends, and their presence could be detrimental to welfare and achievement. Mora & Oreopoulos (2011); Lavy & Sand (2012) show that "non-reciprocal friends" (peers that consider you as a friend while you do not, or the other way around) seem to have no or negative effects on outcomes.

[2] Carrell *et al.* (2011) built an algorithm designed to optimize peer effects and failed to do so partly for this reason.

Another close contribution is related to the strand of the literature on school mobility and school choice. Policies that enhance school choice or expand students' access to high-performing schools have been unexpectedly inefficient to improve students' educational outcomes, especially for males (see for example Cullen *et al.*, 2006; Kling *et al.*, 2007). This might be a direct consequence of the disruption generated on students' social network and environment by such programs, as suggested by Lavy & Sand (2012). Our findings support this interpretation, as they stress the importance for students of already knowing some peers when transitioning to high school.

Finally, the present paper puts into question the relevance of the structural transition that takes place during secondary education in many countries. The results suggest that comprehensive schools that include all secondary education grades may be less detrimental to low-ability students and may foster their completion of secondary schooling. Note that 30 percent of private middle schools in France also include high school grades, while only 5 percent of public middle schools do. This is one potential explanation of the relative attractiveness and performance of private schools compared to public schools, while the existing literature has mainly focused on differences in resources and composition between both sectors.

Section 1 describes the French education system and the data. Section 2 describes the identification strategy. We show the results and discuss the distribution of the effect and its mechanisms in section 3. Robustness checks are then provided in section 4. Section 5 discusses the implications of our results and concludes.

1 Institutional context and data

1.1 The high school curriculum in France

1.1.1 The transition to high school

By the end of middle school (grades 6 to 9), students apply for either vocational or general studies, with the approval of middle school teachers. Around two thirds of 9th grade students enroll in the general track, in which case they have to apply to general high schools in their district[3]. Rules of admission then differ with school districts and years[4], but they usually depend upon the students' home address, socioeconomic status and school performance (9th grade scores). The allocation is over by the end of June and high school administrations hold the registration files of their future 10th graders in the first week of July.

Simultaneously, 9th grade students take national anonymous exams in the end of June in three core subjects: mathematics, French and history-geography. These exams are not graded by teachers from students' middle school, but externally (with scores between 0 and 40). The resulting *anonymous scores* are combined with *continuous assessment* or *in-school scores*, *i.e.* scores obtained in 9th grade class in all courses and graded by students' own teachers (between 0 and 20). The anonymous scores and the in-school scores are combined to compute a total score that determines whether they pass the middle school graduation diploma (*Diplôme national du brevet* or DNB hereinafter) or not[5].

The anonymous scores are only available by mid July. By that time, students have completed their administrative registration to high school and classes' compositions are already determined. In addition, these scores are not sent to the high school during the summer[6]. Therefore, the principals assign freshmen to grade 10 classes without knowing their anonymous scores, but only their continuous assessment scores.

[3] Students enrolling in the vocational track have to choose a specialty, and vocational high schools have usually only one or two classes for each specialty. We decided therefore to exclude vocational high schools from this study, since classroom composition is very constrained and is not really policy-relevant in these schools.

[4] In 2008, a major reform has implemented an automatic procedure (*Affelnet*) to allocate students among general high schools.

[5] Note that students do not need to pass to pursue in high school.

[6] Some students do inform the high school of their results at the anonymous exams once they get them (even if it is not required), but according to informal discussions we had with some high school administrations, this hardly ever happens. In any case, principals do not know these scores for all students, so that they are very unlikely to use them to sort students across classrooms.

1.1.2 Tracking in high schools

In France, high school freshman year marks an important and difficult milestone for students. By the end of the year, they have to choose a major that will determine their 11th and 12th grade courses, their Baccalauréate (high school graduation exam) specialty, and the university tracks they will be able to apply for at the end of high school. First, students have to apply for for the academic or the technological track, the former being historically more prestigious, with more academic and difficult courses[7]. If students are not accepted in any of the majors they applied for, they can either opt for an alternative major suggested by teachers, or repeat grade 10 to apply again the following year[8]. Therefore, students' outcomes at the end of freshman year (retention, enrollment in an academic or a technological major, or drop-out) depend on both students' achievement and preferences.

Finally, high school ends at grade 12 with the *Baccalauréat* exam. This high school graduation exam includes anonymous tests in several subjects depending on students' major, and is almost entirely graded by teachers outside students' high school.

1.2 The class-assignment mechanism

In France, students are assigned to the same class for the entire school year and for all courses. Classmates have therefore even more influence on each other's outcomes, as they will spend most of the day together during the whole school year. In practice, the class assignment is made in early July, right after students have completed their registration to high school, and two months before the beginning of the school year in September. It is done completely by hand, without using any computer algorithm. Freshmen are not assigned randomly among classes, but contrary to other grades, high school principals do not know the students personally when they assign them to classes, which is the key feature used in this paper for identification. As a

[7] Within the academic track, they choose between three majors: science, humanities or social science. The scientific major is the most prestigious among them, as it allows the students to enroll in virtually any field of studies after high school, including humanities. Within the technological track, students major predominantly in industry, administration or healthcare. There are actually quite a few other majors, but they represent a very small share of the students.

[8] Students who are not allowed to move to the next grade may appeal the decision to committee exterior to the school, which has the final decision. Those who did not get the permission to enroll in their major of choice may appeal the decision to the principal or even negotiate with a different high school. Anyway, the final decision of allowing a student in a given major remains in the hands of the principal of the high school attended in 11th grade. According to the principals we met, very few students each year actually manage to go against their teachers' advice.

consequence, high school principals only use the set of formal information on students that is available in their registration files and observable in our dataset[9].

Principals first consider specific courses chosen by students. While most courses belong to the common core curriculum and are thus identical for all (*e.g.* mathematics or French), some specific courses are chosen by students, *e.g.* the foreign languages they want to study (*e.g.* English or Spanish) or additional optional courses (*e.g.* Latin or ancient Greek). Students who take the same specific courses are often grouped in the same class, for convenience with regard to classes' timetables.

Conditional to students' specific courses, school principals generally (but not necessarily) try to equilibrate classrooms in terms of gender and ability composition[10]. They rely on formal information contained in students' personal registration file: scores obtained in 9th grade courses (between 0 and 20), 9th grade teachers' comments on students' behavior, or personal information on the student and their family (mainly gender, age and parents' occupations). Contrary to other grades in which principals know their students, they cannot take into account personal knowledge about them such as motivation, mental strength or well-being[11]. Families do only rarely intervene directly with principals to influence the classroom assignment[12]. Strategical behavior to get one's child assigned to a better classroom works mainly through the choice of specific courses[13].

There are good reasons to believe that principals do not use all the detailed formal information they have on students during class assignment. As revealed by the sessions of class assignment we attended, proceeding to a simple allocation based on specific courses is already

[9] High schools are usually distinct from middle schools. However, 16 percent of French students are enrolled in schools that include the whole secondary curriculum (mostly private schools). In these schools, principals might know 10th grade students who come from their own middle school. However, the middle school and the high school still have separate assistant heads to which principals generally delegate classroom composition. These assistants do not necessarily coordinate during the class assignment of 10th grade students, so that the high school assistant heads may not use more information than the registration file. This is supported by our exogeneity test (see section 2.1.1), which suggests that students are conditionally randomly assigned even in this case. Therefore, we chose to keep students from these schools in our sample. Removing them from the sample has virtually no impact on the results.

[10] There is no legal requirement to do so, but the 1975 Haby law that implemented middle school comprehensiveness established a tacit rule urging schools to favor within-class heterogeneity. Besides, principals may not want to group all low-achieving students together in a class that will more likely get out of the teachers' hands.

[11] For other grades, they might for example separate two friends who are disturbing lessons, or allocate a fragile student with their friends to ensure them emotional support.

[12] They do so only in very specific cases, such as a special need for students in rural areas to be grouped in the classroom to help organizing shared car transportation.

[13] Families may incite their children to take "elite" courses (*e.g.* German as first foreign language, or Latin) to get them assigned to a better classroom. This has no consequence on our identification since we only compare students who chose the same specific courses.

highly complex and time-consuming. Again, they have to do it by hand and to take a multiplicity of constraints into account, while a host of other tasks are waiting to be carried out, both to close the current school year and to prepare the upcoming one. If two freshmen's registration files look broadly similar, principals are thus very unlikely to spend more time to investigate their characteristics further, and try to find some small detail to distinguish them (through phoning the family for example). In practice therefore, two 10th grade students do not need to be exactly identical on the paper to be undistinguished during the class assignment process. Section 2 will provide empirical evidence that support this observation made on the field.

1.3 Data

1.3.1 Data sets

The empirical analysis is based on two administrative datasets from the French Ministry of Education.

- *Administrative registration records*: for all students who are enrolled in French public and publicly-funded private middle and high schools from 2001 to 2012, this dataset contains personal information on students' identity (*e.g.* date and region of birth, gender, parents' occupation) and schooling: in particular grade, school and class attended, specific courses, grade and school attended in $t-1$ (but not the class attended in $t-1$).
- *Examination records*: for all students from 2004 to 2011, this dataset contains personal information and informal scores obtained at the 9th grade DNB (both at the anonymous test and the in-school scores) and the 12th grade *Baccalauréat* exams.

Unfortunately, there is no unique identification number that allows us to track each student through the different datasets. Yet for each 10th grade student, we need to know at least which class they attended in 9th grade, as well as the grade attended and major chosen in $t+1$ (repeating 10th grade, or moving to the 11th grade). We also have to match the administrative and the examination records.

For this purpose, we use a matching procedure based on students' personal information contained in each dataset. The procedure is mainly based upon the date and region of birth, the gender, and the grade and school attended in years t and $t-1$. We manage to match all needed information for 80 percent of new 10th grade students. In the rest of the paper, all regressions include controls for the share of unmatched students within the classroom, although they do not change the estimates.

Our identification will rely on the set of information on students we observe in our dataset with regard to the information observed by principals in their registration files at the time of classroom allocation. Therefore, it is useful at this stage to summarize what variable is observed by whom:

- *Covariates observed by both the principal and the econometrician (X_i):* Date of birth, city of residence, gender, parents' occupation, foreign languages and specific courses chosen, 9th grade in-school scores in all subjects, middle school and 9th grade classroom. We also observe a numerical measure of students' behavior as graded by the student's head teacher, but this information is missing for the first two cohorts (out of eight).
- *Covariates observed by the principal but not the econometrician (U_i):* Students' first and last name (from which, in particular, ethnicity could be inferred), precise home address. The principal also observe the 9th grade teachers' written comments that may signal behavioral issues.
- *Covariates observed by the econometrician but not by the principal (A_i):* Anonymous DNB test scores.

Most information observed by the principal is thus contained in the dataset. Although we do not observe the teachers' written comments, we do observe a behavioral score for three quarters of the sample, which contains precisely the information we expect the principals to infer from the written comments[14]. As we will show, the A_i variables are key in this study, since they enable us to test the our main assumption that freshmen are conditionally randomly assigned to classrooms.

1.3.2 Descriptive statistics

Descriptive statistics are presented in Table I, for the whole population of 10th grade students (column I). The population contains 2.9 million individuals (over eight cohorts), 55 percent of

[14] Note that it would be hard anyway to deal directly with written comments, even if we could observe them. In such a case, we would precisely try to build a score to summarize the information contained in the comments, which is the purpose of this behavioral score.

which are girls and 31 percent have parents with high socioeconomic status (SES) as measured by the father's occupation[15]). 10 percent of them have already repeated a grade at least once before reaching grade 10. The DNB anonymous exam is graded over 40 points and has mean 23.9.

The second part of Table I displays the average outcomes of the population during high school. 15 percent of students repeat 10th grade, while 62 percent and 17 percent enroll in an academic or a technological major, respectively. The 6 percent remaining students pertain to attrition due either to drop out or to "unmatched" students during the panelization procedure. Finally, 71 percent of new 10th grade students are taking the *Baccalauréat* exam three years later, and 57 percent do graduate then.

As reported on the bottom part of Table I, the average freshman has 1.7 *persistent classmates* (PC), over the 8.3 former classmates that enrolled in their high school[16]. This is very low compared to other grades of secondary education, as shown in Figure II that plots the classroom composition of a typical student in each grade. As a benchmark, the share of persistent classmates remains fairly constant across middle school at around 30 percent[17], meaning for instance that students in grade 9 have 30 percent of their current classmates who were also in their class in grade 8, in average. But in grade 10, the number of PC drops dramatically. Only 5 percent of their classmates come from the same class and 20 percent from the same middle school. Assuming that students only rarely know students coming from other middle schools, this means that students do not know at least 80 percent of their classmates at the beginning of the year. This figure then rises up again in following grades and amounts to 55 percent in grade 12 due to the partial conservation of major-specific classes from grade 11. Overall, Figure II shows how exceptional and intense the disruption in students' social environment is during the transition to high school.

[15] The SES field in our database actually contains the father's occupation if available, the mother's or legal guardian's otherwise.

[16] The ratio is roughly equal to 5, the average number of classes in high schools, suggesting that principals do not take the students' class of origin into account when allocating them among 10th grade classes.

[17] In grade 6, we are only able to identify students coming from the same elementary school, as we don't have any information on the classes in grade 5.

2 Identification

2.1 Identification

The identification strategy used in this paper is based on the conditional-on-observables random assignment of freshmen students. Since principals do not know freshmen students personally, those who were similar "on paper" and separated across different classrooms were presumably assigned randomly to their classrooms. Therefore, differences between classrooms of such students are uncorrelated to differences in individual unobserved factors of achievement, allowing for causal inference.

2.1.1 The class allocation of similar-file students

Following the notations defined in section 1.3.1, each 10th grade student comes to high school with a set of characteristics $(X_i, U_i, A_i, \epsilon_i)$, where ϵ_i is the vector of all remaining unobserved factors of achievement in high school. Yet to the principal's eye, they can only be distinguished based on (X_i, U_i), *i.e.* the limited set of formal information that they do observe in the registration files.

The X_i vector, that we also observe, contains most of this information. We thus partition students into groups (denoted g) within which they share the exact same value of each X_i covariate[18]. Since this vector is very large, this situation does not occur frequently and most groups will contain only one student. However, some groups may contain two (or even more) students, *i.e.* 10th grade students who are exactly or very similar on paper: we call them "similar-file" or SF students.

Such students may either be grouped in the same class or split in different classes. For instance, if two SF students chose a rare foreign language (*e.g.* Chinese) and only a few other students did so, then the school principal will probably group them in one class to simplify the making of the time schedules. But in other cases, they may be separated: if Chinese-learning students can be allocated to any of two (or more) classes, the school principal may want to equilibrate classes as much as possible by separating these students with similar characteristics. Whatever the reason for separating a group of SF students across different classes might be,

[18] As explained in details below, some X_i covariates are continuous and are thus discretized in precise enough quantile groups.

we assume that in this case, the principal will decide randomly which student they assign to which class.

In other words, we use sporadic cases of *similar-file students allocated to different classes* as natural experiments to identify the causal effect of the classroom environment. This approach relies on two distinct but related assumptions. First, we need to assume that the specification of the X_i vector of variables - defining the "similar-file groups" - is sufficiently narrow. This issue arises in particular from the continuous variables contained in X_i such as students' scores in grade 9. To get a reasonable amount of SF groups containing more than one student, we cannot consider only students with the exact same value of each score. At the same time, it is unlikely that students need to have exactly the same scores to be undistinguished by principals, and matching students that belong to the same quantile group of these scores (like quintile or decile) may be enough. But since we have no idea on the actual degree of precision used by principals, some *ad hoc* assumption has to be made and needs to be tested. Second, even conditional to the accurate specification of X_i, principals may use the U_i information we do not observe to assign SF students to their classrooms.

Formally, these two assumptions come down to the same one. Denote $\mathcal{S}(X_i)$ a given specification of X_i, *i.e.* a projection of X_i. All the remaining information that principals have on students, *i.e.* the complementary of $\mathcal{S}(X_i)$ to (X_i, U_i), denoted V_i. Therefore, V_i includes both U_i and the information included in X_i that is not accounted for by $\mathcal{S}(X_i)$, *e.g.* remaining differences in grades within a given decile. Using these notations, the core identification assumption of this paper is that *10th grade classroom characteristics C_{ic} are not correlated with V_i, conditional to $\mathcal{S}(X_i)$*:

$$C_{ic} \perp V_i | \mathcal{S}(X_i) \qquad i.e. \qquad \text{Cov}(C_{ic}, V_i | \mathcal{S}(X_i)) = 0 \tag{1}$$

where C_{ic} is the vector of 10th grade classroom characteristics. As long as $V_i|\mathcal{S}(X_i)$ is correlated with potential outcomes (or, to put it differently, with unobserved factors of achievement in high school ϵ_{ic}), any violation of (1) will lead to biased estimates of the causal effect of C_{ic} on achievement in high school. Assume for instance that there is substantial variation in ethnicity captured by names, even conditional to $\mathcal{S}(X_i)$, and that these variations are correlated to potential outcomes. If they are also taken into account by the principal when assigning these similar-file students to different classrooms, then our results would be biased.

Fortunately, assumption (1) can be tested since we do observe an information that principals do not: the students' anonymous scores A_i obtained at the national DNB exam just before entering high school. If conditional to $\mathcal{S}(X_i)$, principals do assign students based on information contained in V_i *and* that is correlated to potential outcomes[19], then we argue that some correlation between class characteristics C_{igc} and anonymous test scores A_i should be observed, again, conditional to $\mathcal{S}(X_i)$[20]. A_i can thus be used to test whether a given specification $\mathcal{S}$ is precise enough for assumption (1) to be credible. We do so by estimating the following model for each specification we tested:

$$C_{igc} = \alpha_g + \beta \cdot A_i + u_{igc} \tag{2}$$

where C_{igc} is any of class c's characteristics and α_g captures the SF-group fixed effect, *i.e.* the effect of sharing a specific vector of values for $\mathcal{S}(X_i)$. Adding α_g in the model constrains the regression to compare students only with their SF mate(s). If the null hypothesis $\beta = 0$ can be rejected, we conclude that the specification $\mathcal{S}$ is not precise enough. We thus reject the specification, considering that students with the same $\mathcal{S}(X_i)$ are probably still too different and are hence distinguished by principals[21], and we try more restrictive specifications until the balancing test is satisfied.

2.1.2 The optimal specification of SF groups

Within the set of specifications that satisfied the balancing test, we chose the least restrictive one, in order to maximize the number of SF students[22]. This optimal specification, denoted

[19] Any information that is not correlated to potential achievement would not bias our estimates of the causal effect of classroom characteristics on academic outcomes, be it used by principals or not.

[20] For instance, if 9th grade teachers see a student as disruptive enough to signal it by written comments, then they probably underscored his performance in class (as measured by in-school scores). Therefore, disruptive students should exhibit in average higher anonymous scores than their SF mate(s) with no behavioral issues, since SF students have very close in-school scores by construction. This can be shown empirically with our data on 2006-2011 cohorts, for which the NVS behavioral score is available. A regression of A_i on NVS$_i$ controlling for $\mathcal{S}^*(X_i)$ (our main specification described in section 2.1.2 and which includes in-school scores) exhibits a negative correlation of -0.059 with a 0.020 standard error. It shows that when teachers signal a student to have a worse behavior than their SF mate, this student gets a higher score at the anonymous DNB exam, revealing that they have been undergraded in class.

[21] This method led us to reject all specifications that did not imply exact matching. Controlling for very flexible functional forms of each X_i but without interacting them never allowed for satisfactory results on the A_i balancing test.

[22] $\mathcal{S}$ is considered more restrictive than $\mathcal{S}'$ if $\mathcal{S}'(X_i)$ can be deduced from $\mathcal{S}(X_i)$ for all X_i. A specification is said "minimal" if any other specification that is less restrictive does not satisfy the balancing test. As this order is not total, we found several such minimal specifications: among them, we chose the one that led to the largest number of SF students.

$\mathcal{S}^*(X_i)$, is used throughout section 3. It defines as "similar-file" students who come from the same 9th grade class in middle school; enroll in the same high school in the same year; select the same specific courses (*i.e.* same foreign language and optional courses); share the same gender, age[23] and social background (low- or high-SES) based on father's occupation; belong to the same quintile of in-school 9th grade average score in scientific subjects (mathematics, physics-chemistry, biology); belong to the same quintile of in-school 9th grade average score in humanities subjects (French, history and foreign languages)[24]; and belong to the same decile of the in-school 9th grade average score of all subjects enumerated above[25].

This specification leaves us 32,586 *groups* of SF students, 13,723 of which include students who do not end up in the same 10th grade class[26]. The total sample of SF students on which classroom effects can be estimated is thus made of 28,140 students, starting from an initial population of 2,888,258 10th grade students over eight cohorts[27]. In the rest of the paper, "the SF sample" denotes the 28,140 students that have at least one similar-file mate ("SF mate") when considering the specification $\mathcal{S}^*$ of X_i.

Table II reports the estimated β parameters of model (2) on the entire SF sample for a number of classroom characteristics C_{igc} (columns I and II). Column I measures the raw sample correlations between ability and classroom characteristics, *i.e.* without the α_g fixed effect. In the SF sample, more able students are assigned to larger classes and with more persistent classmates; their classmates are besides higher-achieving students, more often female and high-SES. All these correlations are statistically significant at the 1 percent level, except for the number of females. However, these correlations vanish within SF groups: as soon as we include the SF fixed effect, the estimates for β become very small and non-significant for all

[23] We do not look at the exact date of birth but only whether the students have been held back at least once or not: age is broken down to just one dummy variable.

[24] The foreign languages score is the weighted average of student's main foreign language (weight = 2/3) and second foreign language (weight = 1/3). Using different weights does not change the results of the paper. Besides, the in-school History score is missing for 5.4 percent of observations. For these students, the average humanities score is the average of the French and foreign languages scores only.

[25] Two students who belong to the same decile of average score in all subjects may have very different subject-specific profiles: one may have high marks in sciences but not in humanities, and *vice versa*. Most probably, principals do distinguish such students. This explains why we add separately quintiles of scientific and humanities scores, aside from the decile in the average score.

[26] 7,976 of the 32,586 SF groups are characterized by a set of optional courses that were only available in one classroom of the high school. Thus, these groups could not have been separated in any case. We conclude that principals have split 13,723 SF groups out of 24,610 (56 percent) groups that could be splitted.

[27] This population excludes 10th grade repeaters, but also newcomers for whom data on 9th grade exam scores is missing. Note that the optimal specification $\mathcal{S}^*$ allows only 1 percent of the population to come to high school with at least one other student (only one for 93 percent of them) who shares the same values for $\mathcal{S}^*(X_i)$, while ending up in different classes. This illustrates how much our identification approach requires a very rich database: only a large initial pool of students yields a sample of SF students that is large enough to get precise estimates of peer effects.

classroom characteristics (column II). In other words, for students who were similar regarding $\mathcal{S}^*(X_i)$ at the time of the class assignment, remaining differences in ability (unobserved by principals) have no correlation with differences between classrooms characteristics[28].

This is a very strong result in favor of assumption (1). It clearly suggests that principals do not use any achievement-related information in V_i^* to decide the class assignment of separated SF students, and thus separate them in a random way (or at least in an exogenous way regarding potential achievement outcomes). Actually, since X_i includes many variables, the specification $\mathcal{S}^*(X_i)$ is very narrow and the U_i vector of information very small in comparison, this result is far from irrealistic. It is absolutely consistent with the observations we made on the field that principals do not have time (or do not take it) to distinguish students that are very similar.

Another possible storyline is that principals do consider the remaining information in V_i^* and that the latter are indeed correlated to potential outcomes in high school, but *without being correlated to DNB test scores*. In other words, testing for unbalances in A_i would not be relevant since anonymous DNB scores are not a good measure of all unobserved determinants ϵ_i of achievement in high school[29]. Yet we argue that this is very unlikely, since principals neither observe DNB scores nor ϵ_i. It is therefore hard to imagine that principals would use the information contained in V_i^* to allocate SF students in a way that is correlated with ϵ_i but without *any* correlation with A_i as shown on Table II.

Table II also reports the results of the exogeneity test for a subsample of "at risk" SF students (columns III and IV). We define them as SF students who are low-achieving (below the median score of their middle school of origin), low-SES, and experiencing a disruption in their school environment (*i.e.* whose high school contains a higher share of high-SES students than their middle school). As we will show in section 3.2.1, our main results on the positive effect of keeping classmates during transition to high school are mostly driven by this specific subsample. For this reason, we checked that the exogeneity test performed well for these students, which is done in columns III and IV of Table II. Therefore, we conclude that SF students driving our main results are credibly exogenously assigned to their classrooms.

[28] We provide two additional tests in the online appendix. First, we show that the results hold when we use more detailed classroom characteristics regarding the number of persistent classmates of each type (low- or high-ability, same or opposite gender). Second, we estimated equation (2) the other way around, *i.e.* regressing A_i on all classroom characteristics C_{igc} at the same time, thus measuring *partial correlations* between ability and each of the classroom characteristics. The conclusions of both these tests are identical to Table II.

[29] For instance, students' level of autonomy may matter more for high school achievement than for achievement at the DNB exam.

Overall, specification $\mathcal{S}^*$ appears to be very adequate to estimate the causal effect of classroom environment on students' outcomes. To check the robustness of our results to other specifications, we show in section 4.1 that our results barely change when using alternative specifications that are more or less restrictive.

2.1.3 Additional evidence of random assignment using the behavioral score

The behavioral score (*Note de vie scolaire* or NVS) is not included in the optimal specification of SF groups $\mathcal{S}^*$. We do so because this score is not available for the first two cohorts, hence including it would force us to remove one fourth of our sample. However, Table II suggests that it does not constitute a threat to our identification assumption. Otherwise, the resulting allocation of SF students would create a correlation between C_{igc} and A_i conditional to $\mathcal{S}^*(X_i)$, since anonymous DNB scores are correlated to behavioral issues conditional to teachers' grades (see footnote 20).

Additional evidence that principals do not use behavioral issues to distinguish SF students can yet be provided. This is done by checking whether differences in classroom characteristics C_{igc} are correlated with potential differences in the NVS score in cohorts 2006 to 2011 (when the NVS score is available). We do so by estimating (2) after substituting A_i with a dummy for having a NVS score under the 10th percentile (equal to 15 over 20[30]). Results are presented on Table III. Most correlations between students' behavior and classroom characteristics are both very small and non significant at standard levels as long as comparisons are restricted within groups of SF students. This is true for both the whole SF sample (column II compared to column I) and the subsample of SF students at risk who drive our main results (column IV compared to column III). Finally, all the analysis presented in this paper has been reproduced focusing on the 2006-2011 cohorts and constraining SF students to share a similar NVS score. Similar results are systematically found, though the estimates are less precise[31].

Overall, results presented on Table II and Table III strongly support assumption (1) that high school principals randomly assign separated SF students (as defined in section 2.1.1) to their classes, or at least exogenously to achievement potential outcomes. The separation of

[30] The NVS score has a very specific, negatively skewed distribution. 33 percent students have the maximum 20 score since they exhibited no disruptive behavior, the average score is 18 while the median score is 19. Therefore, we chose to define students with disruptive behavior as students with a score below the 10th percentile, which is precisely equal to 15 over 20. Our results are not sensitive to the choice of the threshold.

[31] Results available on demand.

similar-file students across 10th grade classes creates differences in educational outcomes that can therefore be attributed in average to differences in classroom characteristics.

2.2 Description of the SF sample

Descriptive statistics on the SF sample compared to the initial population are presented on column III of Table I. Overall, students from the SF sample appears to be slightly higher achievers than the whole population of high school freshmen. But the differences are not always large in magnitude, though statistically significant. For example, students' average DNB test score is 25.1 (sd = 5.6) in the SF sample compared to 23.9 (sd = 5.1) in the whole population. 15.0 percent of the SF sample repeat 10th grade while 15.3 percent of the whole population does, a difference that is again very small[32]. Column III reports the same descriptive statistics for the subsample of at risk SF students. By construction, these students are very low on the ability distribution. They have an average normalized DNB score of −0.76, repeat grade 10 almost 2.5 times as often as the average student in the population, and graduate high school almost half as often.

[32] In terms of schools attained, SF students are found in 1,851 high schools out of 2,679, *i.e.* 69 percent of all high schools. The high schools that do not get SF students are mostly very small schools, in which the chances to get two students with the same X_i characteristics are small. They have in average 66 students in grade 10, versus 259 in average for high schools that do have SF students. Overall, these high schools account for 91 percent of all 10th grade students.

3 Results

3.1 Freshman-year classroom characteristics and achievement

Because students' assignment is assumed random conditional to $\mathcal{S}^*(X_i)$, differences in classroom characteristics between SF mates are orthogonal to differences in individual unobservable characteristics. Conditional to $\mathcal{S}^*(X_i)$ therefore, regressing outcomes on any classroom characteristic – *e.g.* classmates' average ability – identifies a contextual effect that is not attributable to unobserved individual characteristics. Yet, since classrooms differ in several dimensions simultaneously, the result of such regression could be driven by some correlated, omitted classroom characteristics – *e.g.* the number of females. Hence as a first step, we attempt to figure out what dimension of classroom environment is correlated to achievement, by regressing outcomes on several observed characteristics at once. Formally, we estimate the following model by OLS[33]:

$$y_{igc} = \alpha_g + \beta \cdot C_{igc} + \epsilon_{igc} \tag{3}$$

where y_{igc} denotes high school outcomes for student i, assigned to 10th grade class c and belonging to SF group g; as in model (2), α_g is the SF group fixed effect restricting the analysis to comparisons within groups of SF students; C_{igc} is a vector of classroom characteristics commonly studied in the literature (average ability[34], number of female students, number of high-SES students and class size), completed by students' number of persistent classmates; ϵ_{igc} captures individual unobserved heterogeneity and is orthogonal to classroom characteristics under assumption (1).

We estimated model (3) for different outcomes measured throughout the high school curriculum. The four first outcomes pertain to students' possible outcomes at the end of freshman year: repeating grade 10, dropping out[35], enrolling in an academic or in a technological major[36].

[33] Although our estimation strategy is similar in spirit to exact-matching methods, we chose not to use matching estimation as the regressors examined in this paper are not binary. As far as we know, the literature is very poor on the estimation of average causal effects of multi-valued treatments through propensity score or exact matching methods (see Imbens, 2000, in this direction).

[34] As measured by the DNB score. Because this data is missing for all retained students (around 10 percent of classmates) and for another 20 percent of classmates (not matched, see section 1.3.1), we also include quadratic controls for the shares of retained students and missing data.

[35] As described in section 1, this "drop-out" measure picks up attrition due both to matching issues and actual drop-out. Since classroom environment is unlikely to affect substantially the matching procedure though, we believe this measure adequately captures the effect of classroom characteristics on the risk of drop-out.

[36] We estimated model 3 without controlling for the SF fixed effect to get the raw sample correlations. In brief, the number of persistent classmates exhibits positive and significant correlations to all outcomes, with larger estimates than those obtained with model 3. Contrary to Table IV's estimates, classmates' average ability

Results are reported in columns I to IV of Table IV. The number of persistent classmates is positively associated with achievement at the end of grade 10. In average, each additional persistent classmates decreases by -0.3 percentage point (pp.) the risk of retention and translates into a similar increase of enrollment in academic major (se $=$ 0.1 pp.), with small and non-significant effects on drop out or enrollment in technological major[37]. An additional female classmate decreases the risk of retention by 0.1 pp. in average. This positive relationship between the female peers and school achievement is consistent with the results found by other studies (Hoxby, 2000; Lavy & Schlosser, 2011)[38]. Classmates' average ability is negatively associated with performance. A one standard deviation of classmates' average DNB score[39] increases significantly the risk of retention by 3.8 pp. in average, but decreases both the drop-out rate and the probability of enrollment in an academic major. Peer ability exhibit effects that are thus unclear[40]. The number of high-SES students also has a negative effect as it increases significantly the risk of dropping-out, but its magnitude is rather small. Finally, we find no effect of class size, most likely because of their small variance between the classrooms of a given high school (the standard deviation of class size is only 1.9 students within SF groups).

Table IV also reports results for two outcomes measured later than the end of 10th grade. Column V shows the effect of freshman-year classroom characteristics on students' probability to take the *Baccalauréat* exam "in time", *i.e.* three years after entering high school, meaning that they did not repeat grade 10 or grade 11 and that they made it through grade 12 without dropping out. Then, column VI investigates whether students with more persistent classmates in grade 10 are also more likely to pass the exam at that time. Interestingly, only the number of persistent classmates has a clear and persisting effect over time. Three years after entering high school, SF students who got an additional persistent classmate during their freshman year are still more likely to take the *Baccalauréat* exam at the end of grade 12. This result

and female share are respectively positively and negatively associated with achievement. Detailed results on raw correlations are reported on the online Appendix.

[37] We examined whether one academic major was driving the effect, but found the same positive, non-significant effect on enrollment in sciences, humanities or social sciences. Results on specific major enrollment are not reported for brievety, but are available on demand.

[38] When adding an interaction term between own gender and the number of female classmates, we find that this effect is driven entirely by female students (no effect on males). Note that controlling for this interaction term does not change the estimate of the PC effect. This rules out the interpretation of the PC effect as capturing the impact of assignment to same-sex classmates.

[39] The standard deviation of classmates' average ability *within SF groups* is only 27 percent of an average DNB score standard deviation.

[40] The results obtained in the literature for the effect of peer ability are also mixed and inconclusive. Here, the negative peer effect is consistent with the impact of a lower relative position within the class, because students may look weaker to teachers when assigned with better classmates. This may have little effect on drop-outs, but it would increase their risk of being retained in grade 10 and reduce at the same time their chances of admission in an academic major.

implies that the reduction in 10th grade retention was not cancelled by a higher propensity to repeat grade 11 or to drop out during grade 12. Furthermore, they do not seem to perform worse than others during these years, since they are also more likely to graduate high school. In comparison, all other classroom characteristics display estimates that are rather small in magnitude and never statistically significant. Thus, the number of PC seems very relevant to capture the dimension(s) of classroom environment that matter, even more so than the other, classic peer characteristics commonly studied in the literature.

Yet it is unclear what the number of PC actually measures or captures. SF students' persistent classmates could affect them through any sort of unobserved characterics that generate peer effects, such as ability or motivation. In what follows, we provide strong evidence suggesting that the PC effect does not capture an ability peer effect, but rather works through a social network mechanism. As we will show, the most consistent interpretation with the data is that students simply benefit from getting peers that they know and with whom they are used to interact.

3.2 The protective role of familiarity with classmates

We first check that the PC estimate is not affected by controlling or not for other classroom characteristics. In Table V, we report the previous estimates of the effect of PC from regressions where C_{igc} is the full vector (column I). In column II, the effect of the number of PC is estimated without controlling for the other classroom characteristics. The estimates are virtually identical between the two regressions, indicating that PC is not correlated with these other class characteristics. This first piece of evidence strongly suggests that the number of persistent classmates does not capture any other omitted dimension of classrooms environment. Suppose for instance that persistent classmates of SF students had specific characteristics linked to higher performance, which benefit to SF students without any link to familiarity. Considering Table V's results, such characteristics should be uncorrelated with DNB score, gender and SES. In other words, to be consistent with Table V's pattern, any credible omitted class characteristic driving the PC effect would have to be uncorrelated to all other observed classroom dimensions in C_{igc}. It is very unlikely that such a characteristic exists.

In the remaining tables of the present section, we systematically include quadratic controls for other classroom characteristics when we estimate the effect of PC. As seen in Table V, these controls do not affect our estimates.

3.2.1 Distribution of the PC effect

Another reason to believe in the "familiarity" interpretation bears upon the heterogeneity of the PC effect. If the historical familiarity with classmates really matters, it is probably not equally important for all students. In particular, we expect that keeping something unchanged in your environment is helpful precisely when everything else gets disrupted during the transition to high school. This is exactly what we find.

The analysis of the distribution of the PC effect can be found on Table VI. We first look separately at low- and high- ability students, defined by their relative position to the median in-school score of their middle school of origin (rows B and C)[41]. The PC effect is strikingly heterogenous across these two categories. While the number of PC has virtually no effect on high-ability students, low-ability ones are strongly, positively impacted. For brevity, we do not comment the magnitude of the estimates, since the effects are actually very heterogenous again within this subgroup, between low- and high-SES. As reported in rows D and E, the effects observed on low-ability students are almost exclusively driven by low-SES ones. In average, each additional PC reduces their risk of retention by 1.4 pp, though not their risk of dropping out. As a consequence, they are significantly more likely to enroll in either an academic or a technological major, with a similar increase in magnitude. No backlash to this strong short term impact can be found in following grades. In average, each PC during freshman year increases their chances to be taking the *Baccalauréat* exam and to graduate by the same amount. In comparison, the estimates are very small in magnitude and never statistically significant at conventional levels for low-ability high-SES students (column V) or high-ability low-SES students (not reported) This indicates that keeping some classmates matters only for students who may be experiencing a hard transition *both* academically – they were already performing badly in middle school – and culturally – their parents come from the working class and might not have studied in high school.

In rows F and G, we investigate further this distribution pattern by looking how the PC

[41] We use the in-school score since SF groups are defined with regard to it, so that two SF students are necessarily both below or above the median. Although the anonymous DNB score would be a better measure of ability, two SF students may be on different sides of the median DNB score. We would thus lose part of the SF sample by analyzing the PC effect separately at each side of the median DNB score. However, doing so brings out the same conclusions as in Table VI, though the estimates are often less precise.

effect varies with the difference in school-level social environment (measured by the share of high-SES students). This gap, denoted Δp, is negative for one third of low-ability low-SES students only[42]. Twice more often, these low-ability, low-SES students experience a positive Δp, meaning that they get into an environment with more high-SES students than they used to have. The heterogeneity of the effect is smaller in this dimension than in the previous one, but goes in the same direction. While estimates remains substantial in magnitude for students with $\Delta p \leq 0$, the precision of the estimates is systematically low, *e.g.* -0.8 pp. (se $=$ 1.3 pp.) for grade retention (row F). The PC effect is much larger in magnitude on all outcomes for the subsample of low-ability, low-SES students with $\Delta p > 0$, referred to as the "at risk" SF subsample. In average, each additional PC in grade 10 mitigates the risk of retention by -1.8 pp. (se $=$ 0.7 pp.), and this effect persists until graduation.

We also estimated the PC effect separately for male and female at risk students. However, the results do not exhibit any clear heterogeneity in the gender dimension, as reported in rows H and I of Table VI. Persistent classmates seem to have more impact more on the males' retention rate than the females'. But the discrepancy goes in the opposite direction for the *Baccalauréat* outcomes, with larger estimates for females. Both male and female students thus seem to benefit from persistent classmates in the freshman year, although the benefits are slightly different considering the stage of the high school curriculum[43]. We analyzed further the distribution of the effect with regard to the middle and high school contexts. The results are not reported since no other interesting pattern could be found. For example, the effect does not seem to vary significantly with the middle or high school sizes, the share of middle school classmates attending the high school, or the 10th grade classroom context[44].

Overall, the results of this investigation are consistent with our interpretation of the PC effect. The estimates reported on Table IV were a diluted version of the very strong PC effect located on SF students that experience a strong disruption during the transition to high school[45].

[42] This is a mechanical consequence of the lower probability of low-SES students to enroll in general high schools after grade 9.

[43] In fact, an interesting pattern appears when examining the precise academic major in which males and females enroll. If the raise in academic major enrollment is similar between both gender, persistent classmates only drive males towards science and females towards humanities. More precisely, both male and female PC increase male enrollment in science, while females enroll more in humanities only when they get more female PC. Results available on demand.

[44] We checked in particular whether the degree to which your new classmates are grouped with their former classmates increased your need to be with yours. Yet again, we found no result in this direction This is noteworthy as it suggests that grouping former classmates would not drive negative spillovers on their other classmates, who did not necessarily have many former classmates in the high school. Though, it would be helpful to confirm such a conclusion with a controlled field experiment that would allow direct examination of externalities within the classroom.

[45] We checked whether the other peer characteristics studied on Table IV also had a larger effect on this specific

Already knowing some peers in the classroom matters a lot for low-ability students with low socioeconomic status who came from a deprived environment compared to the high school. This is very consistent with the interpretation of the PC effect as reflecting the impact of familiarity. By contrast, it is unlikely that the former classmates of these low-achieving deprived students have higher unobservables than the average, which would drive the PC effect. The following section adds supplementary evidence in this direction.

3.2.2 Do all former peers matter?

If the effect of persistent classmates is explained by familiarity, then at risk students should be more affected by peers with whom they have been more likely to interact during middle school[46]. For example, they may have interacted much more with their former classmates than with peers from their middle school but assigned to other classrooms. In Table VII, panel A, we add to the previous regressions the number of these former *middle school mates from other classes*. We find a small, negative effect on grade retention, though not significant. Surprisingly, this effect is related to a small increase in the risk of dropping out, statistically significant at the 5 percent level. Other estimates are very small in magnitude and never statistically significant. Therefore, students seem to benefit only from their middle school mates that came from the same classroom, with whom they probably interacted much more.

Students may also be more likely to interact with same-gender peers. In panel B, we thus separated the number of persistent classmates into the numbers of same- and different-gender. As a matter of fact, students seem to benefit from both types of persistent classmates, even if the effect is slightly stronger and more precisely estimated for same-gender PC. The main difference pertains to grade 11 major enrollment. Same-gender PC drive students strongly towards academic majors only, while opposite-gender PC only increase their chances to enroll in a technological major. All together, these results suggest that students benefit slightly more from keeping same-gender than opposite-gender persistent classmates.

Finally, we examine whether these low-ability students are more impacted by their persistent classmates who exhibit similar achievement (*i.e.* who were themselves low achieving in grade 9),

category of students "at risk". Results are provided on the Online Appendix. Again, other peer characteristics (average ability, number of females, etc.) display non-significant and non-persisting effect on achievement in high school, even for these students.

[46] We only focus on the subsample of at risk SF students since section 3.2.1 showed that they were the only one driving the PC effect.

with whom they are again more likely to interact (panel C). The discrepancy is much stronger here. Each low-ability PC has a tremendous positive effect on disrupted students, reducing the risk of retention by 2.8 pp. in average and increasing chances of enrollment in all kinds of major by a similar amount. Around 75 percent of this effect can still be observed three years after on students' probability to take and pass the *Baccalauréat* exam. In comparison, the estimates for high-ability persistent classmates are very small and never statistically significant[47].

To summarize these results, among all 10th grade classmates coming from their middle school, students only benefit from the presence of those who were in their classroom. And the more they have in common with these PC, the stronger the impact. This is additional evidence that the PC effect comes from the high familiarity they have with these peers. The finding that low-ability students are almost exclusively impacted by their low-ability PC is particularly relevant in this perspective. Since under-achieving PC are very unlikely to have some positive unobservable characteristics that is not shared by high-achieving PC, our results most probably capture a social network dimension. In other words, we argue that students benefit from their persistent classmates only because they know each other well.

[47] We carry on defining "low-ability" by students with in-school scores below the school median, for consistency with Table VI. However, anonymous test scores are a better measure of ability and could be used here to define classmates' ability without loss of precision (by comparison to SF students, see again footnote 41. Actually, doing so leads to an even greater discrepancy, with low-ability PC driving an effect of -4.0 pp. on grade retention while high-ability PC have virtually no effect.

4 Robustness checks

4.1 Alternative specifications of SF groups

All the results presented in section 3 are based on the specification $\mathcal{S}^*$. It is the optimal specification in the sense of the least restrictive specification under which our exogeneity tests are valid. In this section, we investigate the sensitivity of the results to the choice of the specification $\mathcal{S}$.

We tested both specifications that were more restrictive and less restrictive than $\mathcal{S}^*$. In Table VIII, we report the effect of persistent classmates on grade retention for four alternative specifications on the primarily affected sample of "at risk" students (low-ability, low-SES, $\Delta p > 0$). The reference specification $\mathcal{S}^*$ is reproduced in column IV; columns I to III show the results for less restrictive specifications and column V shows a more restrictive specification. The details of each specification is given in the table.

All specifications lead to a negative, significant effect of the number of persistent classmates on grade retention, although the magnitude of the effect varies from one specification to another. One should keep in mind that the balancing test presented in section 2.1 leads to less convincing results for the alternative specifications[48]. The results in column IV therefore remain our reference results. However, the fact that the effect keeps the same sign and order of magnitude is reassuring regarding the validity and robustness of our results.

4.2 Estimation based on the impact of SF students' allocation on their classmates

In section 3, identification of peer effects within classrooms derives directly from the comparison of SF students who were randomly assigned to different classes. But the random assignment of SF students may also be considered itself as a shock to the classrooms' composition. Receiving one or the other SF mate in the class can make a difference for the other students in these classrooms. Focusing on the consequences of SF students' allocation on their 10th grade classmates allows us to provide new estimates of the effect of classmates' persistency, thus testing the robustness of the results presented in section 3.

[48] Results available on demand.

4.2.1 Principle

Using the notations from Figure III, we now compare with each other students C to H instead of comparing student A to student B. If A and B are defined as sharing the same values of $\mathcal{S}^*$, the result of the random allocation of A and B should have no impact on these students *a priori*, since A and B have the same characteristics. However, if we allow A and B to come from *two different classes* of the same middle school, the result of the allocation will have an impact on students C to H if some of them come from A or B's 9th grade class. Because of the similarity of A and B in most dimensions except for their exact class of origin, the result of the allocation will only affect this dimension of the class characteristics vector C_{ic}. For example, if students A and C come from the same 9th grade class, C would have one more PC in case 1 than in case 2.

Therefore, in this section, we use another specification of SF students called $\mathcal{S}_d^*$ which is similar to $\mathcal{S}^*$ except that we allow students to come from *different classes*. However, principals do observe classroom of origin and may distinguish students if their previous classroom were too different. This was confirmed empirically, as we had to require the classes of origin to be similar for the exogeneity test 2 to be satisfied. Thus, students with the same $\mathcal{S}_d^*$ had to come from classrooms with a similar level (quintile of the average DNB score), allowing or not for elite optional courses, and sending a similar number of students in the high school (same quintile of the number of former classmates in the high school).

Formally for each SF group j produced by the $\mathcal{S}_d^*$ specification, we can define an instrument Z_{ij} equal to the number of persistent classmates that student i obtained from that SF group. The variable is defined only for students who were in the same 9th grade class as one of the SF students in group j, and who are in one of the 10th grade classes attended by these SF students.. We denote this sample $\mathcal{P}_j$. Note that in this context, i belongs to the sample of *former classmates* of the SF sample, $\mathcal{P} = \bigcup_j \mathcal{P}_j$, and not to the SF sample itself.

Formally, we estimate the following reduced form model:

$$Y_{ijkc} = \alpha_{jk} + \beta \cdot \mathrm{Z}_{ij} + \epsilon_{ijkc} \tag{4}$$

where c denotes the 10th grade class and k denotes the 9th grade class. The α_{jk} fixed effect ensures that comparisons are made between students who belong to the same $\mathcal{P}_j$ sample and come from the same 9th grade class. Alternatively, this fixed effect can be replaced with a

α_{jc} fixed effect, where we only compare students ending up in the same 10th grade class. β identifies the causal effect of getting one additional persistent classmate[49].

4.2.2 Validity of the test

The exogeneity of our instrument relies on a stronger assumption than the main model. Even if SF students are randomly split, the assignment of other freshmen is not exogenous. In particular, if they were assigned regarding the number of PC, the instrument Z_{ij} would not be exogenous. Therefore, model 4 is identified under the hypothesis that $\mathcal{P}$ students are not allocated to classes in correlation to PC.

In order to check this additional hypothesis, we estimate the correlation between the value of the instrument Z_{ij} (the number of persistent classmates received through random allocation) and the individual characteristics of the students $i \in \mathcal{P}_j$:

$$Z_{ij} = \alpha_{jk} + \gamma \cdot X_i + u_{ijkc} \tag{5}$$

where X_i is the vector of observable characteristics tested.

We show the results of this test in Table IX. In column I, we find that individual characteristics are correlated with the instrument when the controls for α_{jk} fixed effects are not included. However, these correlations vanish when we include them (column II) or when we replace it with a 10th-grade-class fixed effect (column III)[50]. These results suggest that the students who obtained a PC through the random assignment are comparable on observed dimensions to those who did not, within $\mathcal{P}_j$ samples. Although these students might be different on an unobserved level, we argue that this test is satisfactory enough to run a robustness check or our main results presented in section 3.

4.2.3 Results

The results of the estimation of model (4) are produced in Table X. Like in Table IX, the α_{jk} fixed effect is omitted in column I, included in column II and replaced with a 10th-grade-class

[49] Z_{ij} is a "perfect" instrument for PC, as it has a correlation of one with PC and as there is no compliance issue here. This is why we estimate the reduced form model directly.

[50] We used DNB quintile dummies instead of the DNB score to avoid losing the students with missing values. Students for which the DNB score is missing have all five dummies equal to zero.

fixed effect in column III. Since Z_{kc} has the same value for all students in 10th grade class c and coming from the same 9th grade class k, the standard errors are clustered within kc groups (following Moulton's formula). Similar to section 3, we find that a higher number of persistent classmates are associated with lower grade retention and higher enrollment in the academic major. We also observe a positive, long term effect on high school graduation. The orders of magnitude are similar to the results using the first strategy and do not vary drastically depending on the fixed effect that we include.

Overall, these results confirm our results from the main strategy. Besides, this approach present some advantages, although it relies on a stronger assumption. First, the main strategy focuses on students who have been separated from a very similar former classmate, likely a friend. Getting more persistent classmates may have more impact than usual in such settings. With the current approach, we find a similar impact on a different sample, thus removing doubts about the external validity of our results. Furthermore by allowing comparisons of students within the same classroom (column III), the effect can be estimated of a pure variation in the number of PC, holding other classroom characteristics constant[51] This mitigates our concerns about omitted classroom characteristics driving our results in section 3. Last but not least, it shows that the positive effect of persistent classmates does not operate only through improvement in the global classroom context, that would affect everyone similarly[52]. Freshmen do therefore benefit from familiar peers through channels that operate at the individual level, such as higher sense of belonging or social and academic support.

[51] Suppose on Figure III that C comes from A's class and D from B's class. For C and D, getting A or B does only change their relative number of PC. This is true *by construction*, because A and B have the same characteristics regarding all other dimensions.

[52] One could have expected for example that it is more comfortable to teach a class if more students already know each other in the classroom. This beneficial impact on teachers could then affect all students in the class, even those who are not directly affected by having more former classmates. Yet in this case, we should not find any difference between students in the same 10th grade class.

5 Discussion and conclusion

This paper documents how classrooms influence students' achievements in high school. Empirical evidence suggests that freshmen students with very similar registration files, when separated across different classrooms, were randomly assigned to their class. Therefore, differences in classroom environments can be credibly assumed orthogonal to potential outcomes. After examining the correlations between several measures of classrooms' composition and students' outcomes, we find a robust and significant effect of being assigned again with more former classmates. But this effect is all but homogenous. It is almost exclusively driven by low-achieving, low-SES freshmen who enroll in high schools with more high-SES mates than they used to have. These students "at risk" during high school benefit almost only from the presence of low-achieving persistent classmates. It may be a surprising result, since these peers are unlikely to be of any academic help compared to high-achieving persistent classmates. Most probably, low-achieving students are better-off by keeping similarly low-achieving PC through social channels.

Mechanisms implying direct interactions could be at work. For example, persistent classmates could be friends or acquaintances to whom freshmen may talk during the first weeks, ask for help, or even work as a team[53]. Even without interacting with them though, being surrounded by peers they know and who experience the same difficulties may also be a psychological relief, fostering their sense of belonging in the high school. Quite clearly, more data would be needed to understand how freshmen take advantage of familiarity with peers. Most importantly though, results show that students do not bear the brunt of increased enrollment in grade 11 by lower performance in subsequent years. So, whatever the mechanisms at work, we know that persistent classmates do increase achievement[54].

Basically, this result is sufficient to draw relevant policy recommendations on classroom composition. Whereas very expensive efforts are usually spent on students at risk of under-achievement, we show that assigning them to some persistent classmates could increase their performance for no cost. And the potential gains could be substantial. According to our analy-

[53] In this way, grouping students who come from the same class may be an efficient tool to help teachers to develop cooperative learning within the classroom, as they might rely on existing friendships and social links between students right from the start of the year.

[54] In particular, mechanisms implying solely a change in preferences are ruled out by this result. For example, persistent classmates could make students less likely to repeat grade 10 only by increasing their propensity to appeal teachers' decision (to avoid losing friends) In that case though, negative drawbacks should be observed in following grades since students enroll in grades and majors where requirements would be to high.

sis, each low-achieving persistent classmate reduces their risk of retention by 2.8 pp. in average. This figure is estimated using only the variance of PC observed within groups of at risk SF students, with 98 percent of variations ranging from 0 to 3 low-achieving PC (no conclusion should be drawn on the PC effect beyond this range). Students at risk in the freshmen population have 0.61 low-achieving PC in average: increasing this figure to 3[55] could thus reduce their risk of retention by 6.7 pp. (meaning 18 percent of their current rate) while increasing their graduation rate by the same amount. Non-linearities of the PC effect might result in the same (or even a higher) benefit with less than 3 PC, but the small variance of the number of PC in our sample does not allow us to investigate it[56].

We think this study makes an important contribution to the existing literature on the role of school environment on achievement. While a large strand of studies have been looking for non-linearities in peer effects, the finding that low-achieving students may benefit from low-achieving peers may seem counter-intuitive in the first place. But this is true only as long as they know each other, and as the rest of their environment gets largely disrupted by the transition to high school. In fact, the need for some minimum stability when one faces large unstability is rather intuitive, but emphasizes the high complexity of peer effects and social interactions.

[55] 3.14 low-achieving former classmates are enrolled in their high school in average.

[56] We tried to regress outcomes on the number of high-ability PC, the number of low-ability PC and a dummy for the latter being strictly positive. The estimate for the number of PC decreases slightly in magnitude from -2.8 to -2.0 pp. and is not significantly different from 0 anymore (se = 1.2 pp.). The estimate for the dummy amounts to -6.5 pp. but has also a bad precision (se = 4.9 pp.). These results are difficult to interpret, so we prefer not to draw any conclusion on the non-linearity of the PC effect.

References

ANGRIST, JOSHUA D., & LANG, KEVIN. 2004 (dec). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *The American Economic Review*, **94**(5), 1613–1634.

CARRELL, SCOTT E., SACERDOTE, BRUCE I., & WEST, JAMES E. 2011 (Feb.). *From Natural Variation to Optimal Policy? The Lucas Critique Meets Peer Effects*. NBER Working Paper 16865. National Bureau of Economic Research.

CULLEN, JULIE BERRY, JACOB, BRIAN A., & LEVITT, STEVEN. 2006 (Sept.). The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica*, **74**(5), 1191–1230.

DUFLO, ESTHER, DUPAS, PASCALINE, & KREMER, MICHAEL. 2011 (aug). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *The American Economic Review*, **101**(5), 1739–74.

FOSTER, GIGI. 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.

GOUX, DOMINIQUE, & MAURIN, ÉRIC. 2007 (oct). Close Neighbours Matter: Neighbourhood Effects on Early Performance at School. *Economic Journal*, **117**(523), 1192–1215.

HALLIDAY, TIMOTHY J., & KWAK, SALLY. 2012. What is a peer? The role of network definitions in estimation of endogenous peer effects. *Applied Economics*, **44**(3), 289–302.

HOXBY, CAROLINE M. 2000. *Peer Effects in the Classroom: Learning from Gender and Race Variation*. NBER Working Paper 7867. National Bureau of Economic Research.

IMBENS, GUIDO W. 2000 (Sept.). The Role of the Propensity Score in Estimating Dose-Response Functions. *Biometrika*, **87**(3), 706–710.

KLING, JEFFREY R, LIEBMAN, JEFFREY B, & KATZ, LAWRENCE F. 2007. Experimental Analysis of Neighborhood Effects. *Econometrica*, **75**(1), 83–119.

LAVY, VICTOR, & SAND, EDITH. 2012 (Oct.). *The Friends Factor: How Students' Social Networks Affect Their Academic Achievement and Well-Being?* NBER Working Paper 18430. National Bureau of Economic Research.

LAVY, VICTOR, & SCHLOSSER, ANALIA. 2011 (Apr.). Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics*, **3**(2), 1–33.

MORA, TONI, & OREOPOULOS, PHILIP. 2011. Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. *Economics of Education Review*, **30**(4), 575–581.

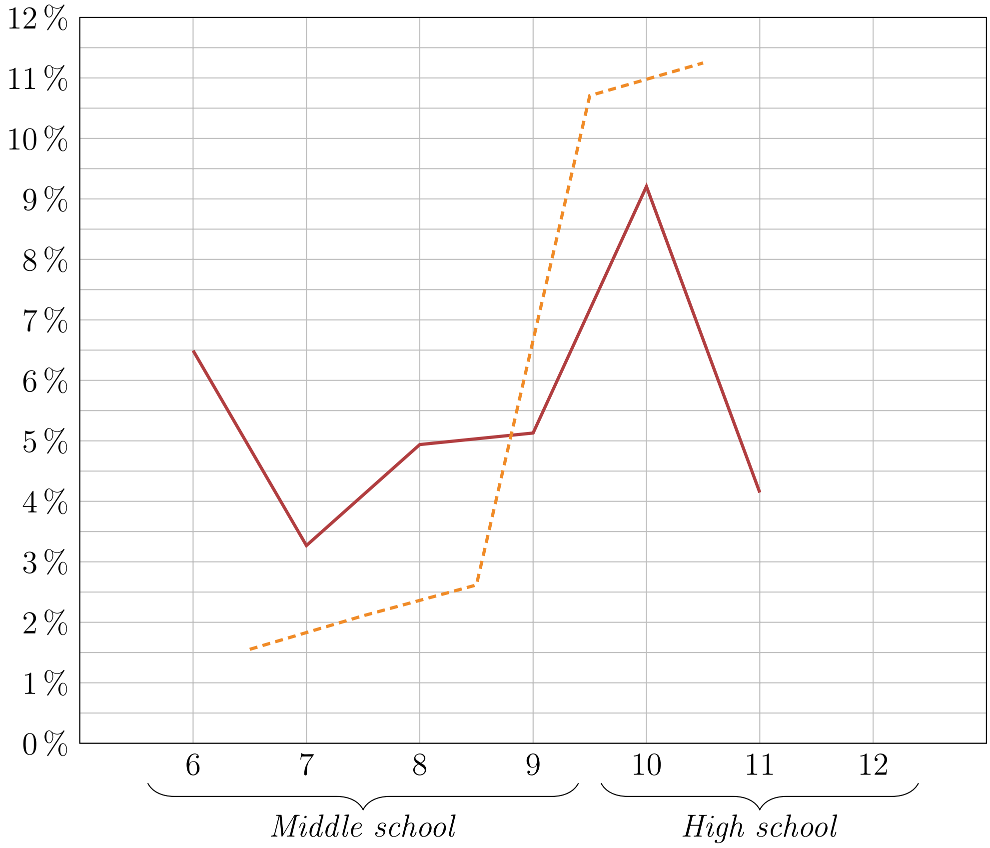
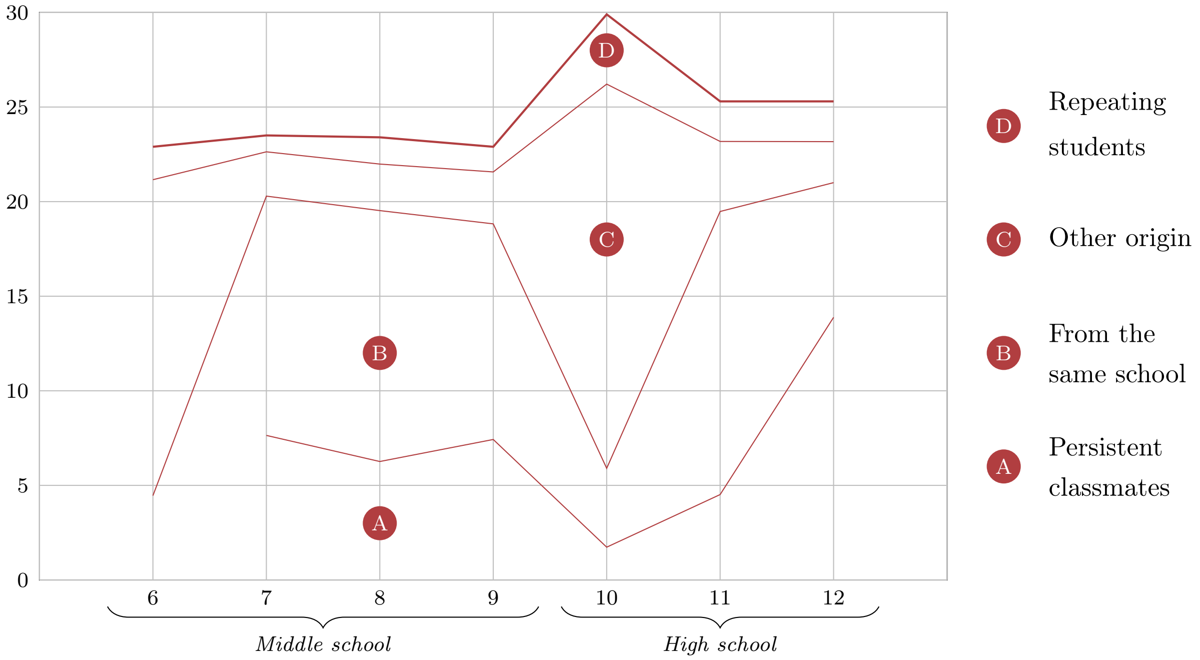


Figure I: Retention rates of students who did not already repeat the current grade. Students who repeat the year to choose a new major are not counted as repeaters.



Sample for grade g consists of all students entering grade g for the first time (non-repeating) between years $1994 + g$ and $2001 + g$.
First year missing for grades 6 and 7; last year missing for grade 12.
Reading: In grade 10, in the average non-repeating student's class, there are 1.7 persistent classmates, $5.9 - 1.7 = 4.2$ former schoolmates, $26.3 - 5.9 = 20.4$ students from other origins and $29.9 - 26.3 = 3.6$ repeating students, *i.e.* 29.9 in total.

Figure II: Composition of the typical classroom from a non-retained student's point of view

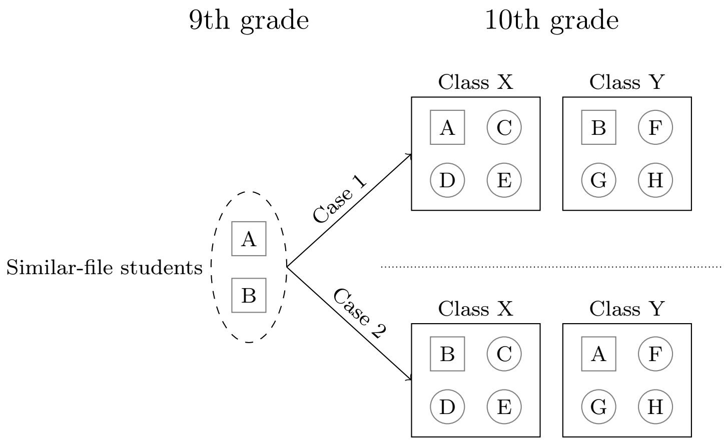


Figure III: Class assignment of similar-file students

Table I: Descriptive statistics on students' characteristics

	Population	SF sample	At risk
	(I)	(II)	(III)
Girl	0.547	0.618	0.600
	(0.498)	(0.486)	(0.490)
High-SES	0.306	0.301	0.000
	(0.461)	(0.459)	(0.000)
High quality optional course	0.150	0.093	0.020
	(0.357)	(0.291)	(0.142)
DNB national exam score	23.935	25.136	20.073
	(5.077)	(5.554)	(3.882)
Normalized DNB national exam score	0.000	0.245	−0.764
	(1.000)	(1.099)	(0.760)
Had repeated at least once before grade 10	0.102	0.037	0.101
	(0.302)	(0.190)	(0.302)
Repeats 10th grade	0.153	0.150	0.371
	(0.360)	(0.357)	(0.483)
Attrition (drop out or unmatched in panel)	0.058	0.041	0.084
	(0.234)	(0.197)	(0.277)
Academic major in grade 11	0.620	0.693	0.295
	(0.485)	(0.461)	(0.456)
Technological major in grade 11	0.169	0.117	0.251
	(0.375)	(0.321)	(0.433)
Takes *Bac* in time	0.707	0.736	0.464
	(0.455)	(0.441)	(0.499)
Graduates high school	0.574	0.630	0.307
	(0.494)	(0.483)	(0.461)
Number of PC in 10th grade class	1.721	1.999	1.532
	(2.523)	(2.258)	(1.781)
Number of former classmates in high school	8.325	12.679	10.639
	(6.431)	(5.925)	(4.783)
N	4,129,926	28,140	6,054

Standard deviations are reported in parentheses.

Table II: Student's classroom characteristics regressed on own anonymous exam score: Evidence of the random assignment of similar-file students

	All	All	At risk	At risk
Dependent variable	(I)	(II)	(III)	(IV)
Persistent classmates (PC)	0.062***	0.000	0.046***	−0.010
	(0.002)	(0.005)	(0.006)	(0.012)
Normalized[1]DNB score	0.040***	0.001	0.046***	0.001
	(0.001)	(0.001)	(0.001)	(0.002)
Number of girls	0.004	0.001	−0.003	−0.036
	(0.006)	(0.011)	(0.017)	(0.031)
Number of high-SES students	0.316***	0.013	0.283***	0.018
	(0.007)	(0.009)	(0.017)	(0.021)
Class size	0.067***	0.002	0.057***	−0.016
	(0.004)	(0.006)	(0.011)	(0.017)
N	28,095	28,095	6,040	6,040
SF fixed effect	No	Yes	No	Yes

[1] The normalization is done over the whole population; the sample's mean is 0.245.
Each cell is from a separate regression of the classroom characteristic of interest on the student's standardized average anonymous score at the DNB exam. All regressions include quadratic controls for the share of retained students and of missing DNB scores. Robust standard errors are reported in parentheses.

Table III: Student's classroom characteristics regressed on behavior score: Evidence of the random assignment of similar-file students

	All	All	At risk	At risk
Dependent variable	(I)	(II)	(III)	(IV)
Persistent classmates (PC)	−0.177***	−0.114	−0.113	0.089
	(0.054)	(0.082)	(0.072)	(0.134)
Normalized[1]DNB score	−0.157***	−0.005	−0.133***	−0.036
	(0.013)	(0.013)	(0.020)	(0.026)
Number of girls	−0.552***	−0.048	−0.617***	0.183
	(0.127)	(0.160)	(0.219)	(0.330)
Number of high-SES students	0.026	0.067	−0.103	−0.324
	(0.191)	(0.119)	(0.233)	(0.227)
Class size	−0.292***	−0.037	−0.236*	−0.071
	(0.088)	(0.106)	(0.129)	(0.162)
N	15,825	15,825	3,330	3,330
SF fixed effect	No	Yes	No	Yes

[1] The normalization is done over the whole population; the sample's mean is 0.245.
Each cell is from a separate regression of the classroom characteristic of interest on the student's standardized average anonymous score at the DNB exam. All regressions include quadratic controls for the share of retained students and of missing DNB scores. Robust standard errors are reported in parentheses.

Table IV: Effect of classroom characteristics on high school outcomes

	Repeats 10th grade	Drops out	Academic major	Tech. major	Takes *Bac* in time	HS graduate
Independent variable	(I)	(II)	(III)	(IV)	(V)	(VI)
PC	-0.003***	-0.001	0.003**	0.001	0.005**	0.004*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Average DNB score	0.038***	-0.012*	-0.021**	-0.005	-0.008	0.009
	(0.010)	(0.006)	(0.010)	(0.010)	(0.013)	(0.014)
Number of girls	-0.001**	0.000	0.001	0.000	0.001	0.001
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Number of high-SES students	-0.000	0.001**	-0.000	-0.001	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Class size	-0.000	0.001	0.000	-0.000	0.002	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
R^2	0.68	0.56	0.79	0.63	0.68	0.71
N	28,140	28,140	28,140	28,140	23,019	23,019
SF fixed effect	Yes	Yes	Yes	Yes	Yes	Yes

[1] *Bac* data was not available for the last two cohorts.
Each column is from a separate regression of students' outcomes on their classroom characteristics. All regressions include controls for the share of retained students and of missing DNB scores. Robust standard errors are reported in parentheses.

Table V: Effect of persistent classmates on high school outcomes with and without controlling for other class characteristics

Dependent variable	(I)	(II)
Repeats 10th grade	−0.003***	−0.003***
	(0.001)	(0.001)
Drops out	−0.001	−0.001
	(0.001)	(0.001)
Academic major	0.003**	0.003**
	(0.001)	(0.001)
Technological major	0.001	0.001
	(0.001)	(0.001)
N	28,140	28,095
Takes *Bac* in time	0.005***	0.005***
	(0.002)	(0.002)
HS graduate	0.004**	0.004*
	(0.002)	(0.002)
N	23,019	22,981
Control for other class characteristics	No	Yes

Each cell is from a separate regression of students' outcomes on their classroom characteristics. All regressions include controls for the share of retained students and of missing DNB scores. Robust standard errors are reported in parentheses.

[1] *Bac* data was not available for the last two cohorts.

Table VI: Distribution of the PC effect

	Repeats 10th grade	Drops out	Academic major	Tech. major	Takes *Bac* in time	HS graduate
Population	(I)	(II)	(III)	(IV)	(V)	(VI)
(A) All	−0.003***	−0.001	0.003**	0.001	0.005***	0.004*
($N = 28,095$)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
(B) Low ability	−0.009***	−0.002	0.005	0.006**	0.010***	0.008**
($N = 11,409$)	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
(C) High ability	−0.000	−0.001	0.001	−0.001	0.001	0.001
($N = 16,686$)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
(D) Low ability, low-SES	−0.014***	−0.001	0.008**	0.007*	0.014***	0.012**
($N = 9,004$)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)
(E) Low ability, high-SES	0.002	−0.003	−0.003	0.003	0.003	0.001
($N = 2,405$)	(0.005)	(0.003)	(0.005)	(0.003)	(0.007)	(0.007)
(F) Low ability, low-SES, $\Delta p \leq 0$	−0.008	−0.005	0.007	0.005	0.012	0.003
($N = 2,964$)	(0.013)	(0.007)	(0.012)	(0.011)	(0.014)	(0.013)
(G) Low ability, low-SES, $\Delta p > 0$	−0.018***	0.001	0.010*	0.007	0.018**	0.014**
($N = 6,040$)	(0.007)	(0.004)	(0.006)	(0.006)	(0.008)	(0.007)
(H) Low ability, low-SES, $\Delta p > 0$, male	−0.023**	0.007	0.009	0.007	0.009	0.012
($N = 2,418$)	(0.010)	(0.007)	(0.010)	(0.008)	(0.013)	(0.012)
(I) Low ability, low-SES, $\Delta p > 0$, female	−0.014	−0.004	0.011	0.007	0.023**	0.016*
($N = 3,622$)	(0.009)	(0.005)	(0.008)	(0.008)	(0.010)	(0.009)

Each cell is from a separate regression of students' outcomes on their number of persistent classmates. All regressions include similar-file fixed effects and quadratic controls for class characteristics. Robust standard errors are reported in parentheses.
Bac data was not available for the last two cohorts, therefore the sample size is smaller for the last two columns.

Table VII: Which peers do matter? Decomposition of the PC effect

	Repeats 10th grade	Drops out	Academic major	Tech. major	Takes *Bac* in time	HS graduate
Independent variable	(I)	(II)	(III)	(IV)	(V)	(VI)
(A) Persistent classmates and persistent schoolmates						
PC	−0.020***	0.002	0.011*	0.007	0.018**	0.014*
	(0.007)	(0.004)	(0.006)	(0.006)	(0.008)	(0.007)
Persistent schoolmates from other classes	−0.006	0.005**	0.001	−0.000	−0.001	−0.002
	(0.004)	(0.002)	(0.003)	(0.003)	(0.005)	(0.004)
(B) Persistent classmates by gender						
Same sex PC	−0.021**	0.000	0.019**	0.002	0.018*	0.015
	(0.009)	(0.006)	(0.009)	(0.009)	(0.011)	(0.010)
Opposite sex PC	−0.015	0.001	0.000	0.013	0.018	0.014
	(0.011)	(0.006)	(0.010)	(0.009)	(0.012)	(0.011)
(C) Persistent classmates by ability						
High-ability PC	−0.009	0.001	0.003	0.005	0.005	0.004
	(0.010)	(0.006)	(0.009)	(0.008)	(0.011)	(0.011)
Low-ability PC	−0.028***	−0.000	0.019**	0.010	0.034***	0.027**
	(0.010)	(0.007)	(0.009)	(0.009)	(0.012)	(0.011)
N	6,040	6,040	6,040	6,040	5,092[1]	5,092[1]

Each cell is from a separate regression of students' outcomes on their classroom characteristics. All regressions include similar-file fixed effects and quadratic controls for class characteristics. Robust standard errors are reported in parentheses.

[1] *Bac* data was not available for the last two cohorts.

Table VIII: Robustness check: effect of PC on low-ability students' retention rate using different specifications of the SF fixed effect

	Specifications				
Independent variable	(I)[1]	(II)[2]	(III)[3]	(IV)[4]	(V)[5]
PC	−0.004**	−0.017***	−0.016***	−0.018***	−0.021**
	(0.002)	(0.006)	(0.006)	(0.007)	(0.010)
R^2	0.67	0.73	0.68	0.68	0.69
N	109,967	12,574	7,559	6,040	2,132
SF students share...					
Options	✓	✓	✓	✓	✓
Middle school	✓	✓	✓	✓	✓
9th grade class	Indifferent	Same	Similar	Same	Same
In-school score	Decile	Decile	Decile	Decile	Decile
Science score		Quintile	Quintile	Quintile	Decile
Humanities score		Quintile	Quintile	Quintile	Decile
Held back	✓	✓	✓	✓	✓
Gender			✓	✓	✓
2-category SES			✓	✓	✓

Sample: at risk students only.
Each cell is from a separate regression of grade retention on PC. All regressions include similar-file fixed effects (different in each column) and quadratic controls for class characteristics. Robust standard errors are reported in parentheses.
Column IV is the original specification, columns I to III are less restrictive and column V is more restrictive.

Table IX: IV exogeneity test

Independent variable	(I)	(II)	(III)	(IV)
Held back	−0.023	−0.024*	−0.016	−0.018
	(0.014)	(0.013)	(0.012)	(0.012)
Girl	−0.013	0.003	0.006	−0.004
	(0.010)	(0.008)	(0.007)	(0.007)
High-SES	0.062***	0.004	0.001	0.009
	(0.010)	(0.008)	(0.007)	(0.007)
High quality optional course	0.041*	−0.005	0.006	0.002
	(0.022)	(0.021)	(0.017)	(0.017)
DNB Quintile 1	−0.051***	0.020	0.005	0.008
	(0.017)	(0.016)	(0.015)	(0.015)
DNB Quintile 2	−0.044***	−0.007	−0.010	−0.000
	(0.016)	(0.015)	(0.014)	(0.013)
DNB Quintile 3	−0.015	−0.007	−0.008	−0.014
	(0.016)	(0.013)	(0.012)	(0.012)
DNB Quintile 4	0.031**	0.021*	0.014	0.013
	(0.014)	(0.013)	(0.011)	(0.011)
DNB Quintile 5	0.035**	0.020	0.001	0.015
	(0.015)	(0.013)	(0.011)	(0.011)
DNB Missing	Ref.	Ref.	Ref.	Ref.
	—	—	—	—
R^2	0.01	0.21	0.42	0.41
N	33,663	33,663	33,663	33,663
Fixed effect	None	High school	HS × 9th grade class	10th grade class

Each column is from a separate regression of the instrument Z on students' characteristics. All regressions include high-school fixed effects. Robust standard errors are reported in parentheses.

Table X: Effect of PC on high school outcomes using the IV strategy

Dependent variable	(I)	(II)	(III)	(IV)
Repeats 10th grade	−0.014***	−0.009***	−0.007**	−0.008***
	(0.003)	(0.003)	(0.003)	(0.003)
Drops out	0.002	0.001	0.002	−0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Academic major	0.029***	0.013***	0.009**	0.011***
	(0.004)	(0.004)	(0.004)	(0.004)
Technological major	−0.018***	−0.004**	−0.004**	−0.003
	(0.002)	(0.002)	(0.002)	(0.002)
N	33,663	33,663	33,663	33,663
Takes *Bac* in time	0.010**	0.009**	0.003	0.009**
	(0.004)	(0.004)	(0.004)	(0.004)
HS graduate	0.025***	0.013***	0.008*	0.012***
	(0.005)	(0.004)	(0.004)	(0.004)
N	26,608	26,608	26,608	26,608
Fixed effect	None	High-school	HS × 9th grade class	10th grade class

Each cell is from a separate regression of students' outcomes on the instrument. Robust standard errors are reported in parentheses.