Adjusting Your Dreams?

High School Plans and Dropout Behavior[1]

Dominique Goux (CREST), Marc Gurgand (Paris School of Economics-CNRS)
and Eric Maurin (Paris School of Economics-EHESS)

July 2014

Abstract

At the end of middle school, many low-achieving students have to abandon hope of getting into selective high-school programs, which may be a source of disengagement and eventually lead them to drop out of high school. Based on a randomized controlled trial, this paper shows that low-achievers can be helped to formulate educational objectives that fit better with their academic aptitudes, through a series of meetings facilitated by the school principals. By affecting high school plans of the less realistic fraction of students, the intervention is able to reduce grade repetition and high-school dropout by 25% to 40%.

[1] This research was supported by a grant from the French Experimental Fund for Youth. We are very grateful for the support of the schools and administrative teams of the *Rectorat de Versailles*. We thank Adrien Bouguen, Bastien Michel and the many J-Pal Europe research assistants that worked on this project. We also thank seminar participants at the Paris School of Economics, London School of Economics, CREST (Paris), Tinbergen Institute (Amsterdam), Louvain-la-Neuve University, Aarhus University, University of Warwick, Austin-Bergen Education and Labor Workshop, Bergen-Stavanger Labor Workshop and the fourth IWAEE conference in Cantazaro. Goux: goux@ensae.fr; Gurgand: gurgand@pse.ens.fr; Maurin: maurin@pse.ens.fr.

1. Introduction

School dropout is a subject of major concern in most developed countries. High-school dropouts experience considerably larger unemployment rates than their better-educated peers and are much more exposed to poverty and delinquency (Belfield and Levin, 2007). In periods of economic stagnation, the gap between dropouts and other young people tends to increase over time, leading to rising polarization and concern for social cohesion.

There are many potential causes of dropout, pertaining to the individual, the school and the social environment (Rumberger and Lim, 2008; Murnane, 2013), but the mechanisms leading to dropout are not completely understood. Much of the economic literature views education as an investment, involving the comparison of immediate costs and anticipated gains. Students drop out from school when the anticipated rewards from staying are too low compared with the financial or psychological costs of doing so (Eckstein and Wolpin, 1999). In this context, dropout may reflect an accurate evaluation of the costs and benefits of staying in education, but it may also be driven by students underestimating the rewards from obtaining further qualifications (Oreopoulos, 2007). Specifically, there is growing evidence that perceived returns to schooling are low compared with actual average returns and that improved knowledge of returns to schooling may help to improve school attendance, at least in developing countries (Nguyen, 2008; Jensen, 2010). On the cost side, Bettinger et al. (2012) show that providing information and assistance on financial aid in the US can also affect educational choices.

In this article we explore another basic cause of early dropout behavior, namely the difficulty for some adolescents and their families to formulate realistic educational objectives. In many developed countries, a uniform schooling system terminates at adolescence, and gives place to a highly stratified system of schools and tracks which typically involves a prestigious academic track and a complex system of vocational schools, with a number of

specializations and locations.[2] Given that only the best students can get access to the most demanded tracks, such a system may be a source of disappointment and disengagement for many students, especially those whose academic results are weak and whose information is incomplete about available options and student assignment mechanisms. Many find themselves obliged to choose among tracks that they never planned to attend, and this may eventually lead them to drop out from education.

This paper reports the results of a large-scale randomized experiment showing that a simple program of meetings facilitated by school principals and targeted at low-achieving students at the end of middle school can help students to identify tracks that fit both their tastes and their academic capacity, the consequence being a very significant reduction in grade repetition and dropout rates in high-school. In 37 middle schools in the suburbs of Paris, mostly in deprived neighborhoods, school principals were asked to preselect the 25% students most exposed to the risk of dropout in ninth grade. At the end of this grade, French students are either allowed to enter an academic three-year high school or asked to submit a choice of four vocational tracks, ranked in order of preference, to a centralized allocation system, where acceptance is based on average academic performance in the ninth grade. Once the lists of preselected students had been completed, we randomly chose about half of the classes, in which the parents of the preselected students were invited by school principals to attend two collective meetings during the second term. During those meetings, principals discuss the specific aspirations of each family in view of the academic performance of their child and, whenever necessary, provide them with specific feedbacks and targeted information on alternative options.

At the end of the treatment year, we find that parents become more involved in schools, more satisfied with information, and have formed educational expectations that fit better with

[2] See OECD (2008) or European commission (2013). Almost half of the high-school students in OECD countries are enrolled in vocational education (OECD, 2008).

the very low academic record of their children. This is reflected in their children's applications at the end of the treatment year: building on exhaustive administrative data, we show that the proportion of students who include at least one 2-year vocational program in their list of possible high-school assignments increases by about 30%, whereas the proportion who wish to repeat (with the aim to access the more selective 3-year programs) decreases by about the same proportion.

This adjustment in the preselected students' applications is followed by very significant shifts in their actual assignments one year after the treatment. Preselected students obtain their demanded tracks more often and end up less often repeating ninth grade or directly dropping out after middle school. Specifically, one year after the treatment, their grade repetition rate is reduced from 13% to 9% and their dropout out rate is reduced from 9% to 5%. Two years after the intervention, we find that treated students were not induced to make choices not suited to them, or to simply postpone dropping out: they are now in the second year of vocational education or apprenticeship in the same proportion as one year before, and there are even less dropouts. As it turns out, by inducing many students to opt for a 2-year vocational program rather than repeating ninth grade, the intervention did not harm their education prospects, but helped to reduce their dropout rates further.

In contrast, the intervention had no impact on the share of preselected students who choose (and end up in) 3-year programs. This important fact shows that principals were able to target their intervention so as not to decrease the aspirations of the best performing students among them.

We cannot identify spillovers among preselected students, but building on the partial population design of our experiment and on detailed information on friendship networks, we explore whether the intervention has an influence on *non-selected* students. Evidence of such spillover effects is only obtained when we focus on the 20% of non-selected students who

have both preselected friends and relatively low academic records: we find that a significant fraction of those with treated friends are induced to enter a vocational high school just after middle school instead of trying to pursue the academic track. These results suggest that students may be influenced by their classmates, but only when they closely interact with them and face similar education alternatives.

Overall, our experiment brings new evidence on the mechanisms leading to dropout and more specifically on the role of students' and families' preference over possible tracks, and expectations on the chances to access them. The social science literature has long emphasized the role of expectations and aspirations in shaping educational attainment, but there is still little evidence on whether they really make a difference, because they are strongly connected to cognitive performance (Jacob and Wilder, 2011). As our intervention generates exogenous changes in expectations and preference without affecting academic records, it provides a unique tool for exploring this issue. Our findings imply that the lack of objectives adjusted to one's school capacity is a source of dropout. A simple intervention facilitated by someone with detailed information on students can have large effects: instrumental variable estimates suggest that a 10 percentage points increase in the proportion of low-achieving students who receive actual feedback on their high-school plans generate a decrease in their dropout rate by about 1.1 percentage points.

These results are in line with the psychological literature, which has long emphasized the idea that unrealized expectations may have devastating consequences for individuals (see e.g., Gottfredson, 1981; Higgins, 1987; Walker and Pettigrew, 1984; Wheaton, 1994). Also, they are consistent with Stinebrickner and Stinebrickner (2013), who observe that over-optimistic expectations at entry into university are associated with higher dropout rates in US colleges.

By showing that a simple intervention facilitated by the school principal can induce a significant fraction of would-be dropouts and would-be repeaters to identify and opt for programs in which they can persevere and pass grades, our experiment also contributes to the literature on dropout prevention policies. Many interventions involve tutoring and academic support (Dynarski et al., 2008), financial incentives (Dearden et al., 2009, Oreopoulos et al., 2009), early childhood interventions (Heckman, 2008) or, more radically, compulsory schooling age (Oreopoulos, 2006 and 2007). Compared to such interventions, the set of meetings considered here are extremely low cost: although the school district had to define guidelines and edit a DVD presenting former students of vocational tracks, the marginal cost of the intervention is very limited, as it simply involves the school principal's time for two meetings and the effort of inviting parents.

It is likely, however, that receiving information from the school principal has played an important role because of his/her credibility and knowledge of individual situations. This adds to the recent understanding of the importance of school principals in education policy. Extending the literature on the role of leaders (Bertrand and Schoar, 2003), Branch et al. (2012) and Duhey and Smith (2013) suggest that school principals have a significant influence on student performance, and Griffith (2001) notes their importance in developing a welcoming school climate, especially towards families from lower socio-economic backgrounds.

We also contribute to the literature on peer effects and social interactions. In particular, we show that the overall effect of a policy intervention may depend on the actual friendship ties between initially targeted individuals and the rest of the population. This finding is consistent with recent research on the impact of the network position of the first individuals to receive information about a new product on the eventual diffusion of that product (Banerjee et al., 2013), and on endogenous peer interactions (Carrell et al., 2013).

The next section presents the institutional context of the experiment and specifically the rules governing track choice at the end of middle school in France. Sections 3 and 4 present the experimental design and data respectively. We then move to measurement of the effects of the intervention on applications at the end of the intervention year (section 5) and on school status one year and two years after the intervention (section 6). Section 7 provides an analysis of spillovers and friendship networks. Section 8 builds on the experiment to provide class-level estimates of the effect of program participation intensity on dropout behavior and section 9 concludes.

2. Track choices at the end of middle school: the institutional context

Middle school in France runs from grade 6 to grade 9. Students complete ninth grade the year of their 15th or 16th birthday (16 also being the minimum school-leaving age), depending on whether or not they have already repeated a grade. The curriculum is the same in all middle schools and there is no streaming by ability.

For ninth graders, a typical week consists of 29 school hours, distributed across 11 different subjects, with a different teacher teaching each subject. Pupils stay in the same class of 20 to 30 pupils throughout the school year and for every subject. The class is therefore a very distinct and closed entity where most of the interactions with same-age children take place.

At the end of the third term, French ninth graders have six basic options for the next year, four within the national education system and two outside it. Those who decide to stay in the school system can either pursue a 3-year academic track in high school or enter a vocational school and pursue either a 3-year or a 2-year vocational program. Both 3-year academic and vocational programs open access to higher education; the 2-year program doesn't. Students are also entitled to repeat ninth grade at least once.

Students who do not stay in the school system can either enter an apprenticeship centre or drop out from education and training. Apprentices can pursue the same 2-year or 3-year programs as students in a vocational school, the difference being that apprenticeship involves periods of on-the-job training with a specific tutor in a specific host company. Apprenticeship centers are funded by the private sector and do not depend on the national education system.[3]

Students who pursue vocational studies have to choose not only a specific training institution (vocational school or apprenticeship centre) and a specific diploma (2-year or 3-year), but also a specific qualification: plumbing, hairdressing, sales, car repair, etc. In the district of Versailles alone (where the experiment took place), there are about 60 different types of 3-year programs, almost the same number of 2-year programs and more than 300 apprenticeship centers.

At the end of the academic year, school principals and teachers decide who will be allowed to pursue the academic track. They base their decisions on students' academic performances during the school year. In 2010, about 60% of French ninth graders nationally were admitted to a 3-year academic program. Students who are not admitted can either ask to repeat ninth grade or apply for a vocational school.[4] In the latter case, they are asked to list up to four specific choices in descending order of preference. Each choice corresponds to a specific qualification (plumbing, hair dressing, etc.) in a specific school. A central assignment system (called *Affelnet*) uses average marks obtained during the ninth grade to rank applications and to assign as many students as possible to one of their listed choices, using a deferred acceptance algorithm (Roth, 2008). The initial outcome of the centralized assignment procedure is known in early July. Assigned students then have a few days to actually register

[3] The French law stipulates that firms with more than 250 employees must either have 4% (or more) apprentices in their labor force or pay additional taxes. The actual rate of apprentices in these firms is only about 1.7% and many employers therefore have to pay the additional taxes.

[4] In parallel, they can also appeal against the school's decision. If they win, they free the slot to which they were assigned by the central system. About 1.5% of ninth graders appeal in this way.

in their new school.[5] After this initial registration period, a small fraction of students remain unassigned and a more informal second round takes place during the summer, in which non-assigned students are asked to reformulate their choices and to apply for tracks that were under-subscribed during the first round.

The final outcome of the assignment procedure is observed at the beginning of the subsequent academic year. In our experimental schools, a large majority of students are admitted into a high-school program, where they start either a 3-year academic (58%), a 3-year vocational (25%) or a 2-year vocational program (3%). Only about 3% enter an apprenticeship centre. Finally, about 6% drop out from education and about 6% opt to repeat the ninth grade. One year later, an additional proportion of students fail to complete the first year of their program and choose to drop out. Two years after the initial assignment procedure, the proportion of dropouts rises to about 13%. For the sake of comparison, the comparable national figure is 7.5% (Ministère de l'éducation nationale, 2011).

Given the limited number of possible applications, students' lists (and final assignments) depend not only on their preference over the different tracks, but also on their initial expectations about the outcomes of the assignment procedure. For example, students who initially expected having no chance to get into a 3-year vocational program may end up applying (and being assigned) to a 2-year program even though it is not their preferred option. In Online Appendix A, we develop a simple track choice model which captures the main features of this institution. This model provides us with a simple tool for interpreting the effects of our intervention on students' applications and assignments in terms of changes in either preference over tracks or expectations about the assignment procedure.

[5]A small fraction of assigned students decide not to register, either because they have chosen to leave the education system (apprenticeship, direct entry into the labor market) or to repeat ninth grade (which is always possible, at each stage of the process) or because they won their appeal against the school. This non-registration frees slots for non-assigned students, some of whom will finally end up assigned to one of their initial choices.

3. Program and experimental design

3.1. School and student selection

The experiment took place in the local education authority (LEA) of Versailles, which includes all suburbs located to the west of Paris, with about 5.5 million inhabitants. It is the largest French LEA, with more than 1.1 million students, or 9% of the French total. In 2010, the LEA launched a preventive program against dropout at the end of middle school. It invited about 50 middle schools to take part in the experiment, of which 37 accepted: they represent 9% of the 400 middle schools of the district. Low-income areas are over-represented in this sample: about two thirds of the volunteer schools are in the 25% poorest neighborhoods in the district (and less than 10% are in the richest 50%). The universe of the experiment is the 4,291 ninth grade students of those 37 middle schools, in 179 classes.[6]

Early in the 2010-2011 academic year, in every ninth grade class, the school principal preselected the students most exposed to the risk of dropping out. In December 2010, the lists were finalized and they contained 1,130 students, representing about 25% of all ninth grade students in the experiment, and about 6 students per class. As much as possible, principals were asked to base their selection on objective data, specifically academic performance. First trimester information confirms that the preselected students are academically weaker than the rest of their peers. Specifically, principals preselected about two thirds of the students in the three lower deciles of the pre-treatment distribution of academic performance, but only a very small proportion of those in the five upper deciles (see Figure B1 in Online Appendix B). Also, half of the preselected students have repeated a grade at some point, compared with 25%, and one third of preselected students are from low-income families (measured through scholarships eligibility), compared with 23%. Nationally, about 25% of families are eligible for scholarship.

[6] Special education students were excluded from the intervention because their orientation towards vocational education had already been determined. Eight students with missing identification numbers (one of them among selected students) were also dropped from the data.

3.2. Randomization

Once the lists of preselected students had been finalized in all schools, we randomized treatment classes within each school. Half of the classes – or half rounded up to the nearest integer when there was an odd number of classes – were put in the treatment group, and randomization was stratified by the number of preselected students, the number of girls and the number of students who had repeated a year in the class. In the end, we have 97 treatment and 82 control classes. Online Appendix B, Table B1, panel A, compares the initial characteristics of the two groups of preselected students, in terms of characteristics for which we have no missing information (parents' income, gender and grade repetition), and there are no significant differences between them.

Only the parents of preselected students in the 97 treatment classes were then invited to attend the meetings facilitated by the school principal. Most of the results in this paper are based on the comparison of preselected students in treated and control classes. Under the basic assumption that preselected students in control classes remain unaffected by the treatment (SUTVA), this comparison provides the estimate of an intention-to-treat parameter, namely the impact of being invited to the meetings on the outcomes of students at risk of dropping out from education. Additional results will be based on the comparison of non-selected students in the two types of classes, so as to identify potential spillover effects on these students.

3.3 Program content

After randomization, the principals invited eligible families to attend two collective meetings about their children's choice of track. This invitation is part of the intervention because it may elicit the feeling among parents that their input is appreciated, which may in

itself improve their involvement. In a previous control trial with sixth graders' families, Avvisati et al. (2011) observed that participation in school meetings was strongly improved when parents were called directly rather than informed about the meeting by letter.

The meetings took place in the school, typically at 6 pm, between January and early April. LEA experts prepared guidelines for the meetings and a DVD showing vocational high-school students sharing their experience. The guidelines explain that the first important objective of the meetings is to make parents understand that important choices have necessarily to be made by the end of the academic year, to help them understand the procedures, and to encourage them to get involved. The guidelines also suggest that the aspirations of each family should be identified and discussed in view of the actual performances of the child. If necessary, families should be provided with information on alternative options and should be helped to adjust their expectations. The principals are also suggested to warn families that grade repetition does not necessarily lead to grade improvement and to illustrate that apprenticeship can be a solution outside of the National Education system, something that is rarely emphasized by internal staff. By contrast, principals are not asked to provide information on the labor market outcomes associated with the different possible tracks and no information is provided on this issue.

Broadly, the program is likely to change the overall value given by students and their families to various options, either because they get to change their perception of, say, vocational education, or because they measure better their own chances to perform well (for instance upon grade repetition). It also makes salient the potential outcomes of the assignment procedure, thus the expected consequences of applying to the different tracks. As mentioned earlier, this is formalized in Online Appendix A.

The cost of this program is mostly related to the conception and production of the guidelines and DVDs. These are largely fixed costs that do not increase with the scale of the

program, as the schools did not receive a specific budget. As the intervention consists of two collective two-hour meetings, the opportunity cost of the school principal's time is limited to a few hours, plus the time taken to contact parents.

3.4. Program take-up

At the beginning of each meeting, the principal asked the families to complete an attendance sheet. Based on this information, Table 1 presents the attendance of four groups of families: preselected students in treatment classes; preselected students in control classes; non-selected students in treatment classes, and non-selected students in control classes. Reassuringly, take-up is only large for the preselected students in treatment classes: about 52% attended one of the meetings and 21% attended both. By contrast, only a tiny fraction of other families attended. As the principals were not required to invite them, it seems that the protocol has been followed.

Given that preselected families are mainly from modest social backgrounds and did not initially volunteer to participate in this program, the take-up rate of over 50% is rather high. In a similar school involvement experiment targeted only at *volunteer* families, Avvisati et al. (2014) obtain a similar participation rate in a comparable social environment. This high take-up rate suggests that the principals have made a genuine effort to convince families to attend.

4. Data

We first collected a number of administrative data from the schools and from the LEA. Data from the schools include a census of students at the start of the academic year 2010-2011 which provides us with baseline demographic and social characteristics. For each student and each term of the academic year 2010-2011, these data also include information on truancy, disciplinary warnings and sanctions (taken at any time by an ad hoc commission) and the

average marks given by teachers in each subject (Maths, French, Physics, etc.). These marks are particularly important in our context because they determine a student's chances of obtaining his or her preferred choices in the centralized allocation system. For all these outcomes, attrition is low (between 5% and 8%) and balanced across treated arms (see Online Appendix B, Table B1, panel B). It mostly reflects the fact that some children change school during the academic year.

We also know the attendance and marks obtained at the end-of-year national exam taken by ninth-grade students (*Diplôme National du Brevet*). This exam is held on the last day of the academic year; it is not compulsory and does not determine allocation to tracks. Some students are absent on exam day, despite the fact that this is very often the moment when students are informed of their track allocation. As we will see, the rate of absence is much higher among future dropouts. Therefore, information about absence on exam days provides us with an early signal of the likelihood of dropout.

We then have exhaustive administrative data from the LEA, providing information about the application and allocation process for each student in our sample. This includes:

(a) Application to the four preferred choices at the end of the treatment year. For each choice, we know the choice rank and the corresponding type of school and program. We also know whether the student appealed or asked to repeat the ninth grade, unless they entered the 3-year academic program.

(b) Students' actual situation one year after the treatment year (i.e., in 2011-212), and two years after the treatment year (2012-2013). We can track each student in the administrative databases based on a national identification number. In particular, we know who is still present in a school or apprenticeship centre at any given date; dropouts can be inferred from this group by deduction. They represent 12% of our students after two years and 20% in our preselected group.

Data on students' applications and subsequent allocation (and on the end-of-the-year exam) are directly obtained from the LEA and do not suffer from attrition.

Lastly, we conducted a series of surveys. In June 2011, parents were asked to complete a questionnaire sent for us by the schools. In early July, in order to increase the response rate, we called non-respondent parents of preselected students, and asked them exactly the same set of questions. Overall, the response rate for the preselected students' families is 75%, balanced over the treatment and control groups (Online Appendix B, Table B1, panel B). We also checked that observed baseline characteristics of the respondents are similar in both groups.

We asked parents questions about their involvement in their children's track choices. Some questions measure how parents reach out for information from school staff (attending general meetings, meeting teachers or career counselors) and how satisfied they are with that information. Other questions relate to information-sharing with other parents (attending parent association meetings, talking with other parents about track choices). A last question measures their educational expectations for their children, specifically the highest secondary education diploma they expect them to obtain.

In order to measure friendship networks, we asked sports teachers to complete a table recording, for each student in the class, up to five best friends within that class. These teachers spend three hours per week with the students and are in a much better position to observe social relations than other teachers. This survey was conducted during the first trimester, with a response rate of about 92%, similar across treatment arms (Online Appendix B, Table B1, panel B). We also verified that baseline characteristics of the respondents are similar in both groups.

5. Outcomes at the end of treatment year

5.1. Parental involvement and expectations

The first basic objective of the program was to increase parents' involvement at school, improve their information about the education system and help them to form more realistic educational expectations for their children. The survey conducted at the end of the treatment year suggests that the intervention was successful in meeting these goals (Table 2). In particular, the survey reveals a strong positive effect on parents' involvement at school, with a 24 percentage points increase in the proportion of preselected parents who participated in information meetings at school. This is the most direct effect of the intervention and it shows that treated families did not simply substitute participation in the program for participation in regular information meetings at school. We also observe a significant impact on the proportion of preselected parents who participated in meetings organized by parent associations (+3.5 percentage points) and who declare interactions with other parents (+9.3 percentage points). Overall, the intervention significantly increases the proportion of preselected parents who are satisfied with the information received from school (+5.6 percentage points).

Preselected parents tend to be more involved and better informed in test classes, but they also tend to form more realistic expectations for their children. In particular, the intervention reduced the proportion of parents expecting that their children will complete a 3-year academic or vocational program, and thereby obtain the *Baccalauréat*, leading to higher education, by about 8 percentage points, with 69% in treatment classes versus 77% in control classes. By construction, the vast majority of preselected students have low or very low academic records and a very small minority can only be expected to complete a 3-year

program.[7] Such overconfidence is not an isolated situation, and it has been recently documented in the US college context (Jacob and Wilder, 2011; Stinebrickner and Stinebrickner, 2013). In our context, a proportion of 69% expecting to complete a 3-year program in treatment classes remains unrealistic, but obviously less so than the 77% observed in control classes.

The intervention does not induce more pessimistic expectations, but more modest ones (+3.4 percentage points increase in the proportion expecting a 2-year vocational degree) and more uncertain ones (+5.4 percentage points say that they do not yet know what to expect). In fact, very few parents expect that their children will drop out from high-school in either treated or control classes. This result is clearly suggestive that high-school dropout is very rarely perceived by families as an optimal choice for children: it is consistent with the fact that the labor outcomes of high-school dropouts are much more problematic than those of individuals who completed additional years of education after middle school. According to the 2011 Labor Force Survey, unemployment rates at entry into the labor market is about 50% for early high-school dropouts whereas it is only about 27% for graduates of a 2-year vocational program and 18% for graduates of a 3-year vocational program (see Online Appendix Table C).

5.2. Performance and behavior at the end of the treatment year

The program was not designed to help pupils improve their performance at school. It is nonetheless possible that it induced an improvement in school performance if only because it

[7] According to the longitudinal administrative database constructed by Ly and Riegert (2013), the probability of completing a 3-year program is about 8.2% nationally for those who fail to pass the national examination at the end of middle school (average marks below 10/20) and 30% for those who pass it without honors (average marks between 10/20 and 12/20). Given that the vast majority of selected students either fail to pass this exam (58%) or pass it without honors (40%), we can expect that only a small minority will complete a 3-year program. Furthermore, based on the control group in our data, we can observe that, two years after the intervention, the proportion of selected students still present in a 3-year program is only about 68% (and only about 49% succeeded in completing the first year). Hence, two years after the intervention, the upper bound for selected students' actual graduation rate in the control group is already about 9% lower than their parent's expected graduation rates at the end of the treatment year (77%).

contributed to a better understanding of the importance of the academic record in the track assignment process.

To test this assumption, Table 3 shows the effect of the intervention on marks obtained during the third term of the treatment year (i.e., the post-treatment term) and on average marks obtained during the treatment year, and specifically the annual average marks used by the assignment software to rank pupils' applications. We do not find any significant effect on either outcome. This is an important result: any effect of the program on track choices can be interpreted as resulting from how tracks are perceived by students and parents, not as an indirect effect of improved marks on students' choice sets. Furthermore, we do not find any significant effect on behavior, as measured by truancy or official sanctions. We only detect a marginally significant improvement in work effort at the end of the treatment year, with a reduction in the number of official warnings given by the pedagogical team for lack of work in treated classes.

Finally, we have information on whether students register for the national examination held at the end of middle school, whether or not registered students were actually present on exam day and whether they passed the exam or not. As discussed above, this exam is not compulsory and the results are not taken into account in the track assignment process (nor at any other subsequent point in the school career). In fact, the results of the assignment process are sent to schools a few days before exam day. Most schools take advantage of the fact that the vast majority of students actually return to school on that day to give them the official results of the centralized assignment process, after the last test. In this context, it is likely that students who are absent on examination day are not interested in their future assignment anymore, and are potential dropouts. As a matter of fact, in the control group, student absence on examination day is very strongly correlated with subsequent dropout: the probability of high-school dropout is about 20 percent points higher for preselected students who are absent

than for those who are present. Overall, this absence can be interpreted as a leading indicator of subsequent decisions to dropout.

Under this assumption, the key question is whether the intervention has any effect on student absence on that day. As it turns out, Table 3 does not show any effect of the treatment on registration rates, but reveals that the proportion of preselected students who were absent on examination day is significantly lower in treated than in control classes (5.2% versus 10.6%). Also, the table shows that increased presence on examination day in treatment classes is not accompanied by any increase in the overall pass rate, which suggests that increased presence rate in treated classes is driven by very low achieving students only. The simplest interpretation is that these students in the treatement group did not want to start the holidays without knowing whether (and where) they were assigned by the educational system. It is a first indication that the program succeeded in convincing them to stay in the school system for at least one more year.

5.3. Track choices

The treatment only has a very weak effect on pupils' academic records. The principal's intervention may nonetheless have a significant impact on the way students and their families perceive the value of different school options at the end of middle school and, consequently, on the tracks to which they choose to apply. Table 4 compares the choices made by preselected students in treatment and control classes at the end of the treatment year.

First, the table confirms that the program has no significant effect on the proportion of students applying for a 3-year academic program (about 17%). This finding is consistent with the fact that the program has negligible effect on academic outcomes and, consequently, negligible effect on the proportion of students allowed to pursue the more academic track.

By contrast, the table reveals a very significant impact of the intervention on the choices made by preselected students who are not allowed to enter a 3-year academic program. Invitation to the program induces an increase in the proportion of preselected students who include 2-year vocational programs in their list of applications (+4.9 percentage points, corresponding to a 30% increase in this proportion) and a symmetrical decline in the proportion who either focus exclusively on a 3-year vocational program or ask to repeat the year (-5.5 percentage points). We observe a decline in the proportion of preselected students who apply only for 3-year vocational programs (-2.5 points) and in the proportion who appeal or ask to repeat the year (-3.0 points). As 2-year programs are less selective than 3-year ones,[8] treated students may have only included them somewhere in their choice list as a strategy to avoid having all their applications rejected. Table 4 suggests that the effect of the intervention is deeper than that. The increase in the proportion of choice lists that include 2-year programs is driven mostly by students who ask for a 2-year program as their first choice (+ 3.8 percentage points, a 34% increase).

Overall, our findings are suggestive that the intervention did not just make low-achieving students more realistic about their chances to get admitted into a 3-year program, but changed the way they actually perceive the different programs, and namely increased the perceived value of entering into a 2-year program compared to that of entering a 3-year one. In the simple track choice model developed in Online Appendix A, we distinguish three types of students: those who prefer entering a 2-year program; those who prefer a 3-year program, but would rather attend a 2-year program than repeat; and those who favor 3-year programs so much that they would repeat if not admitted. The latter are the vast majority of low-achieving students: they do not want to consider 2-year vocational programs as a possible alternative, although they are the least selective and, on average, the least difficult to access. These

[8] In the control group, the difference in average marks between students whose first choice is a 3-year vocational program and those whose first choice is a 2-year vocational program represents about half a standard deviation (i.e., 9.0/20 versus 8.1/20).

students either apply for 3-year programs only or ask directly for grade repetition depending on whether they expect to have any chance to obtain a 3-year program at the end of the current year. In this theoretical set-up, if preferences were not affected by the program and the students were *only* made aware of their low chances to obtain a 3-year vocational track, there would be less students applying for 3-year program only and more students asking for repetition, if anything. The fact that there are both fewer students asking for repetition *and* fewer students applying for 3-year program implies that principals did not simply affect expectations about the outcomes of the assignment procedure, but contribute also to deeper changes in the values given to the different tracks. Specifically, school principals convinced a significant fraction of preselected students that grade repetition was not necessarily a better option than direct entry into a 2-year program and induced them to include 2-year programs in their choice list as a possible high-school assignment. This can result either from lowering the value of repetition by explaining that it may not be productive, or from increasing the value of 2-year programs for instance through apprentices testimony as appeared in the DVD.

6. Assignment outcomes

6.1. Outcomes one year after the intervention

The program has significant effects on students' applications at the end of the treatment year, but it does not necessarily follow that it has an effect on students' assignments for the next academic year. If the students who are induced to modify their applications all belong to the subset of students who actually intend to enter apprenticeship (or dropout), the subsequent effect on assignment is likely to be small. Similarly, if students convinced by school heads to apply for vocational programs fail to be admitted into these tracks, they may finally end up having to choose among the same second-best options as if they had not been treated (typically, repetition or dropout). In this scenario, the impact of the intervention on final assignments would again be much weaker than on initial applications.

To explore these issues, Table 5 focuses on preselected students and shows the effect of the intervention on their actual assignment one year after the treatment. Consistent with our previous results, we do not observe any significant difference across treatment and control classes in the proportion of preselected students enrolled into a 3-year academic program. Most importantly, we do not observe any difference in the proportion of students enrolled into a 3-year vocational program either. As it turns out, one year after the intervention, there is no evidence that principals convinced would-be graduates from a 3-year program to opt for a 2-year one. The principals only diverted would-be dropout and would-be grade repeaters from their initial choices and we only observe significant changes in the distribution assignments among student who do not pursue into a 3-year track.

As such, the intervention is followed by a significant increase in actual enrolment into 2-year high-school programs (+3.3 percentage points, almost doubling the proportion) and a symmetrical decrease in the share who actually repeat ninth grade (-3.5 percentage points, a decrease of about 28%). The increase in enrolment into 2-year programs is driven mostly by students who include this type of track in their initial choice list. By contrast, the decrease in grade repetition is driven mostly by students who did not include this specific option in their choice list: the program induced a 2.5 percentage points fall in these "second-round" grade repetitions. As made explicit in the model in Online Appendix A, these results are suggestive that principals were able to target students with both high and unrealistic expectations. By convincing these students to broaden their choice lists to include less selective 2-year programs, school principals succeeded in increasing the proportion who entered high-school tracks corresponding to their initial choices and thus in decreasing the proportion who ended up repeating.

The intervention also induced a significant increase in the proportion of students who enter apprenticeship and a symmetrical decrease in the proportion of students who drop out of

education and training. One year after the treatment, we observe a 3.7 percentage points fall in the proportion of dropouts in the test group, corresponding to a 43% reduction of this proportion.

As discussed above, apprenticeship is not one of the options that students are asked to rank at the end of ninth grade. Strictly speaking, therefore, we cannot determine whether the increased enrolment in apprenticeship one year after the treatment corresponds to initial choices. However, it is unlikely to be a reaction to the first-round outcome of the assignment procedure. Getting enrolled in a training centre is a long and difficult process and it is highly unlikely that the observed increase in enrolment was driven by decisions made in mid-July (after the first round of the assignment process). Would-be apprentices have to find not only a place in a suitable training centre (i.e., one that provides the qualification they want), but also a sponsoring firm in the relevant industry prepared to hire them as apprentices.[9] After a prospection period during which they send CVs and motivation letters to potential sponsors, students attend hiring interviews in April or May. Many of them start working for their sponsors in July, before the start of the formal training period. In such a context, school principals can only boost successful applications for apprenticeship centers by getting students to invest significant time and effort long before the end of the academic year and the start of the assignment process.

Overall, the treatment provoked significant changes in the distribution of assignment both *within* the school system (more vocational high schools, less grade repetitions) and *outside* the school system (more apprenticeships, less dropouts). As discussed in Online Appendix A, one simple interpretation of these shifts is that they correspond to two different processes, one affecting students intending to pursue secondary education and the other

[9]Apprentices and their sponsoring firms sign a specific 2-year or 3-year labor contract. The wages are fixed by labor laws, and vary between 25% and 50% of the national minimum wage during the first year of training (depending on the age of the apprentices) and between 50% and 75% of the minimum wage during the last year of training.

affecting those intending to leave education. For the first group, the principals convinced students and their families that they should not expect grade repetition or appeal to lead to better outcomes than 2-year vocational education; and that 3-year vocational programs are more difficult to obtain than they might think. As a result of more realistic objectives, more of these students obtain one of their listed choices and less of them repeat the ninth grade. As shown in the next section, this in turn generates lower dropout in the longer run. For the second group, which is the group with the lowest academic level,[10] the principals convinced students and families that apprenticeship is a much better way to get a foothold in the workplace than direct entry into the labor market. Furthermore, it represents a way to obtain additional education that is very different from formal schooling and likely to fit better with their aspirations. Traditionally, schools tend to promote choices within the education system at the detriment of apprenticeship, and it is likely that the image of apprenticeship among teenagers suffers from this. There is thus a significant margin of action for principals to alter this perception.

If this interpretation of the two mechanisms is correct, it implies that the treatment did not induce any potential stayers to leave the school system or vice versa. This has testable implications: the treatment must have no effect either on the proportion of students who chose to stay in the school system or on their pre-treatment characteristics.[11] Any change in size or composition of the group of stayers induced by the treatment necessarily involves something more complex than simple shifts within each group. Online Appendix Table B2 shows that

[10] Students in this second group (i.e. outside the school system) belong to the lower end of our selected students: in the control group, their average mark is 7.2/20 instead of 8.9/20 for the first group (80% of a standard deviation) and the share of former repeaters is 78% instead of 50%. It is likely that the principals had a different approach to their situation.

[11] Equivalently, the treatment must have no effect either on the proportion or on the pre-treatment characteristics of students who chose to leave the school system. Note that other interpretations of our empirical results are possible. For instance, they could also be obtained if the program involved (a) a rise in the value of staying in the school system specifically for those who, in the counterfactual, would have remained outside *and* would have preferred dropping out to apprenticeship; (b) a rise in the value of apprenticeship specifically for those who, in the counterfactual, would have remained inside the school system and would have preferred grade repetition to a two-year program.

these conditions hold true. The proportion of students who chose to stay in the school system is very similar across treatment arms (about 85%) and so are their pre-treatment characteristics. This finding is consistent with the assumption that school heads simply induced potential repeaters to choose 2-year vocational high school and potential dropouts to go into apprenticeship.

6.2. Outcomes two years after the intervention

One of the major effects of the intervention is to induce some of the preselected students not to repeat ninth grade. It may be, however, that these students would have benefited from an additional year in middle school. Symmetrically, the program induced some students to enter apprenticeship (rather than drop out), but these students may be disappointed by this choice, and dropping out may only have been delayed. More generally, it could be that the school principals influenced students' perceptions in favor of choices that were perhaps more realistic in the short run, but did not really fit with their specific aspirations or potential.

The most direct way to test this assumption is to compare grade advancement between test and control students two years after the treatment. If the intervention simply led to a delay in repetitions and dropouts, we should observe much weaker differences in grade advancement and dropout rates two years after the treatment than one year after it. Table 6 shows that this is not the case.

Two years after the treatment, the gap in the proportion of students who completed the first year of their high school program and moved to the second year is no smaller than the initial gap in access to high-school programs across treatment and control groups (+4.4 versus +4.1 percentage points). This is mostly driven by the 2-year vocational program, where the gap in the second year is very similar to that observed in the first year (+3.4 vs. +3.3 percentage points). It suggests that those who were induced to enter this specific track (rather

than repeat ninth grade) did not subsequently experience higher repetition or dropout rates in high school. Meanwhile, the program still didn't have any significant effect on the proportion enrolled in 3-year academic or vocational programs.

A second important finding is that the difference in dropout rates across treatment and control students is even larger than one year after the treatment (-5.1 vs. -3.7 percentage points). Not only does the intervention reduce the proportion of students who drop out just after the treatment year, it also significantly reduces the proportion of students who drop out after repeating ninth grade (-1.5 percentage points), which accounts for the impact increase at the two-year horizon.

Overall, the intervention reduced dropout rates through two different channels. First, it helped already-disengaged middle-school students to define new prospects outside the school system. Specifically, it induced a fraction of would-be dropouts to opt for apprenticeship just after the treatment year, and most of them then succeeded in completing the first year of their training program. This generated a gap in early dropout rates between treatment and control groups which persists over time. Second, the intervention helped low-achieving students still wishing to pursue education to focus on more realistic prospects. In particular, it induced a fraction of would-be grade repeaters to enter a 2-year vocational high-school program and, again, most of them then succeeded in completing the first year of their high-school program, whereas a large proportion would probably have dropped out from school at the end of their repeated ninth grade, had they not been treated. This is a longer-term drop out mechanism that school principals managed to alter.

To further explore the mechanism driving this latter result, we have also analyzed the impact of the intervention on students' academic records at the time they leave middle school. Specifically, we have considered academic outcomes observed at the end of the *last* ninth grade: outcomes are thus measured at the end of the treatment year for those who did not

repeat and one year later for those who repeated. Assuming that the intervention has no direct effect on academic performance (as suggested by Table 3), the difference in marks at the end of the last ninth grade across treatment and control groups identifies the effect of not repeating ninth grade on those induced not to repeat by the treatment. Online Appendix B, Table B3 suggests that this effect is not significant. Similarly, the program has no significant effect either on the proportion of students that enter an academic track or on the proportion that enter a 3-year vocational track at the end of last ninth grade. These findings suggest that a proportion of grade repetitions at the end of middle school has no impact on students' academic records at entry into high school, which is consistent with earlier findings obtained in different institutional contexts (Jacob and Lefgren, 2009; Manacorda, 2012).[12]

The table also reveals that the gap in dropout rates between test and control students tends to be larger at the end of the second ninth grade than after the treatment year. In fact, the gap in dropout rates observed at the end of the last ninth grade is almost as large as the gap observed two years after the treatment (-4.8 vs. -5.1 percentage points). Hence, most of the increase in this gap between year 1 and year 2 seems to be driven by the fact that a large proportion of students induced not to repeat ninth grade at the end of year 1 would have dropped out of school just after the end of their second ninth grade, had they not been treated.

7. The role of friendship networks

Peer pressure has long been identified as a potential determinant of pupils' perceptions and choices.[13] It is, however, very difficult to provide robust evidence on whether peers really exert a causal influence on pupils' decisions. Progress in this direction has been limited by the difficulty of observing independent variation in the influence exerted by peers, as pupils

[12] In our case, this is true of the "compliers" whose preference for repetition was affected by the program.

[13] For recent contributions to the large literature on peer group influence on students' behavior, see e.g. Avvisati et al. (2014), Card and Giuliano (2013), Kremer and Levy (2008). Recent research in cognitive science also suggests that the brain regions involved in considering both the long term consequences of behavior and peer opinions develop most rapidly during adolescence (Mc Clure et al., 2004; Blakemore, 2008).

within the same class are generally subject to similar influences. In this section, we exploit the "partial population" design of our experiment to overcome this issue and to provide estimates on how independent changes in the aspirations of preselected classmates affect non-selected pupils.

The first two columns of Table 7 focus on the full sample of non-selected students and provide estimates of the indirect effects of the intervention on their average marks and applications at the end of the treatment year. Specifically, for each possible track choice, we show the impact of being in a treatment class on the probability of including this choice in the list of applications. We do not find any significant effect on either average marks or choices.

One possible reason why we do not detect spillover effects on non-selected students is because students do not in general interact with all their classmates, but mostly with a subset of friends. According to our baseline friendship network survey, a large proportion of non-selected students (about 60%) have no friends among the preselected students: the absence of spillover effects on non-selected students thus comes as no surprise.[14]

To further explore this assumption, columns 4-6 of Table 7 provide a separate analysis of spillover effects on the subgroup of non-selected students who have at least one friend among their preselected classmates, as measured in the baseline survey. We still find no spillover effects on marks, but we can now detect some spillover effects on applications, although not significant at standard levels. Another reason why spillover effects on non-preselected students may be relatively weak is that the majority of non-selected students has a good academic level and is not at risk of being denied the academic track. For a large fraction of non-selected students, there is in fact little room for peers to have any influence on their choices. To address this issue, the last three columns of Table 7 further focus on the 20% or so among non-selected pupils with some preselected friends, and whose academic level is

[14] Some descriptive statistics on friendship ties are given in Appendix B, table B4. Consistent with a long-standing literature on friendship networks, we find that pupils are very similar to their friends in terms of gender or academic status (see e.g., Shrum et al., 1988; Tuma and Hallinan, 1979).

relatively low compared with other non-selected pupils (those whose average marks during the pre-treatment term were12/20 or below). These mid-ability pupils are the ones for whom the question of choosing between general and vocational tracks is most likely to remain open until the very end of the year.[15] When we replicate our analysis on this subset of non-selected pupils, we detect a significant negative effect on the proportion of students applying for an academic track (-7.7 percentage points) and a symmetrical positive effect on the proportion applying for a vocational track (+8.7 percentage points). These results clearly suggest that pupils' choices may be influenced by their friends, especially when they have a similar academic level and face similar education alternatives.[16]

Table 8, upper panel, demonstrates that these spillovers on choices translate into similar significant spillovers on assignment one year after the treatment. When we focus on the group of non-selected students with both preselected friends and relatively low marks, we observe that they have been induced to enter a vocational track rather than an academic track by about 9.2 percentage points. However, when we replicate the same analysis *two* years after the treatment, these spillover effects tend to fade out (Table 8, lower panel). At this point in time, the intervention is still associated with a higher proportion of non-selected students in vocational tracks, but this effect is about 30% lower than one year after the treatment (+7.0 percentage points versus +9.2 percentage points) and no longer significant at standard levels. This finding suggests that a proportion of the students who have been induced by their preselected friends to enter a vocational track rather than an academic one at the end of middle school, would have moved to this type of track anyway before the end of high school, even if they had not been treated.

[15] Within this sub-group, the proportion of pupils who enter the general track is actually only about 40% (in the control group), i.e., not larger than the proportion who enter a vocational track. By contrast, within the subgroup whose average marks are above 12/20 during the first term, the proportion who enter the general track is about 90%.

[16] If there were significant friendship ties across classes so that selected students in treatment classes influenced non-selected students in control classes, this would tend to bias our spillover estimates towards zero.

Spillover effects on non-selected students' choices may be driven by treated students being able to pass on information received from the principal to their non-selected friends. Alternatively, these effects may be driven by the desire of non-selected students to attend the same schools as their preselected friends.[17] It is typically very difficult to provide robust evidence on the channels through which eligible individuals affect their non-eligible peers. However, using schools' identification numbers, we have checked that the treatment has no effect on the probability that non-selected students apply for (or enter) the same schools as their friends, even when we focus on the 20% of non-selected students with preselected friends and relatively low marks. This result suggests that spillovers on non-selected students are not driven by the desire to enter the same school as friends, but reflect a deeper influence.

The program generated significant spillover effects on non-selected students whose academic level was just above that of preselected students. It is therefore likely that the program also generated spillover effects *within* the group of preselected students. Preselected students being exposed to more irreversible decisions than non-selected ones (such as the decision to drop out), it is even likely that these spillover effects have been more persistent than those on non-selected students. Unfortunately, spillover effects on eligible individuals cannot be robustly identified in a set-up like ours, where there is no random variation in the proportion of eligible individuals across experimental units (Baird et al., 2012).[18]

Building on our longitudinal information on friendship ties, we are nonetheless able to explore whether the intervention induced preselected students to have more or less interactions with specific classmates during the treatment year. As recently emphasized by Carrell et al. (2013), the impact of public policy interventions on network structure may be a

[17] Vocational and academic education programs are usually provided in different high schools. Hence, pupils who wish to attend the same establishment as friends who have opted for a vocational track must also enter a vocational track, though not necessarily of the same kind.

[18] We have checked that the impact of the intervention on selected students' propensity to dropout is significantly stronger for those who have some selected friends than for those who only have non-selected friends. This result is clearly consistent with the existence of spillover effects across selected classmates, but it may also be because the same unobserved factors explain both selected students' propensity to have selected friends and students' responsiveness to the intervention.

channel through which interventions can affect behavior. Recent research also shows that network stability per se may be a source of improvement in school outcomes (Lavy and Sand, 2012). Reduced dropout rates in test classes could also be partly the consequence of more stable friendship networks and the increased school integration of preselected students in these classes.

To explore these issues, Online Appendix B, Table B5 focuses on the sample of preselected students for whom we observe friendship networks both before and after treatment[19] and compares the changes in network size and composition in test and control classes over the treatment year. It shows that one effect of the treatment was to increase the stability of the friendship network (less lost friends and less new friends over the year). Another effect is that whatever changes in friends that occur decreases the proportion of future dropouts among friends, and increases the proportion of friends who will be enrolled in a 3-year program, when the student is in the treatment group. In contrast, it doesn't increase friendship ties among preselected students.

Overall, by adjusting and harmonizing aspirations, the principal's intervention seems to protect social ties from both disruption and dispersion. It also increases interactions between students who were at risk of dropping out and students who are going to enter selective high-school programs. Given previous evidence on spillovers, it is likely that these students influence each other and maintain their friendship ties as their aspirations converge. One way in which the principal's intervention may have been amplified is through the strengthening and qualitative improvement of social ties.

[19]The sample in this analysis is constrained by the roughly 45% response rate to the endline network survey. Appendix B, Table B1 provides evidence that the selection in the sample used for Appendix B, Table B5 is ignorable: the treatment has no effect on the probability of being selected in this sample. Further, we checked that the estimated direct effects of the treatment on students' behavior are similar in this specific sample to the full sample of selected students.

8. Treatment on the treated parameter

While the previous sections have focused on the reduced-form effects of invitations to meetings (intention-to-treat analysis), in this section we use invitations as a source of identification for the effects of actual participation to meetings on subsequent outcomes (treatment-on-treated analysis). Specifically, we focus on preselected families, and evaluate the extent to which their actual participation to meetings (and their benefiting from principals' specific feedback) is followed by a reduction in their children's dropout rates.

With respect to the exact channels through which meetings affect outcomes, we cannot distinguish between direct and spillover effects, since there is no random variation in the proportion of eligible peers across classes or friendship networks. The exclusion restriction would thus not hold at the individual level. Given this, we define outcomes at the aggregate (class) level and thus provide estimates of the impact of average participation to meetings on average class dropout behavior of preselected students. Specifically, we assume the following class level model:

$$Y_c = \theta P_c + \delta X_c + \varepsilon_c,$$

where Y_c represents the average educational outcome of preselected students in class c, P_c represents the proportion of preselected families in class c who actually participate to meetings, X_c is an average of baseline control variables, and ε_c denotes unobserved random characteristics. In this model, assuming that treatment status T_c is independent from both observed and unobserved characteristics (X and ε) and affects outcomes through P_c, we can use T_c as an instrumental variable to obtain robust identification of parameter θ, the causal effect of preselected families' participation to meetings on their subsequent average educational outcomes.[20]

[20] This identification hypothesis excludes the possibility that invitations to the meetings have an impact irrespective of actual participation. In that sense, the ITT estimates presented in other sections are more robust; The TOT parameter quantifies the impact more precisely, but under this restriction.

This class level model could be derived from a standard linear-in-means individual-level model (Manski, 1993) where preselected students interact in groups within classes and individual outcomes are affected by own participation, the average level of participation and average outcomes of preselected peers. In this set-up, parameter θ represents the sum of direct and indirect effects of participation, inflated by the social multiplier.

Table 9 shows the results of this class-level regression analysis. We consider the two most important outcomes of the experiment: dropout at the end of the treatment year and dropout one year later. The first stage is very significant and reflects the take-up rate. Consistent with the linear-in-means assumptions, we obtain very similar reduced-form estimates at the class level as we previously do at the student level.

The instrumental variable results imply that a 10 percentage points increase in parent's participation to the meetings would decrease immediate dropout by 0.76 percentage points (when it is 8.8% in the control group) and longer-term dropout by 1.06 percentage points (when it is 20% in the control group).

Overall our results are suggestive that students' and families' choices are strongly responsive to specific feedbacks and information provided by principals, which may explain how a relatively simple and inexpensive program can substantially affect dropout rates. As suggested earlier, facilitation by the school principals must have made this intervention particularly efficient, whereas providing information as such may not have been sufficient. For instance, experiments by Bettinger et al. (2012) and Hoxby and Turner (2013) show that providing information on aid or admission into college is only efficient when direct help or administrative simplification accompanies it.

By adjusting preferences and expectation on assignment outcomes, an extension of the program would be likely to increase the overall proportion of low-achieving students who are willing to enter tracks better suited to their actual academic abilities. Hence, an extension of

the program has the potential to significantly reduce overall dropout rates and potentially increase future wages, provided that it is accompanied by a parallel increase in school supply. Given that potential dropouts would be induced to enter vocational tracks at the bottom of the track scale, an extension of the program would not generate negative spillover effects on students who enter the more selective tracks.

9. Conclusion

Based on a randomized controlled trial, this paper documents aspects of schooling decisions and dropout behavior that have received limited attention. Largely irreversible qualitative choices have to be made at some point in most educational systems, and we consider a decisive track assignment process that takes place at the end of middle school in France. We observe that many low-performing students in mostly deprived areas have unsuitable expectations in the face of a complex choice set and assignment system. Many either undervalue vocational education or overestimate their chances of entering the academic system, and therefore fail to consider less selective programs as possible options. This leaves room for intervention and we show that even a non-intensive, inexpensive treatment facilitated by the school principals can adjust those students' and their parents' preferences and their expectations about the outcome of track choices. Because expectations and aspirations are important determinants of school outcomes, this has substantial effects on the educational situation of these students: dropout is reduced by 25% in this target population (from 20% to 15%), in part through a decrease in ineffective grade repetition, in favor of vocational education. In contrast, the intervention did not affect those students whose aspirations were more in line with their academic capacity.

Overall, this experiment shows that it is possible to influence adolescent school choices in a way that is likely to improve their lifelong outcomes at little direct cost. We interpret this

in relation to two features: a gap between education plans and capacity is a source of negative school outcomes, specifically dropout; and preferences and expectations are malleable, the school principals being able to influence them significantly for the relevant margin of the population, without altering the plans of better performers. Whereas the literature has long considered the low aspirations of high-performing, low-background students, this paper show that the educational outcomes of low-performing students can also be improved under the same approach.

References

Avvisati F., Gurgand, M., Guyon N. and Maurin E. (2014), "Getting Parents Involved: a Field Experiment in Deprived Schools", *Review of Economic Studies,* vol. 81, pp. 57-83.

Avvisati F., Gurgand, M., Guyon N. and Maurin E. (2011), "Communication des collèges et implication des parents d'élèves", *Rapport pour le Fonds d'expérimentation pour la jeunesse*.

Baird S., Bohren A., McIntosh C. and Ozler B. (2012), "Designing experiment to measure spillover effects and thresholds effects", Unpublished Manuscript.

Banerjee A., Chandrasekhat A., Duflo E. and Jackson M. (2013), "The diffusion of microfinance", *Science*, vol. 314, 1236498-1.

Belfield C. and Levin H. (2007), *The price we pay: economic and social consequences of inadequate education,* Washington, DC: The Brookings Institution.

Bertrand, M. and Schoar A. (2003), "Managing with Style: The Effect of Managers on Firm Policies." *Quarterly Journal of Economics,* vol. 118, pp. 1169-1208.

Bettinger E., Long B., Oreopoulos P. and Sanbonmatsu L. (2012), "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment", *Quarterly Journal of Economics,* vol. 127, pp. 1205-1242.

Blakemore S.-J. (2008), "The Social brain in adolescence", *Nature Reviews Neuroscience*, vol. 9, pp. 267-277.

Branch G., Hanushek E. and Rivkin S. (2012), "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals", NBER Working Paper 17803.

Card D. and Giuliano L. (2013), "Peer Effects and Multiple Equilibria in the Risky Behavior of Friends", *The Review of Economics and Statistics*, vol. 95, pp 1130-1149.

Carrell S., Sacerdote B. and West J., (2013), "From natural variation to optimal policy? The importance of endogenous peer group formation", *Econometrica*, vol. 81, pp. 855-882.

Dearden L., Emmerson C., Frayne C. and Meghir C. (2009) "Conditional Cash Transfers and School Dropout Rates", *Journal of Human Resources.* vol. 44, pp. 827-857.

Dhuey E. and Smith J. (2013), "How Important Are School Principals in the Production of Student Achievement?", mimeo.

Dynarski M., Clarke L., Cobb B., Finn J., Rumberger R. and Smink J. (2008), *Dropout Prevention: A Practice Guide* (NCEE 2008–4025). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Eckstein Z. and Wolpin K.I. (1999), "Why Youth Drop out of High School: The Impact of Preferences, Opportunities and Abilities", *Econometrica,* vol. 67, pp. 1295-1339.

European Commission (2013), "The Structure of European Educational Systems 2012-2013: Schematic Diagrams", *Eurydice Highlights*.

Gottfredson L. (1981), "Circumscription and compromise: a developmental theory of occupational aspirations", *Journal of Counseling Psychology*, vol. 28, pp. 545-579.

Griffith J. (2001), "Principal leadership of parent involvement", *Journal of Educational Administration*, vol. 39, pp. 162-186.

Heckman J.J. (2008), "Schools, Skills, and Synapses", *Economic Inquiry*, vol. 46, pp. 289-324.

Higgins E. (1987), "Self-discrepancy: a theory relating self and affect", *Psychology Review*, vol. 94, pp. 319-340.

Hoxby C. and Turner S. (2013), "Expanding college opportunities for high-achieving low income students", SIEPR Discussion Paper no 12-014.

Jacob B. and Lefgren L. (2009), "The Effect of Grade Retention on High School Completion", *American Economic Journal: Applied Economics*, vol. 1, n°3, pp. 33-58.

Jacob, B. and Wilder T. (2011), "Educational Expectations and Attainment", in *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children,* edited by G. J. Duncan and R. J. Murnane, New York, NY: Russell Sage Press.

Jensen, R. (2010), "The (Perceived) Returns to Education and the Demand for Schooling", *Quarterly Journal of Economics*, vol. 125, pp. 515-548.

Kremer M. and Levy D. (2008), "Peer Effects and Alcohol Use among College Students", *Journal of Economic Perspectives*, vol. 22, pp. 189-206.

Lavy V. And Sand E. (2012), "The Friends factor: how students' social networks affect their academic achievement and well-being?", NBER Working Paper 18430.

Ly S.T. and Riegert A. (2013), "Persistent Classmates: How Familiarity with Peers Protects from Disruptive School Transitions", Working paper 2013-21, Paris School of Economics.

Manacorda M. (2012), "The cost of grade retention", *Review of Economics and Statistics*, vol. 94, pp. 596-606.

Manski C. (1993), "Identification of Endogenous Social Effects: The Reflection Problem", *The Review of Economic Studies*, vol. 60, pp. 531-542.

Mc Clure S., Laibson D., Loewenstein G., and Cohen J. (2004), "Separate Neural Systems Value Immediate and Delayed Monetary Rewards", *Science*, vol. 306, pp. 503-507.

Ministère de l'éducation nationale (2011), *Géographie de l'école*, no 10.

Murnane R. (2013), "US High School graduation rates: patterns and explanations", *Journal of Economic Literature*, vol. 60, pp. 370-422.

Nguyen T. (2008), "Information, Role Models and Perceived Return to Education: Experimental Evidence from Madagascar" *unpublished manuscript*.

OECD (2008), *Education at a glance 2008: OECD Indicators,* OCDE, Paris.

Oreopoulos P. (2006), "Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter," *American Economic Review*, vol. 96, pp. 152-175.

Oreopoulos P. (2007), "Do Dropouts Drop Out too soon? Wealth, health and happiness from compulsory schooling", *Journal of Public Economics*, vol. 91, pp. 2213-2229.

Oreopoulos P., Lang D. and Angrist J. (2009), "Incentives and Services for College Achievement: Evidence from a Randomized Trial" *American Economic Journal: Applied Economics*, vol. 1, pp. 136-63.

Roth A. E. (2008), "Deferred Acceptance Algorithms: History, Theory, Practice, and Open Questions", *International Journal of Game Theory*, vol. 36, pp. 537-569.

Rumberger R. W. and Lim S. A. (2008), "Why students drop out of schools: a review of 25 years of research", *California Dropout Research Project,* Report n°15.

Shrum W., Cheek Jr. N., and Hunter MacD. S. (1988), "Friendship in School: Gender and Racial Homophily", *Sociology of Education*, vol. 61, pp. 227-239.

Stinebrickner, T. R. and Stinebrickner R. (2013), "Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model," forthcoming, *Journal of Labor Economics*.

Tuma N. B. and Hallinan M.T. (1979), "The Effects of Sex, Race and Achievements on Schoolchildren's Friendship", *Social Forces*, vol. 57, pp. 1265-1285.

Walker I. and Pettigrew T. (1984), "Relative deprivation theory: an overview and conceptual critique", *British Journal of Social Psychology*, vol. 23, pp. 301-310.

Wheaton B. (1994), "Sampling the stress universe", in *Stress and mental health: contemporary issues and prospects for the future*, edited by W. Avison and I. Gotlib, New York, Plenum Press, pp. 77-114.

Table 1: Participation in the program (in %)

	Preselected students		Non-selected students	
	Test	Control	Test	Control
… first meeting	45,5	2,5	1,3	0.1
… second meeting	27.7	0.4	1.3	0.0
… two meetings	21.0	0.4	0.7	0.0
… one or two meetings	52.2	2.5	1.9	0.1
Number of observations	600	510	1 662	1 415

Note: The sample includes all students, except those from one school who did not complete the presence sheets.

Table 2: Treatment effects on parental involvement and expectations, preselected students

	C	T-C	s.e.	Obs.
Information from school:				
General information meetings	16.0	+24.2**	3.0	836
Individual meetings with career counselor	22.7	+1.5	2.7	837
Individual meetings with teachers	40.6	+2.2	3.7	842
Interaction with other parents:				
Has attended meetings of parent association	9.0	+3.5	2.1	834
Has talked with other parents	43.8	+9.3**	3.2	831
Satisfaction:				
Happy with school information	53.3	+5.6*	3.1	835
Expected diploma:				
Baccalauréat (3-year track)	77.5	-8.2**	2.8	830
Vocational certificate (2-year track)	10.3	+3.4*	2.0	830
No diploma	0.5	-0.6*	0.3	830
Do not know	11.6	+5.4**	2.2	830

Note: Each row corresponds to a different model, based on parent's endline survey. Column (C) shows the average response of preselected students' parents in the control group. Column (T-C) contains the coefficient of a treatment class dummy. We use a linear probability model where we control for school dummies, first trimester average marks, gender, currently repeating dummy, low income dummy, as well as dummies for missing first trimester, and missing repeater information. Column (se) shows corresponding standard error clustered at the class level. *: significant at the 10% level; **: at the 5% level.

Table 3: Treatment effects on behavior in school and academic performance at the end of the treatment year, preselected students

	C	T-C	s.e.	Obs.
• *Academic performance*				
Average marks, term 3 (/20)	8.44	-0.00	0.11	1 047
Annual average marks (/200)	86.5	+0.7	0.9	1 097
• *Behavior*				
Warning work effort	26.7	-4.7	3.0	1 074
Suspension, term 3	10.7	-0.2	1.5	1 070
Truancy (number half days)	16.0	-0.5	0.8	1.031
• *End of middle school examination*				
Not registered	4.2	-1.1	1.0	1 130
Fail	55.4	+1.3	2.3	1 130
Fail. but present on exam day	44.8	+6.7**	2.5	1 130
Not present on exam day	10.6	-5.4**	1.3	1 130
Pass	40.4	-0.2	2.3	1 130
Honors	1.9	+1.2	1.0	1 130

Note: Each row corresponds to a different model, based on administrative data. Column (C) shows the average value for the preselected students in the control group. Column (T-C) contains the coefficient of a treatment class dummy. We use a linear probability model where we control for school dummies, first trimester average marks, gender, currently repeating dummy, low income dummy, as well as dummies for missing first trimester, and missing repeater. For behavior variables, the regression also includes a control for first trimester value. Column (se) shows corresponding standard error clustered at the class level. *: significant at the 10% level; **: at the 5% level.

Table 4: Treatment effects on applications at the end of the treatment year, preselected students

List of applications includes:	C	T - C	s.e.
At least one 2-year vocational program	15.8	+4.9**	1.9
2-year vocational is first choice	*11.0*	*+3.8***	*1.7*
2-year vocational is not first choice	*4.8*	*+1.1*	*1.1*
No 2-year vocational or 3-year academic programs	61.0	-5.5**	2.6
Only 3-year vocational programs	*50.6*	*-2.5*	*2.7*
Repetition or appeal	*10.4*	*-3.0**	*1.6*
3-year academic program	16.7	+0.1	1.8
Other cases (private; other districts)	6.5	+0.5	1.1

Note: Each row corresponds to a different model, based on administrative data. Column (C) shows the average value for the preselected students in the control group. Column (T-C) contains the coefficient of a treatment class dummy. We use a linear probability model where we control for school dummies, first trimester average marks, year-long average marks, gender, currently repeating dummy, low income dummy, as well as dummies for missing first trimester and year-long marks, and missing repeater. The number of observations used is N=1130. Column (se) shows corresponding standard error clustered at the class level. *: significant at the 10% level; **: at the 5% level.

Table 5: Treatment effect on assignment one year after treatment, preselected students

Status one year after intervention	C	T - C	s.e.
• Within national education			
3-year academic	18,5	-0,0	1,8
3-year vocational	50,4	+0,8	2,9
Repetition	12,7	-3,5**	1,6
Repetition or appeal in the choice list	*6,5*	*-1,0*	*1,3*
Repetition and appeal not in the list	*6,2*	*-2,5***	*1,0*
2-year vocational	3,8	+3,3**	1,1
2-year vocational in the choice list	*3,5*	*+2,1***	*1,0*
2-year vocational not in the list	*0,4*	*+1,2***	*0,5*
• Outside national education			
Apprenticeship	5,8	+3,1**	1,4
Dropout	8,8	-3,7**	1,1

Note: Each row corresponds to a different model, based on administrative data. Column (C) shows the average value for the preselected students in the control group. Column (T-C) contains the coefficient of a treatment class dummy. We use a linear probability model where we control for school dummies, first trimester average marks, year-long average marks, gender, currently repeating dummy, low income dummy, as well as dummies for missing first trimester and year-long marks, and missing repeater. The number of observations used is N=1130. Column (se) shows corresponding standard error clustered at the class level. *: significant at the 10% level; **: at the 5% level.

Table 6: Treatment effects on assignment two years after treatment, preselected students

Status two years after intervention	C	T – C	s.e.
First year completed	52.7	+4.4*	2.6
3-year academic (2nd year)	*10.8*	*+1.0*	*1.6*
3-year vocational (2nd year)	*38.7*	*-0.0*	*2.7*
2-year vocational (2nd year)	*3.3*	*+3.4***	*1.0*
First year still not completed	20.0	-1.9	2.2
3-year academic (1st year)	*9.0*	*-1.0*	*1.5*
3-year vocational (1st year)	*9.2*	*-0.9*	*1.6*
2-year vocational (1st year)	*1.7*	*-0.0*	*0.7*
Apprenticeship	7.3	+2.3	1.4
Dropout	20.0	-5.1**	1.9
dropout in year 1	*7.5*	*-3.9***	*1.0*
repeating 9th grade in year 1	*2.3*	*-1.5***	*0.6*
others	*10.2*	*+0.3*	*1.7*

Note: Each row corresponds to a different model, based on administrative data. Column (C) shows the average value for the preselected students in the control group. Column (T-C) contains the coefficient of a treatment class dummy. We use a linear probability model where we control for school dummies, first trimester average marks, year-long average marks, gender, currently repeating dummy, low income dummy, as well as dummies for missing first trimester and year-long marks, and missing repeater. The number of observations used is N=1130. Column (se) shows corresponding standard error clustered at the class level. *: significant at the 10% level; **: at the 5% level.

Table 7: Spillover effects on non-selected students: marks and applications at the end of the treatment year

	All			Some friends preselected			Some friends preselected and pre-treatment marks<=12		
	C	T-C	Obs.	C	T-C	Obs.	C	T-C	Obs.
• Average marks (annual)	124.1	-0.8 (0.5)	2 913	117.4	-0.2 (0.7)	1 183	96.7	+0.4 (1.1)	512
• Applications									
Appeal or repetition	3.1	+0.9 (0.7)	2 972	4.3	+0.2 (1.1)	1 208	9.2	-1.1 (2.3)	528
3-year general	73.2	-1.0 (1.4)	2 972	67.7	-3.6 (2.3)	1 208	38.7	-7.7(3.6)**	528
3-year vocational	22.2	+0.0 (1.4)	2 972	28.4	+2.4 (2.3)	1 208	50.8	+8.7 (3.8)**	528
2-year vocational	6.2	-2.0 (0.9)**	2 972	6.2	+0.3 (1.3)	1 208	13.0	-2.2 (2.6)	528
Others	5.1	+0.1 (0.6)	2 972	3.4	+0.8 (0.8)	1 208	5.5	+0.7 (1.4)	528

Note: Each row corresponds to a different model, based on administrative data. Column (C) shows the average value for the non-selected students in the control group. Column (T-C) contains the coefficient of a treatment class dummy, with corresponding standard-error in parenthesis (clustered at the class level). Friends are measured as of baseline. We use a linear probability model where we control for school dummies, first trimester average marks, year-long average marks, gender, currently repeating dummy, low income dummy, as well as dummies for missing first trimester and year-long marks, and missing repeater. *: significant at the 10% level; **: at the 5% level.

Table 8: Spillover effects on non-selected students: assignment one and two years after treatment

	All		Some friends preselected		Some friends preselected and pre-treatment marks<=12	
	C	T-C	C	T-C	C	T-C
• Status one year after treatment						
Repetition	2.7	+0.5 (0.6)	3.9	-0.5 (1.0)	8.8	-1.3 (2.3)
3-year academic track	73.6	-2.0 (1.4)	67.9	-4.1 (2.3)*	39.9	-7.7 (3.4)**
Vocational tracks	19.5	+0.8 (1.5)	24.1	+3.8 (2.2)*	45.0	+9.2 (3.9)**
Dropouts	4.2	+0.8 (0.7)	4.1	+0.7 (1.1)	6.3	-0.1 (1.9)
• Status two years after treatment						
3-year academic track	70.3	-1.0 (1.5)	65.7	-3.2 (2.4)	38.2	-4.5 (3.8)
Vocational tracks(1)	20.6	+1.7 (1.5)	26.9	+2.8 (2.3)	49.6	+7.0 (4.1)*
Dropouts	9.1	-0.7 (0.9)	7.5	+0.4 (1.4)	12.2	-2.5 (2.4)
Obs.		2 972		1 208		528

Note: Each row corresponds to a different model, based on administrative data. Column (C) shows the average value for the non-selected students in the control group. Column (T-C) contains the coefficient of a treatment class dummy, with corresponding standard-error in parenthesis (clustered at the class level). We use a linear probability model where we control for school dummies, first trimester average marks, year-long average marks, gender, currently repeating dummy, low income dummy, as well as dummies for missing first trimester and year-long marks, and missing repeater. *: significant at the 10% level; **: at the 5% level.
(1) includes second repetition of 9th grade.

Table 9: Class-level IV estimation of the effect of the proportion of parents attending meetings on preselected students' dropout rates

Panel A	Proportion parents participating in 1 or 2 meetings	Proportion of students who dropout one year after treatment	
	(First stage)	(Red. form)	(IV)
Treatment	0.506**	-0.038**	-
	(0.031)	(0.014)	
Proportion parents participating in one or two meetings	-	-	-0.076**
			(0.027)
Panel B	**Proportion parents participating in 1 or 2 meetings**	**Proportion of students who dropout two years after treatment**	
	(First stage)	(Red. form)	(IV)
Treatment	0.506**	-0.053**	-
	(0.031)	(0.024)	
Proportion parents participating in one or two meetings	-	-	-0.106**
			(0.048)
Obs.	174	174	174

Note: Panel A and B show the instrumental variable (IV) regression of the proportion of preselected students who drop out from school one year after the treatment (panel A) or two years after the treatment (panel B) on the proportion of preselected students whose parents took part in the meetings using treatment status as an instrument. This is estimated on preselected students at the class level to account for spillovers. The first column shows the results of the first-stage regression of the proportion of parents who took part in the meetings on the instrument, whereas the second column shows the reduced form regression of the dependent variable (dropout) on the instrument (treatment status). In all regressions, we control for school fixed effects as well as for a set of class-level variables describing preselected students' baseline characteristics: proportion of boys, proportion from low income families, proportion of grade repeaters, first term average marks, annual average marks, proportion with annual average marks less than 10/20 (as well the proportions with missing information on first-term average marks, annual average marks or grade repetitions). *: significant at the 10% level; **: at the 5% level.

ONLINE APPENDIX

Online Appendix A

A model for applications and assignments

We assume that students can apply for either a 3-year program (denoted *H*), or a 2-year program (*L*) or ask for repetition (*R*). They can rank one or more tracks on a list of applications. Consistent with French institutions, we also assume that students can get access to *H* only if their average mark *m* is above a threshold m_0 that is revealed at the end of the application process. By contrast, grade repetitions can be obtained after the revelation of m_0 even if *m* is below m_0 and even if *R* has not been put on the list of applications.

In this set up, students can get access to *H* if they rank it first and if $m>m_0$. They can get access to *L* if they rank it first, regardless of their mark. They can also get access to *L* if they rank it second after *H* and if $m< m_0$. In all other cases (i.e., if they apply directly for *R* or if they apply for *H* only and $m< m_0$), they can only repeat a grade.

A. Applications

Given these rules, students' applications depend on their preference over the different tracks as well as on whether they expect (at the time of applications) that their average mark will be above m_0.

With respect to preferences, we denote V_H, V_L and V_R the discounted values of the three different possible school choices. We do not put any constraint on their possible ranking except that we assume that if a student prefers *H* to *L* then he prefers direct access to *H* than access to *H* through grade repetition[21]. Formally it amounts assuming that $V_H>V_R$ whenever $V_H>V_L$.

With respect to expectations at the time of applications, we denote $\underline{m}_0$ the minimum value of the expected support of m_0. For each student, it indicates the minimum value of m_0 that is believed to be possible. Not all students believe that $\underline{m}_0<m$, that is not all students believe that they have a chance to get admitted in track *H*. We denote *I* a dummy variable indicating that $\underline{m}_0<m$. This variable indicates that a student believes that the probability that m_0 will fall below his own *m* is strictly positive. When *I*=1, students believe that they have some chance to get admitted into *H*, whereas when *I*=0, they believe to have no chance to get into *H*. In the remainder, we assume that when *I*=0, students do not put *H* in their list.

[21] Assuming that V_R can be written $\delta(PV_H+(1-P)V_L)-C$ where *C* is the cost of grade repetition, *P* the expected probability that average mark *m* will be above m_0 after a repetition, and δ discount factor, it is actually easy to check that if $V_H>V_L$ then $V_H>V_R$.

In this set-up, there are six types of optimal applications depending on whether students believe or not to have a chance to get access to *H* (two cases: *I*=0 or *I*=1) and on their preference over the 3 school options (with 3 possible rankings: $V_L>V_H>V_R$ or $V_H>V_L>V_R$ or $V_H>V_R>V_L$). Students whose ranking is $V_H>V_R>V_L$ will apply for *R* if their expectation is *I*=0 and for *H* if their expectation is *I*=1. Students whose ranking is $V_L>V_H>V_R$ will apply for *L* regardless of whether their expectation is *I*=0 or *I*=1. Finally, students whose ranking is $V_H>V_L>V_R$ will apply for *L* if *I*=0 and for *H* if *I*=1. In the latter case they will put explicitly *L* as a second choice. Table A1 summarizes how applications depend on preferences over tracks and expectations about outcomes of the assignment process (as captured by *I*).

Our dataset provides information on applications only (not on preferences or on expected m_0) and the key question is whether such data make possible to separately identify the effect of the treatment on preferences. With respect to this issue, the key feature of the model is that the subset of students applying either for *R* or for *H* only (i.e., who do not include *L* ins their lists of applications) corresponds exactly to the subset whose preferences satisfy $V_H>V_R>V_L$. This result is summarized by the following proposition:

Proposition A1 (applications) : *There is a one-to-one relationship between the group of students who do not include L in their list of applications and the group of students such that* $V_H>V_R>V_L$.

As discussed in the main text, the intervention induced a decline in the proportion of preselected students who do not include *L* in their list of applications of about 5 percent points. In our theoretical set-up and according to Proposition 1, it can be interpreted as meaning that 5% of preselected students changed their preferences over the treatment year, and namely started preferring *L* to *R* ($V_L>V_R$) rather than *R* to *L* ($V_R>V_L$).

B. Offers

After m_0's realization, students get an offer (denoted *O*) from the school system. This offer depends on their initial application, but also on whether realized m_0 falls above or below their own average mark *m*. Specifically, if they applied for track *H* only, the offer they get is *H* if $m>m_0$ and *R* if $m<m_0$. If they applied for (*H*,*L*) the offer is *H* if $m>m_0$ and *L* if $m<m_0$. If they applied for *L* or for *R*, the offer *O* corresponds to their application.

From this (and using results in Table A1), it is not difficult to take one step forward and show the relationship between students' preferences and expectations on the one hand

and, on the other hand, the offers that they get from the school system. As summarized in Table A2, there are four leading cases depending on students' preferred options and expectations. First, students who are offered to get into track *H* correspond to those with both high aspirations (*H* ranked first among all possible school options) and high/realistic expectations (*m* above both $\underline{m}_0$ and m_0). Second, students who are offered to repeat 9th grade even though they did not put *R* on their list of applications correspond to those with both relatively high aspirations (*R* ranked above *L*) and high/unrealistic expectations (*m* above $\underline{m}_0$ but below m_0). Third, students who are offered to repeat 9th grade *R* and who actually put *R* on their list of applications correspond to those with both high aspirations (*R* ranked above *L*) and low expectations (*m* below $\underline{m}_0$). Students who are offered to get into *L* correspond to the remaining cases.

C. Assignments

The last important feature of the model is that final assignments do not necessarily correspond to offers from the school system, since there are two options outside the school system, namely apprenticeship (denoted *A*, value V_A) and drop out (denoted *D*, value V_D). Denoting V_O the value of the offer *O* that students get from the school system, students accept this offer if and only if V_O is above both V_A and V_D. Denoting $V_{EXT}=\max(V_A,V_D)$ the value of leaving the school system, the assignment observed at t+1 is *O* if $V_O>V_{EXT}$, but it is *D* if $V_O<V_{EXT}$ and $V_D>V_A$ and it is *A* if $V_O<V_{EXT}$ and $V_A>V_D$.

From this (and building on results in Table A2), Table A3 provides a characterization of students who eventually stay in the school system in terms of initial preferences, expectations and marks. For example, it is not difficult to check that students eventually assigned to *H* are those who consider *H* as their best possible options (i.e., $V_H>\max(V_L,V_R,V_A,V_D)$), who initially expected to have a chance to be admitted into track *H* (i.e., $m>\underline{m}_0$) and whose average mark *m* turns out to be actually above realized m_0. Students eventually assigned to *R* but who did not put *R* on their list of applications corresponds to those with both high aspirations (*H* and *R* ranked first) and high/unrealistic expectations (*m* above $\underline{m}_0$ but below m_0). In contrast, students eventually assigned to *R* but who did put *R* in their lists corresponds to those with high aspirations but low expectations.

In this set-up, data on assignments within the school system provide direct information on the joint distribution of students across possible types of aspirations (as defined by the most valued option) and possible type of expectations.

Proposition A2 (assignment within the school system) : *There is a one-to-one relationship between (a) the group of students eventually assigned to H and the group with both high aspirations (H most preferred track) and high/realistic expectations (m above* $\underline{m}_0$ *and* m_0*); (b) the group of students who are assigned to R even though they did not put R in their list and the group with both high aspirations and high/unrealistic expectations (m above* $\underline{m}_0$ *but below* m_0*), (c) the group of students who applied to R and are eventually assigned to R and the group with both high aspiration and low expectations (m below* $\underline{m}_0$*).*

For each student, our data provide information on both applications and assignments. In our theoretical set up and according to proposition A2, this information makes possible to identify the effect of the treatment on the joint distribution of preferences and expectations. For example, the fact that the intervention did not affect probability of assignment to *H* can be interpreted as meaning that it did not induce students with high/realistic expectations to apply for less selective tracks. Similarly, the rise in the proportion of students assigned to *L* at the detriment of students assigned to *R* after having applied to *H* can be interpreted as meaning that the intervention induced students with high/unrealistic expectations to adopt more realistic aspirations (i.e., include *L* in their lists rather than focusing on *H* only).

D. Apprenticeship and dropout

With respect to assignment outside the school system (*A* or *D*), they correspond to situation where $V_O<V_{EXT}$. Two cases can be considered, depending on whether V_O coincides or not with the maximum value obtainable within the school system (i.e., $\max(V_H,V_L,V_R)$).

If $V_O=\max(V_H,V_L,V_R)$, then $V_O<V_{EXT}$ means that the best option outside the school system is preferred to any alternative within the school system. In such a case, a rise in *A* (or a decline in *D*) cannot be obtained by simple changes in expectations about the outcome of the assignment procedure. Specifically, a rise in *A* involves necessarily an increase in V_A compared to alternative options whereas a decline in *D* involves necessarily a decline in V_D compared to alternative options.

By contrast, if $V_O<\max(V_H,V_L,V_R)$, then $V_O<V_{EXT}$ does not exclude that there exists an option within the school system which value is higher than V_{EXT} . Specifically, for students with low expectations, we could very well have $V_H>V_{EXT}>\max(V_L,V_R)=V_O$

In such case, a rise in *A* could be induced by a change in V_A, but it could also be induced by students with $V_H>V_A>V_O$ moving from high to low expectations (i.e., from *I*=1 to *I*=0). Similarly, a decline in *D* could be induced by students with $V_H>V_D>V_O$ moving from

low to high expectations. In this scenario, however, the rise in *A* (or the decline in *D*) would be mostly driven by students who do not apply for *H*, which is not what we observe.

Assuming that the rise in *A* and the drop in *D* are mostly driven by change in preferences, the simplest scenario is that the intervention induced some potential school leavers (i.e., satisfying $\min(V_D,V_A)>\max(V_H,V_L,V_R)$) to move from $V_D>V_A$ to $V_A>V_D$. As discussed in the main text, this assumption has testable predictions, and namely an absence of any impact of the intervention on the size and composition of the group of school leavers.

Table A1: Applications as a function of preferences and expectations about $\underline{m}_0$

Preferences	Expectations about $\underline{m}_0$	
	I=0	I=1
$V_L > V_H > V_R$	*L*	*L*
$V_H > V_L > V_R$	*L*	*(H,L)*
$V_H > V_R > V_L$	*R*	*H*

Table A2: Relationships between preferences, expectations and offers from the school system

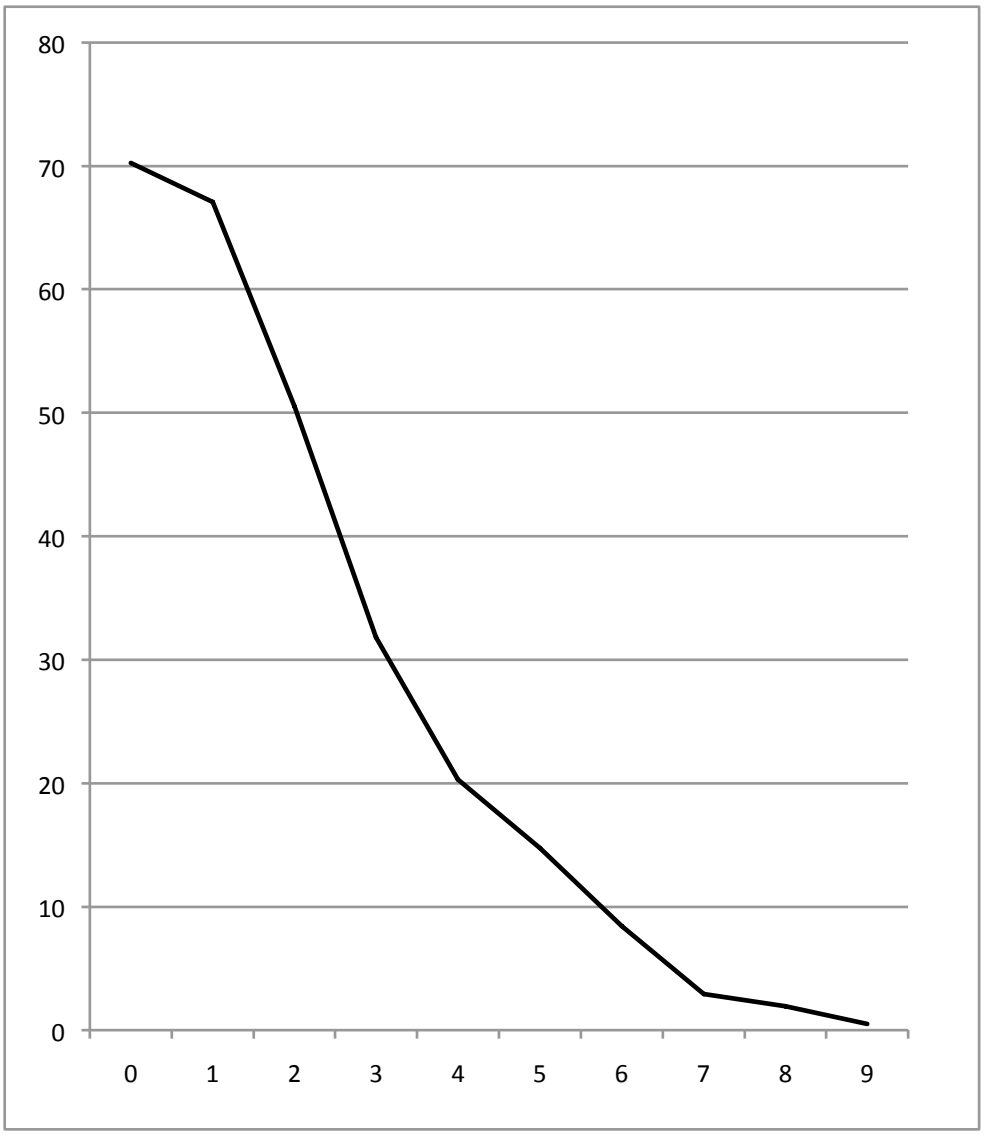
Preferences, marks and expectations	Applications	Offers
$V_H > \max(V_k, k=R,L)$ $m>\max(m_0,\underline{m}_0)$	*H* or (*H,L*)	*H*
$V_H > V_R > V_L$ $\underline{m}_0>m$	*R*	*R*
$V_H > V_R > V_L$ $m_0>m>\underline{m}_0$	*H*	*R*
$V_L > V_R$ $m<\max(\underline{m}_0, m_0)$	*L* or (*H,L*)	*L*

Table A3: Characterization of students eventually assigned within the school system.

Preferences, marks and expectations	Application	Offer	Assignment
$V_H > \max(V_k, k=R,L,A,D)$ $m > \max(m_0, \underline{m}_0)$	*H* or (*H,L*)	*H*	*H*
$V_H > V_R > \max(V_k, k=L,A,D)$ $\underline{m}_0 > m$	*R*	*R*	*R*
$V_H > V_R > \max(V_k, k=L,A,D)$ $m_0 > m > \underline{m}_0$	*H*	*R*	*R*
$V_L > \max(V_k, k=R,A,D)$ $m < \max(\underline{m}_0, m_0)$	*L* or (*H,L*)	*L*	*L*

OnlineAppendix B

Appendix Figure B1: Proportion of preselected students by deciles of academic performance



Note: The horizontal axis gives the deciles of academic performances (measured with average marks, term 1). The vertical axis is the proportion in each decile that were preselected by the principals.

Appendix Table B1: Baseline comparison of preselected students in test and control classes and differential response rates

	C	T-C	se	Obs
A: Variables available for the full sample				
(Preselected students)				
Girls	44.2	-0.8	2.5	1 130
Has already repeated a grade	54.4	-0.4	2.4	1 130
Low income	32.3	0.9	2.6	1 130
B: Differential response rates for different surveys and administrative data				
Survey parents (preselected students)				
General information meetings (parents)	74.6	-1.3	2.6	1 130
Indiv. meetings with career counselor (parents)	74.6	-0.9	2.5	1 130
Indiv. meeting with teachers (parents)	75.8	-2.4	2.4	1 130
Has attended meetings of parents' assn (parents)	75.0	-2.2	2.6	1 130
Has talked with other parents (parents)	74.6	-2.0	2.6	1 130
Happy with school info (parents)	75.0	-1.8	2.5	1 130
Expected diploma (parents)	74.4	-1.5	2.5	1 130
School data (preselected students)				
Annual average marks	97.1	-0.2	0.9	1 130
Average marks, term 3	94.4	-2.5	2.0	1 130
Warning work effort, term 3	95.0	-0.2	1.3	1 130
Suspension, term 3	95.0	-0.8	1.5	1 130
Truancy, term 3	92.7	-1.3	0.9	1 130
Survey on networks				
Tables 7 and 8 (full sample)	92.0	3.3*	2.0	3 161
Tables 7 and 8 (some friends preselected)	36.8	3.4	3.2	3 161
Tables 7 and 8 (some friends preselected and Marks<12)	16.3	1.3	1.7	3 161
Table B5	44.6	-0.2	5.1	1 130

Note: Each row corresponds to a different model, based on administrative data. Column (C) shows the average value for the non-selected students in the control group. Column (T-C) contains the coefficient of the treatment dummy. We use a linear probability model where we control for school dummies.

Appendix Table B2: Treatment effect on the share and composition of preselected students who remain in the national education system

	C	T-C	s.e.	Obs
Stays within the education system	85.4	+0.6	1.6	1 130
Composition of the group:				
Boys (%)	*55.4*	*+0.6*	*2.8*	971
Has just repeated 9th grade (%)	*6.8*	*-0.7*	*1.7*	971
Annual average marks (/200)	*87.5*	*+0.7*	*0.9*	971
Low income parents (%)	*32.9*	*+1.7*	*2.9*	971
First term marks (/20)	*9.32*	*-0.04*	*0.08*	971

Note: Each row corresponds to a different model, based on administrative data. Column (C) shows the average value for the non-selected students in the control group. Column (T-C) contains the coefficient of the treatment dummy. We use a linear probability model where we control for school dummies, and, if the variable is not the dependent one, for first trimester average marks, year-long average marks, gender, currently repeating dummy, low income dummy, as well as dummies for missing first trimester and year-long marks, and missing repeater. Column (se) shows corresponding standard error clustered at the class level. *: significant at the 10% level; **: at the 5% level.

Appendix Table B3: Effect on marks and assignment at the end of the last 9th grade, preselected students

	C	T - C	s.e.	Obs.
• Marks :				
Average marks (/240)	87.7	+0.6	0.9	1102
Prop. whose average marks>110	14.9	+0.9	2.0	1102
• Assignment				
3-year general	22.7	-0.9	2.0	1130
3-year vocational	55.0	-0.3	2.7	1130
2-year vocational	4.4	+3.3**	1.1	1130
Apprenticeship	6.7	+2.3	1.5	1130
Dropout	11.2	-4.8**	1.4	1130
Other	0.0	+0.4	0.3	1130

Note: Each row corresponds to a different model, based on administrative data. Column (C) shows the average value for the non-selected students in the control group. Column (T-C) contains the coefficient of the treatment dummy. We use a linear probability model where we control for school dummies, first trimester average marks, gender, currently repeating dummy, low income dummy, as well as dummies for missing first trimester and year-long marks, and missing repeater. For assignment variables we add two more controls: year-long average marks and a dummy for missing year-long average marks. Column (se) shows corresponding standard error clustered at the class level. *: significant at the 10% level; **: at the 5% level.

Appendix Table B4: Friendship networks: descriptive statistics

	Mean	Standard deviation
Respondents' own characteristics		
Boys (%)	49.1	50.0
Preselected (%)	26.4	44.1
Already repeated a grade (%)	31.7	46.5
Characteristics of respondents' friendship networks		
Number of friends	2.66	1.45
No friends	7.2	25.9
Proportion of friends with same sex as respondent		
For girls	89.6	23.0
For boys	88.9	24.9
Proportion of friends with same selection status as respondent		
For preselected students	35.9	33.9
For non-selected students	77.6	29.9
Proportion of friends with same grade repetition status as respondent		
For students who have repeated a grade	41.2	34.8
For students who have not repeated a grade	74.5	32.0

Note: The sample of respondents with information on friendship network has 4,040 observations. The average proportion of friends with same sex (or same selection status, or same grade repetition status) are computed on the sample of respondents with at least one friend (N=3,748). Reading: 49.1% of respondents with information on friendship networks are boys. For those who have at least one friend, the proportion of boys among their friends is 88.9%.

Appendix Table B5: Treatment effects on the size and composition of friendship networks, preselected students

	Mean C	T - C	s.e.
• Network stability			
Persistent friends	41.0	+10.4**	3.4
Persistent and preselected	*13.9*	*+6.2***	*2.8*
Persistent and non-selected	*27.1*	*+4.2*	*2.9*
Friends lost	66.1	-12.4**	5.5
New friends	47.0	-6.8*	3.7
• Network size			
Final network (June 2011)			
Number of friends	2.72	-.08	.15
0 friend	15.5	-6.0**	2.7
1-3 friends	47.8	+13.7**	4.3
4-5 friends	36.6	-7.7*	4.1
Initial network (October 2010)			
Number of friends	2.97	+.22	.20
0 friend	5.6	-2.8	2.2
1-3 friends	55.6	+0.9	6.2
4-5 friends	38.8	+1.9	6.6
• Network composition			
Final network (June 2011)			
Preselected friends	30.9	+1.4	3.5
Treated friends	1.6	+18.6**	2.5
Friends who will repeat or drop out	10.7	-5.4**	1.8
Friends who will enroll in 2-year program	3.3	+4.8**	1.9
Friends who will enroll in 3-year program	69.6	+7.8**	3.6
Has gained a future dropout/repeater	14.7	-7.1*	3.6
Has lost a future dropout/repeater	12.5	-4.1	2.8
Initial network (October 2010)			
Preselected friends	34.6	+4.0	3.8
Treated friends	1.8	+18.4**	2.1
Friends who will repeat or drop out	12.0	-4.2*	2.3
Friends who will enroll in 2-year program	3.6	+5.3**	1.9
Friends who will enroll in 3-year program	77.6	+2.2	3.1

Note: Each row corresponds to a different model, based on network survey data. The analysis is restricted to classes for which friendship surveys have been completed at both baseline and endline. Column (C) shows the average value for the non-selected students in the control group. Column (T-C) contains the coefficient of a treatment class dummy. We use a linear probability model where we control for school dummies, first trimester average marks, gender, currently repeating dummy, low income dummy, and dummies for missing first trimester and missing repeater. For network stability variables, the regression also includes number of friends in first trimester. For final network size variables, the regression also includes a control for first trimester value. Column (se) shows corresponding standard error clustered at the class level. The number of observations used is N=524. *: significant at the 10% level; **: at the 5% level.

Online Appendix C

Appendix Table C: Unemployment and wages at entry into the labor market, by education groups.

	Unemployment	Monthly earnings	
3-year general	10.6	1 650	172
3-year vocational	18.0	1 260	131
2-year vocational	27.0	1 100	115
Apprenticeship	16.6	1 300	136
2-year vocational	*19.8*	*1 250*	*130*
3-year vocational	*10.2*	*1 390*	*144*
Dropout	43.5	1 070	111
Late dropout	*37.2*	*1 160*	*122*
Early dropout	*50.6*	*960*	*100*

Source: French Labor Force Survey conducted in 2011. Note: The Table reports the average unemployment rate and average monthly earnings of wage-earner with 1-5 years of labor market experience.