Ready for boarding?
The effects of a boarding school for disadvantaged students.*

Luc Behaghel[†] Clément de Chaisemartin[‡] Marc Gurgand[§]

February 8, 2015

Abstract

Boarding schools substitute school to home, but little is known on the effects this substitution produces on students. We present results of an experiment in which seats in a boarding school for disadvantaged students were randomly allocated. Boarders enjoy better studying conditions than control students. However, they start outperforming control students in mathematics only two years after admission, and this effect mostly comes from strong students. After one year, levels of well-being are lower among boarders, but in their second year, students adjust: well-being catches-up. This suggests that substituting school to home is disruptive: only strong students benefit from the boarding school, once they have managed to adapt to their new environment.

Keywords: boarding school, cognitive skills, non-cognitive skills, randomized controlled trial, heterogeneous effects

JEL Codes: I21, I28, J24, H52

*This research was supported by a grant from the French Experimental Fund for Youth. We are very grateful to Jean-Michel Blanquer for his strong support to this project, as well as to Cédric Afsa, Pierrette Briant, Jean-François Bourdon, Bernard Lociciro, Stéphane Lours, Sithi Maricar, Cédric Montésinos, Bénédicte Robert and all National Education personnel who supported operational work and data collection; and to Pascal Bessonneau, Jean-François Chesné, Sylvie Fumel and Thierry Rocher, from the Evaluation Department of the French Ministry of Education, who created the cognitive tests used in this paper. Adrien Bouguen and Axelle Charpentier provided outstanding contributions as research managers at the J-Pal European office: we thank them very warmly, as well as J-Pal administrative team and research assistants. We also thank Karen Brandon, Xavier D'Haultfœuille, Julien Grenet, Francis Kramarz, Victor Lavy, Andrew Oswald, Roland Rathelot, Vincent Pons, Claudia Senik, Fabian Waldinger, Chris Woodruff, seminar participants at Warwick University, the Institute of Fiscal Studies, Uppsala University, Crest and PSE for their helpful comments.

[†]Paris School of Economics - INRA, luc.behaghel@ens.fr

[‡]University of Warwick, clement.de-chaisemartin@warwick.ac.uk

[§]Paris School of Economics - CNRS, gurgand@pse.ens.fr

1 Introduction

Boarding schools are an intensive form of education, in which students live at school, and visit their families only for weekends and vacations. There is a long-standing tradition in American and English upper-class families of sending male children to elite boarding schools even at a very young age. Cookson et al. (2008) argue that by doing so, parents hope to provide their children a sense of discipline, and, thus, prepare them for leadership positions. But boarding schools have also been used to increase the educational opportunities of marginalized and disadvantaged students. In the end of the 19th century, American philanthropists from the Indian Rights Association set up boarding schools for American Indians' children, most often located outside their parents' reservations. These philanthropists were hoping to assimilate these children into mainstream American culture, something they thought would be impossible to achieve through regular schooling: "Placing these wild children under a teacher's care for five hours a day, and permitting them to spend the other nineteen in the (...) degradation of the village, makes the attempt to educate (...) them a mere farce" (Report of the US Bureau of Indian Affairs, 1878). In 1926, 83 percent of the American Indian school-age population was enrolled in one of these boarding schools (see Adams, 1995). More recently, boarding schools have received renewed interest from policymakers seeking ways to enhance the academic progress of disadvantaged students. Two examples are the SEED boarding schools in the United States which serve poor black students, and the "boarding schools of excellence" in France which serve relatively high-ability students from poor families. In both cases, policy makers opened these schools because they were concerned that the poor studying conditions and negative influences students are exposed to in their home environment could impair their academic potential.

The explicit goal of these boarding schools is to operate a substitution between the two main inputs of the education production function, namely school and home environment, under the presumption that this will generate better outcomes for students. However, very little is known on the effects this substitution actually produces. Curto & Fryer (2014) is the only paper we are aware of which studies this question. The authors find that being enrolled in the SEED boarding school in Washington DC increases students test scores by 20 percent of a standard deviation per year spent in the school.

In this paper, we analyze the effects of a French "boarding school of excellence" on students cognitive and non-cognitive outcomes. The school we study was created in 2009, and is located in a rural area south of Paris. It was oversubscribed, and students offered a seat were randomly selected out of the pool of applicants. We followed the treatment and the control groups over two years after the lottery, and gave them cognitive and non-cognitive tests in the end of each academic year.

The boarding school dramatically increases the quality and the quantity of schooling inputs: boarders benefit from smaller classes, spend longer hours in study room, report much lower levels of classroom disruption, and praise the engagement of their teachers. These investments have positive returns: after two years, the treatment group performs substantially better on the math test. The difference is sizeable, and corresponds to a 20 percent standard deviation increase per year spent in the school. However, these positive effects hide two surprising findings. First, returns only kick in after two years: one year after the lottery, test scores are very similar in the treatment and control groups. This is in sharp contrast with papers studying the dynamic effects of educational interventions, which have often found stronger effects for the first year of treatment (see Krueger (1999)), or effects that are linear in the amount of exposure (see Abdulkadiroglu et al. (2011)). Second, returns are very heterogenous: we find that the average effect of the school after two years mostly comes from students in the higher tercile of math scores at baseline. For them, the effect is very large, around 50 percent of a standard deviation per year spent in the school.

We take advantage of the very detailed data we collected to investigate the mechanisms that could underlie these surprising patterns. When students arrive at the boarding school, they need to adapt to their new environment. First, they have to cope with the separation from friends and family. Second, they relinquish a certain amount of freedom. For instance, they have to wear a formal school uniform similar to those of English and American private schools, instead of their usual clothes. They also report spending four times less time watching television than control students, a difference probably due to the strong control exerted by the boarding school staff. Third, boarders face higher academic demands. They are immersed into an environment with peers who are academically stronger, and teachers who are more demanding: most students experience a sharp decline in their grades when they enter the school. These three factors are probably responsible for the lower levels of well-being we observe among boarders in the end of their first year. During their second year, students seem to adjust, and the positive effects of the intervention start kicking in. Boarders' levels of well-being catch-up with those of control students; their motivation becomes higher, and they also report spending more time on their homework, while there were no differences in the end of the first year on these two dimensions. The stark difference between returns to students' first year and second year in the boarding school might therefore arise from the following mechanism. Adjusting to the school reduces students well-being; this might in turn impede their learning, until they have adapted to their new environment. This could also explain why stronger students make more progress than weaker ones. We find some indication that the initial negative shock on well-being and motivation is larger for weaker students, while the recovery is faster for stronger students, although we lack statistical power to make definitive conclusions. Though this interpretation is somewhat speculative, we review other potential mechanisms, and we argue that they cannot fully account for all of our findings. For instance, recent research has shown that higher within-class ordinal position has a positive effect on academic performance (see Murphy & Weinhardt (2013)). This can explain why weaker students do not improve in the boarding school, as they lose many ranks when they join. However, this fails to explain why strong students do not improve during their first year.

Overall, our results suggest that boarding is a disruptive form of schooling for students. Once they have managed to adjust to their new environment, strong students make very substantial academic progress. On the other hand, this type of school does not seem well-suited to weaker students: even after two years we do not observe any test scores gains among them.

Beyond boarding schools, our results might shed new light on recent, puzzling results on elite schools. Many elite schools around the world use entrance exams to admit students. A number of papers have used regression discontinuity designs to measure the effects of these schools on students at the admission cut-off. These papers have consistently failed to find any effects on students' test scores (see Abdulkadiroğlu et al., 2014 and Lucas & Mbiti, 2012) or college enrollment (see Dobbie & Fryer Jr, 2013), and have even sometimes found negative effects on dropout rates among the most vulnerable students (see de Janvry et al., 2012). This has been interpreted as evidence that peer effects do not play a large role in the production of education (see Abdulkadiroğlu et al., 2014). Our analysis might substantiate another interpretation of these findings. When they enter these elite schools, students may benefit from the presence of strong peers, and at the same time, they may also be hampered by the need to adapt to a new, more competitive environment - as happens to students in our boarding school. The absence of any effect for students at the threshold could then be the sum of a positive peer effect and a negative adaptation effect. Moreover, overcoming this adaptation process might be easier for stronger students, so effects for them might be larger than for students at the admission cut-off. If this is the case, regression discontinuity estimates could differ from the average effect of these schools.

The remainder of the paper is organized as follows. In Section 2, we describe our research design, the complex data collection we had to complete for this project, and our study population. In Section 3, we present the main differences between the boarding school and the schools in which control students are enrolled. In Section 4, we present the effects the boarding school produces on students test scores. In Section 5, we discuss potential mechanisms underlying these effects. Section 6 concludes.

2 Research design, data, and study population

In the fall of 2005, important riots took place in the suburbs of Paris and other large French cities. These events triggered a number of political responses, including which the “Internats d'excellence” program. “Internats d'excellence” could be translated as “boarding schools targeting excellence”. These schools are dedicated to motivated and relatively high ability students in poor suburbs of large French cities. Policy makers were concerned that in those suburbs, poor school quality, negative influences from peers, and bad studying conditions at home could impair the academic success of motivated students. The school we study is located in a rural area southeast of Paris. It was the first “Internat d'excellence” to open, and it is also the largest of the 45 “Internats d'Excellence” now operating in France, with an intake accounting for 10% of that of the 45-school program. It serves students from all eastern parisian suburbs, the most deprived ones.

2.1 Research design and statistical methods

Students offered a seat in the boarding school were randomly selected out of a pool of applicants. We study the boarding school's first two cohorts, those admitted in September 2009 and September 2010. In 2009, 129 seats were offered to students in 8th to 10th grades. In 2010, 150 seats were offered to students in 6th to 12th grades. The school received 275 applications in 2009, and 499 in 2010. In the spring of each year, a committee screened applications to make sure that the students met the school's eligibility criteria. The policy was intended to target motivated students living in homes that were considered unconducive to scholastic progress. In 2009, 73 applications were discarded for lack of eligibility. In 2010, 216 were discarded. A few applicants (five in 2009 and seven in 2010) were granted priority admission because they faced particularly tough conditions at home. The boarding school had set a predetermined intake of students at the grade and gender levels, to ensure that male- and female-only dormitories of given sizes could be formed. In each grade $\times$ gender stratum in which the number of applicants still exceeded the number of seats remaining after the screening and priority admission, we randomly allocated applicants a waiting list number. Seats were offered following this order. Our study population is made up of the 395 students who participated in a lottery. Our treatment group consists of the 258 students who received an offer, and our control group consists of the 137 students who did not receive an offer.

The lottery created very similar treatment and control groups. In Table 13 in the Appendix, we compare them on 14 measures of baseline ability and socio-economic background.

We find only one significant difference at the 10% level.

Compliance with random assignment was high. As shown in Table 1, one year after randomization lottery losers had spent only 0.05 years in the boarding school. 6% of these students managed to enrol, because one of their siblings had been admitted to the school, but not all of them stayed for the entire year. On the contrary, by that time lottery winners had spent 0.8 years in the school. 14% of them never joined, and 10% left during the year. Two years after the lottery, winners had spent 1.27 more years in the school than losers. This difference is lower than 2× the difference after one year because the exit rate between the two years was higher among winners.

Table 1: Effect of the lottery on years spent in the school

	Control mean	T-C after 1 year	SE	T-C after 2 years	SE	N
Years of treatment	0.052	0.748***	0.040	1.269***	0.082	790

Notes. This table reports coefficients from a regression of the number of years spent in the school on a dummy for year 1 (column 2), the interaction of this dummy with our lottery offer (column 3), a dummy for year 2, the interaction of this dummy with our lottery offer (column 5), and the statistical controls listed in Section 2.2. We use propensity score reweighting to control for lottery strata. We use two observations per student (one and two years after the lottery). Standard errors reported in columns 4 and 6 are clustered at the student's level. *significant at 10%; **significant at 5%; ***significant at 1%.

In all the regressions we estimate in the paper, we use propensity score reweighting to account for the fact our lottery offer is randomly assigned within grade × gender strata (see Rosenbaum & Rubin, 1983 and Frölich, 2007). Let Z_i be a dummy denoting our lottery offer, and let S_i denote lottery stratum. In our regressions, students in the treatment group receive a weight equal to $\sqrt{\frac{P(Z_i=1)}{P(Z_i=1|S_i)}}$, while control students receive a weight equal to $\sqrt{\frac{P(Z_i=0)}{P(Z_i=0|S_i)}}$.[1] These weights ensure that our coefficients of interest arise from the comparison of lottery winners and losers within and not across strata. Alternatively, we could have estimated unweighted regressions with lottery strata indicators. These regressions estimate a variance-weighted average of within-strata comparisons, which does not give to each stratum its natural weight in the population. Therefore, these regressions do not estimate standard parameters of interest in policy analysis such as intention to treat (ITT) or local average treatment effects (LATE). Notwithstanding, it is worth noting that using one or the other estimation method hardly changes our main results.

[1] Using a GMM representation, it is easy to see that this reweighting is computationally equivalent to standard propensity score reweighting.

2.2 Data

French students do not take standardized tests every year. Consequently, we had to conduct a complex data collection operation to measure students' academic ability and non-cognitive outcomes. This, among other things, involved collaborating with 169 different schools scattered over the whole of France as we detail below.

One and two years after the lottery, we gave students two standardized tests, each one hour and 30 minutes in length. The first test included a one-hour French test and a 30-minute non-cognitive questionnaire. The second test included a one-hour mathematics test and another 30-minute non-cognitive questionnaire. The French Department of Education created the French and mathematics tests. We devised the non-cognitive questionnaires, using validated psychometric scales and questions from the Program for International Student Assessment (PISA).

Tests were taken online in the computer lab of students' schools. Boarders took them with their classmates. To ensure that treatment and control students were taking the test in somewhat comparable conditions, we randomly selected three classmates to take the test with every student not enrolled in the boarding school. We also took extensive steps to prevent cheating: we sent research assistants to the boarding school to serve as test proctors; the programming of the test ensured questions did not appear in the same order on neighboring computers, so that neighboring students would not answer the same question at the same time; students could only bring a pen and a sheet of paper to the test room. Students not enrolled in the boarding school were scattered among 169 schools. Most of them were in the local school district of Creteil, but some of them were in other areas of France. Due to budget constraints, we could not send research assistants to monitor the tests in each of these 169 schools. Notwithstanding, the Department of Education wrote to the principals of all of these schools to require that our test be monitored by someone from the school. Because the tests were taken online, we can check whether students who took the test out of the boarding school spent more time on the test than was allowed. We do not find evidence of this (see Table 14 in the Appendix). A few schools did not have a working computer lab, and we had to send them paper versions. A few students had dropped out of school by the time they were supposed to take one of our tests. These students took the tests at home. Our main results are robust to dropping these observations (see Table 15 in the Appendix).

In order to ensure that our results would not be plagued by differential attrition, extensive effort was required to reach all of the control students, who were scattered among many more schools than treatment students. In the end, more than 90% of students took our tests, and attrition was balanced in the treatment and in the control groups as shown in Table 16 in the

Appendix.

Cognitive tests were partly revised each year by the Department of Education to ensure that students and their teachers could not anticipate which questions would be asked in the following year. We tried not to change our non-cognitive questionnaires from one year to the other, to ensure the comparability of students' responses. However, at the end of the first year of data collection, we realized that students took much less than the allotted 30 minutes to answer our non-cognitive questionnaires. As a result, in the following years, we added more questions. Unfortunately, this means that some questions are not available one year after the lottery for the first cohort of students.

Finally, we also rely on a number of pre-existing sources of information to describe our study population and the treatment. We use students' average marks in mathematics and French from transcripts required in the application process as measures of baseline ability. We use the "Base Scolarité" (Sconet) administrative data set to describe the students' socio-economic background. We also use data from the "Diplôme National du Brevet", the French national exam given to students at the end of middle school, to compare applicants to the boarding school to their classmates and to French students. Finally, we use the "Base Relais", an administrative data set on teachers and supervisors working in French schools, to compare the school staff in the boarding school to the staffs in schools where control students were enrolled.

To increase statistical precision, all of our regressions include the following list of controls: students grades in French, math, and school behavior, as per the transcripts they provided in their application; a dummy for students enrolled in a Greek or Latin optional class at baseline; the level of financial aid students' family receive under the means-tested grant for middle- and high-school students; a dummy for whether French is the only language spoken at home; a dummy for students whose parents are unemployed, blue collar workers, or clerks; dummies for boys, second cohort, and school grade. Our main results are robust to dropping these controls from the regressions (see Table 17 in the Appendix).

2.3 The population of applicants to the boarding school

We measure the effect of the boarding school within the population of students who applied for seats. This population is the product of several layers of selection. In the fall of each year, the Department of Education wrote to school principals asking them to identify motivated students who lacked home environments conducive to studying, and to encourage these students to apply. Students interested in joining the school then had to fill out an application form, write a letter of application, and provide a letter from a parent. Finally, a committee discarded applications which did not match the profile targeted by the policy.

In Table 2, we describe our study population. Whenever data are available, we also compare the student population to several reference populations. Our population comprises a majority of girls (57 percent), and students' average age when they applied was 14. Eligible applicants are higher achievers than their classmates, but median students in the French population. At the time of application, applicants' average marks in French and mathematics were, respectively, two and two and a half points above those of their classmates, and they ranked slightly above the third decile of their class in each discipline. Slightly more than half of our study population had taken the end-of-middle-school French exam before applying for the boarding school. Those students scored 14 percent of a standard deviation higher than the French average in French and mathematics, and 42 percent of a standard deviation higher than their classmates. Under a normality assumption, this implies that eligible applicants stand at the 45th percentile of the French distribution.

Eligible applicants are also underprivileged students. The share of eligible applicants who are recipients of the means-tested grant for middle- and high-school students is twice as large as in the French population, and close to the share observed among students enrolled in "Éducation prioritaire" schools, a program that encompasses French schools located in the poorest neigborhoods. Still, given that the program explicitly targets disadvantaged students, it might seem surprising that this fraction is not higher than 46 percent. This could be due to the fact that a substantial fraction of eligible families do not claim this grant because its amount is low and the application procedure costly. Applicants' parents are as likely to be clerks and blue-collar workers as parents of their classmates, and more likely to be inactive, and the schools from which applicants come are located in one of the poorest areas in France. French is the only language spoken at home for only 40 percent of them: this suggests that many come from families that recently immigrated to France.

3 The treatment

In this section, we compare the amount of educational inputs received by boarders and control students. Specifically, we estimate the following two stage least squares (2SLS) regressions for 40 such inputs Y_i:

$$Y_i = \eta_0 + \eta_1 D_i + X_i' \zeta + \varepsilon_i. \tag{1}$$

Y_i are either objective measures of the resources of the school where student i is enrolled (e.g. class size), or measures of students' i experience (e.g. perceived levels of classroom

Table 2: Economic background and baseline academic ability of applicants

	Applicants	French students	"Éducation prioritaire"	Classmates
Baseline ability				
Mark in French, transcripts	12.566			10.500
Rank in French, transcripts	0.271			
Mark in Mathematics, transcripts	13.001			10.529
Rank in Mathematics, transcripts	0.298			
Middle school exam, French	0.141	0.000	-0.288	-0.335
Middle school exam, Mathematics	0.145	0.000	-0.352	-0.241
Socio-economic background				
Means tested grant, middle school	0.455	0.278	0.468	
Means tested grant, high school	0.417	0.249		
Parent clerk	0.249			0.210
Parent blue collar	0.261			0.278
Parent inactive	0.179			0.082
Parent has completed high school	0.238			
Only French spoken at home	0.405			
Other characteristics of applicants				
Share of girls	0.572			
Average age	14.125			
Number of children in the family	2.832			

Notes. This table compares applicants to the boarding school to a number of reference populations. Socio-economic variables on applicants come from the "Sconet" administrative data set. Transcripts come from their application files. Grades in the end-of-middle-school exam come from the "Base Brevet" administrative data set. Data on French students, students enrolled in "Éducation Prioritaire" schools and in the Créteil school district come from DGESCO (2010). Ranks range from 0 (highest) to 1 (lowest).

disruption). D_i is a dummy for whether student i was enrolled in the boarding school at the time the measure was made. We use the dummy for our lottery offer Z_i as an instrument for D_i. X_i is the vector of statistical controls listed in Section 2.2 and ε_i is a residual. η_1 measures the difference in the amount of input Y_i received by students who comply with their lottery offer when they are in and out of the boarding school. Indeed, it is equal to the difference between lottery winners' and losers' average of Y_i, normalized by the difference in the share of students enrolled in the boarding school between these two groups. Estimates of the mean of Y_i for compliers in the control group are displayed in the second column of Tables 3, 4, and 5 (we follow the method described in Abadie (2003) to estimate this quantity). Estimates of η_1 are displayed in the third column.

To measure students' experiences, we included questions from PISA on levels of disruption in the classroom, relationships between students, etc., in the questionnaires we administered to students. Answers to these questions could take four values: "strongly disagree", "disagree", "agree", and "strongly agree". In Tables 4 and 5, we present the effect of being enrolled in the boarding school on students' standardized answers to these questions. When several questions arguably measure the same dimension, we aggregate them into a score which we also standardize.[2]

The boarding school benefits from more resources than the schools in which control students are enrolled. As shown in Table 3, the teacher-to-student ratio is 35 percent higher in the boarding school, which corresponds to the fact that classes are 25 percent smaller. The supervisor-to-student ratio is almost five times larger, because students must also be monitored at night. Boarding school teachers are better educated and less experienced than teachers of control students. A larger fraction of them hold the "Aggrégation", the highest degree for high school teachers in France. But twice as many of them have less than three years of experience. Based on these two observable dimensions, boarding school teachers appear less likely to generate high test scores than those in control schools. There is indeed little evidence in the literature that more educated teachers generate higher students test scores, while there is some evidence that experienced teachers do. In particular, the first years of experience seem to have higher returns – for a meta-analysis, see Hanushek & Rivkin (2006). But teachers in the boarding school have volunteered to join, so they could differ from control schools teachers on unobservable dimensions such as motivation.[3]

[2]All the tables in this section present results two years after the lottery took place, because some of these questions were not included in the questionnaires administered to the first cohort one year after the lottery. In Tables 19, 20, and 21 shown in the Appendix, we present results one and two years after the lottery, keeping only the second cohort for questions which were not administered to the first cohort one year after the lottery. We find few differences between the two years.

[3]In our policy report (see Behaghel et al., 2013), we conducted a cost-benefit analysis. Using results from Piketty & Valdenaire (2006), we computed that the effects of the boarding school on students cognitive scores

Table 3: Resources allocated to the boarding school

	$E(Y_0\|C)$	LATE	SE	N
Class size	26.060	-6.089***	0.966	352
Teachers per 100 students	8.416	2.974***	0.242	373
Supervisors per 100 students	1.592	6.088***	0.129	375
Teachers with “Aggregation” degree	0.177	0.101***	0.021	378
Teachers with less than 3 years experience	0.187	0.202***	0.012	378
Teachers years of experience	9.891	-3.493***	0.429	378

Notes. This table reports results from 2SLS regressions of the outcomes in the first column on a dummy for being enrolled in the school and the statistical controls listed in Section 2.2, using our lottery offer as an instrument. The third column reports the coefficient of the dummy (η_1 in equation 1). Standard errors in column 4 are clustered at the class level. The second column reports an estimate of the mean of the outcome for compliers not enrolled in the school. We use propensity score reweighting to control for lottery strata. The last column displays the number of observations. We use only one observation per student, two years after the lottery. The class size variable comes from students' questionnaires. The other variables come from the “Base Relais” administrative data set. *significant at 10%; **significant at 5%; ***significant at 1%.

Boarders also benefit from a much better classroom experience than control students, as shown in Table 4. As per our score, levels of classroom disruption are 70 percent of a standard deviation lower in the boarding school. For instance, students are less likely to answer that they cannot work well in the boarding school. Living together in the boarding school increases solidarity and cooperation among students: treated students are more likely to report that they do their homework in groups, and that strong students help weak ones. Boarding school teachers are more engaged: boarders are more likely to report that their teachers keep explaining until all students have understood, that they give them the opportunity to express their opinions, and that they care about students' academic progress. They also perceive their teachers much more positively: overall, our students-teacher relationship score is one standard deviation higher in the boarding school.

But boarders face higher academic demands. They have to take a two-hour test each week, and grading in the boarding school is much harsher than in a regular school. Students from the first cohort experienced a 2.1 point decrease in their marks in math after entering the boarding school.[4] This is a substantial drop, equivalent to 53 percent of the standard

are comparable to those one could expect to obtain by dividing class size by two within a comparable population of French middle- and high-school students. We also found that the expenditure per student is twice as large in the boarding school as in a regular school, mostly due to the boarding school component of the program (monitoring students at night, heating...). We therefore concluded that the boarding school is as cost-effective as class size reduction.

[4]Unfortunately, we do not have marks in the boarding school for the second cohort of students.

Table 4: Students' experience in the classroom

	$E(Y_0\|C)$	LATE	SE	N
Attendance over the last two weeks				
Attendance score	0.215	0.166	0.204	361
Missed school	-0.319	-0.065	0.214	362
Skipped classes	-0.182	-0.156	0.204	361
Arrived late	-0.068	-0.209	0.209	362
Disruption				
Disruption score	-0.145	-0.697***	0.231	360
Teacher often waits students calm down	-0.164	-0.401*	0.226	361
Students start working long after class begins	-0.193	-0.336	0.222	361
Students cannot work well	-0.104	-0.426**	0.209	360
There is noise and disruption in the classroom	-0.128	-0.524**	0.217	361
Students do not listen to the teacher	-0.044	-0.971***	0.247	361
Relationships between students				
Students relationships score	0.101	0.701***	0.200	288
Students are ashamed when they have good grades	-0.048	-0.218	0.213	289
Weak students make fun of strong ones	-0.388	0.093	0.209	334
Students do their homework in group	-0.123	0.532**	0.215	361
Strong students help weak ones	-0.050	0.942***	0.217	360
Teachers' engagement				
Teachers' engagement score	-0.139	1.350***	0.256	361
She cares for students academic progression	-0.064	0.748***	0.202	361
She explains until students understand	-0.148	1.205***	0.215	361
She listens to students opinions	-0.023	0.800***	0.222	361
Teacher-students relationships				
Teacher-students relationships score	0.032	1.000***	0.248	346
Students get along well with their teachers	0.057	0.834***	0.269	362
Teachers care for students	0.073	0.760***	0.224	346
Teachers listen to students	0.031	0.722***	0.230	362
Teachers give supplementary help if needed	-0.026	0.869***	0.233	362
Teachers are fair to students	0.007	0.715***	0.237	362

Notes. This table reports results from 2SLS regressions of the outcomes in the first column on a dummy for being enrolled in the school and the statistical controls listed in Section 2.2, using our lottery offer as an instrument. The third column reports the coefficient of the dummy (η_1 in equation 1). Standard errors in column 4 are clustered at the class level. The second column reports an estimate of the mean of the outcome for compliers not enrolled in the school. We use propensity score reweighting to control for lottery strata. The last column displays the number of observations. We use only one observation per student, two years after the lottery. All the variables come from students' questionnaires. Each score in italics is standardized and computed from the individual items listed just below. *significant at 10%; **significant at 5%; ***significant at 1%.

deviation of math grades in the boarding school. Because school marks in France are not digitized, we could not collect them for control students. Teachers in regular schools might have tougher marking standards for higher grades, in which case control students might also have experienced a decline of their marks following the lottery. To investigate this possibility, we conduct the following exercise. As students from the first cohort entered in 8th, 9th, or 10th grade, they thus went from 7th to 8th, 8th to 9th, or 9th to 10th grade. Transcripts in France usually include both a student's mark and the average mark in her class. The green line on Figure 1 shows class averages in math at baseline for students who applied when they were in 7th, 8th, 9th, or 10th grade. Under the assumption that these four groups of students do not come from schools with very different marking standards, this green line should be a good proxy of the "natural" year-on-year evolution of marks between these four grades. The three blue lines on Figure 1 show the evolution of marks after entering the school for boarders who joined in 8th, 9th and 10th grade, respectively. The green line is mostly flat: the only noticeable pattern is that class averages decrease by 1.2 points between 7th and 8th grade. On the contrary, the three blue lines all sharply decrease. Given that students who applied in 7th grade only account for 20 percent of the first cohort, only $1.2 \times 0.2/2.1 = 11$ percent of the sharp decline in marks this cohort experienced can be attributed to the mechanical evolution of school marks across grades. The remainder seems attributable to harsher grading standards in the boarding school.

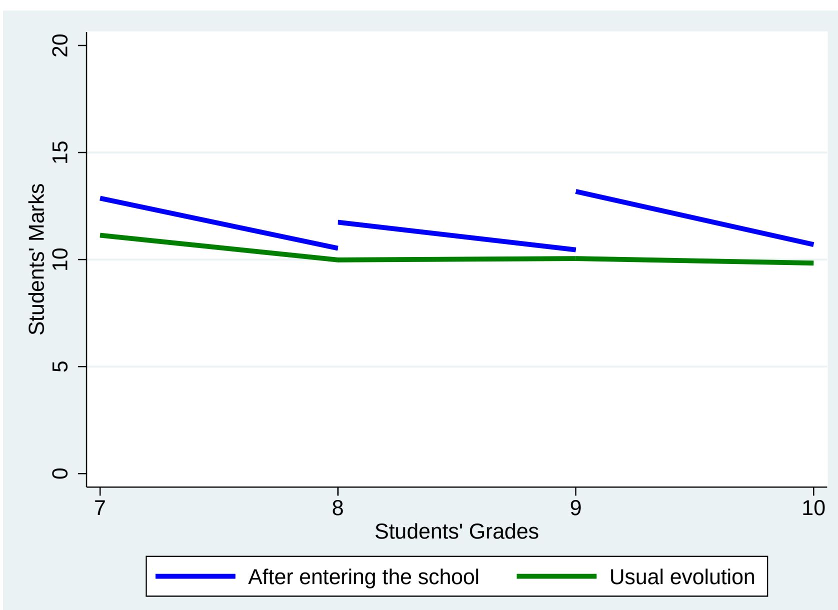

Figure 1: Evolution of Students' Mathematics Marks

Boarders also have to cope with longer studying days and stricter disciplinary rules. Students do not have more class hours in the boarding school than in a regular school, but at the end of their school day they have to spend one hour and a half in a study room in which they are monitored by a supervisor to do their homework. In control schools, spending some time after the school day in a study room is only a non-mandatory option available to students. This is why treated students report spending almost six hours per week in a study room, against one hour and fifteen minutes for those in the control group, as shown in Table 5. Access to TV is strictly regulated in the boarding school, and playing video games is, in theory at least, forbidden. Consequently, treated students report watching TV only 24 minutes per day, against 1 hour and 36 minutes for controls. They also report spending less time playing video games, but the difference is not statistically significant. From the end of the school day to the moment they go to bed, boarders are monitored by supervisors, who have to enforce stringent disciplinary rules. For instance, students have to wear formal school uniforms, a very unusual practice in French schools. This seems to generate conflicts between them and students: our students-supervisor relationship score is 37 percent of a standard deviation lower in the boarding school than in control schools.

Overall, the boarding school offers to underprivileged students an elite education reminiscent of French "Classes Préparatoires" and English and American upper-class boarding schools. Indeed, the important concentration of resources on a small number of students, the interactions with qualified and engaged teachers, the high academic demands, the long school days, and the strict disciplinary rules are common features of all these schools.

4 Effects of the boarding school on students cognitive outcomes

4.1 Effects on the average of test scores

This section presents the impacts of the boarding school on test scores in French and mathematics, one year and two years after the lottery. We present first-stage, intention-to-treat and two-stage least squares estimates in Table 6.

Panel A in Table 6 displays the first-stage estimates, i.e. estimates of the effect of winning the lottery on the number of years spent in the boarding school. Specifically, we estimate the following equation:

$$S_{it} = \gamma_0 1\{t=1\} + \gamma_1 Z_i \times 1\{t=1\} + \gamma_2 1\{t=2\} + \gamma_3 Z_i \times 1\{t=2\} + X_i'\zeta + \varepsilon_{it}. \quad (2)$$

Table 5: Students' experience outside the classroom

	$E(Y_0\|C)$	LATE	SE	N
Students' schedule after the school day				
Hours spent last week in study room	1.254	4.413***	0.862	353
Hours spent last Monday playing video games	0.430	-0.245	0.202	348
Hours spent last Monday watching TV	1.598	-1.180***	0.265	354
Supervisor-students relationships				
Supervisor-students relationships score	-0.058	-0.368*	0.219	289
Students get along well with their supervisors	0.016	-0.508**	0.218	319
Supervisors care for students	-0.149	0.117	0.210	362
Supervisors listen to students	-0.229	-0.032	0.215	332
Supervisors give supplementary help if needed	-0.152	-0.275	0.227	361
Supervisors are fair to students	0.090	-0.648***	0.216	306

Notes. This table reports results from 2SLS regressions of the outcomes in the first column on a dummy for being enrolled in the school and the statistical controls listed in Section 2.2, using our lottery offer as an instrument. The third column reports the coefficient of the dummy (η_1 in equation 1). Standard errors in column 4 are clustered at the class level. The second column reports an estimate of the mean of the outcome for compliers not enrolled in the school. We use propensity score reweighting to control for lottery strata. The last column displays the number of observations. We use only one observation per student, two years after the lottery. All the variables come from students' questionnaires. The supervisor-students relationships score is standardized; it is computed from the individual variables listed below. *significant at 10%; **significant at 5%; ***significant at 1%.

S_{i1} and S_{i2} respectively denote the *total* number of years that student i has spent in the boarding school by the end of the first and second academic years after randomization, respectively; $1\{t = 1\}$ and $1\{t = 2\}$ are dummies for first and second year; X_i is the vector of statistical controls listed in Section 2.2; Z_i indicates whether student i won the lottery; and ε_{it} is a residual. Equation 2 estimates the effect of the lottery after one and two years separately, instead of pooling the two effects together: γ_1 and γ_3 are respectively equal to the difference between lottery winners' and losers' average years of enrollment one and two years after the lottery. Variables $S_{i1} \in [0, 1]$ and $S_{i2} \in [0, 2]$ do not only take integer values: some students dropped out from the boarding school during the academic year, in which case we compute fractions of years based on the number of days actually spent in the boarding school. Estimates of γ_0, γ_1, and γ_3 are displayed in the second, third, and fifth columns of panel A. The seventh column reports the p-value of a test of $\gamma_1 = \gamma_3$.

At the end of the first year, lottery losers had spent 5.3 percent of a year in the boarding school on average. This reflects the fact that about 6 percent of them entered the boarding school during the first year, and most of them stayed for the year. At that point, lottery winners had spent on average 0.773 more years at the boarding school than control students. Two years after the randomization, they had spent 1.327 more years there.[5]

Panel B in Table 6 displays intention-to-treat (ITT) estimates, i.e. estimates of the effect of winning the lottery on students' French and mathematics test scores. Specifically, we estimate the following equation:

$$Y_{it} = \alpha_0 1\{t = 1\} + \alpha_1 Z_i \times 1\{t = 1\} + \alpha_2 1\{t = 2\} + \alpha_3 Z_i \times 1\{t = 2\} + X_i'\zeta + \eta_{it}, \quad (3)$$

where Y_{it} is student i's test score t years after randomization, and η_{it} is a residual. As the first stage equation, this ITT equation estimates the effect of the lottery after one and two years separately: α_1 and α_3 are equal to the difference between lottery winners' and losers' average test scores, one and two years after the lottery respectively. Estimates of α_0, α_1, and α_3 are displayed in the second, third, and fifth columns of panel B. The seventh column reports the p-value of a test of $\alpha_1 = \alpha_3$. Standard errors are clustered at the student level. Clustering standard errors at the class level leaves the results almost unchanged, as shown in Table 18 in the Appendix.

Lottery winners start outperforming losers only two years after the lottery, and only on their mathematics scores. After one year, estimates of the effect of winning the lottery on

[5]The small differences between the first stage estimates reported in Tables 1 and 6 stem from the fact that in Table 1 we use the full sample, while in Table 6 we only use the sample of students who took at least one cognitive test.

French and mathematics scores are small and not statistically different from zero. After two years, the point estimate in French is still rather small and not significant. On the contrary, the point estimate in mathematics is large and significantly different from zero: by then, lottery winners score 25.1 percent of a standard deviation higher than losers.[6] As this panel contains four different estimates of the effect of the boarding school on test scores, one might worry that this significant effect might be a false positive. However, its Bonferroni adjusted p-value is 0.09 (see Abdi, 2007), the Bonferroni adjustment being conservative here because the four outcomes in the panel are highly correlated. The chances that this effect is actually a false positive are low. Finally, the effects on mathematics scores after one and two years significantly differ at the 2 percent level.

Panel C in Table 6 displays local average treatment effects estimates, i.e. estimates of the average effect of spending one year in the boarding school among students who complied with their lottery offer (see Angrist & Imbens, 1995). Specifically, we estimate the following equation by two-stage least squares (2SLS):

$$Y_{it} = \beta_0 \times 1\{t=1\} + \beta_1 S_{i1} \times 1\{t=1\} + \beta_2 1\{t=2\} + \beta_3 S_{i2} \times 1\{t=2\} + X'_{it}\zeta + \mu_{it}, \quad (4)$$

using $Z_i \times 1\{t = 1\}$ and $Z_i \times 1\{t = 2\}$ as excluded instruments for $S_{i1} \times 1\{t = 1\}$ and $S_{i2} \times 1\{t = 2\}$. As the first stage and ITT equations, this 2SLS equation estimates the effect of the boarding school after one and two years separately: β_1 and β_3 are equal to the difference between lottery winners' and losers' average test scores, one and two years after the lottery, respectively, normalized by the difference in the number of years spent in the boarding school between these two groups at each date. Estimates of the mean of test scores for compliers in the control group one year after the lottery are displayed in the second column of the panel. (We follow the method described in Abadie (2003) to estimate this quantity.) Estimates of β_1, and β_3 are displayed in the third and fifth columns. The seventh column reports the p-value of a test of $\beta_1 = \beta_3$.

Two years after the lottery, the magnitude of our 2SLS estimates is consistent with previous findings from the literature. At this date, our estimates indicate that the boarding school increases compliers' mathematics scores by 20 percent of a standard deviation per year spent in the school. Furthermore, it has no effect on scores in French.

Research studying the effects of educational policies in middle and high school has often found low or zero effects in language, and effects on mathematics scores similar to the one we

[6] The number of observations in mathematics and French are different, as these two tests were taken on different days, as explained in Section 2. For instance, some students who took the French test missed the math test because they were sick on the day when it took place.

show here. For instance, in the charter school literature, Dobbie & Fryer (2011) find that the Promise Academy School in Harlem increases students mathematics test scores by 23 percent of a standard deviation per year spent in the school, but has no effect on their English scores. In Boston, Abdulkadiroglu et al. (2011) and Angrist et al. (2010) find larger effects than those we report here, but they also find stronger effects in mathematics than in English (+35 percent versus +12 percent of a standard deviation per year spent in the school). There is no consensus yet on why many middle and high school interventions have larger returns on mathematics than on language test scores. Some cognitive psychologists have argued that language ability might be set during childhood while numerical ability might continue to evolve during adolescence (see e.g. Hopkins & Bracht, 1975) Also, language is learned and manipulated at home, whereas mathematics is more exclusively a school topic - which may make it more dependent on teaching quality. One of the few exceptions to this language versus mathematics divide is Curto & Fryer (2014), who study the SEED Boarding School in Washington, D.C., the closest school to the one we study here for which causal effects on test scores are available. They find comparable effects to ours in mathematics, and larger effects in English (+23 and +20 percent of a standard deviation per year spent in the school, respectively). As a potential explanation for their result, the authors argue that boarding schools might be more efficient than other interventions at raising language ability if students speak no or little English in their home environment. We do not find evidence of this here: even when we focus on students for whom French is not the only language spoken at home, we still find insignificant effects of the boarding school on their French test scores, even though we lack statistical power to make definitive conclusions.

Our most surprising findings are the absence of any effect on test scores one year after the lottery and the evolution of the treatment effect between the first and second year. First, given all the positive inputs the boarding school provides to students – smaller class size, less classroom disruption, more engaged teachers, higher achieving peers, supervised homework time –, one could have expected lottery winners to outperform lottery losers even one year after the lottery. Second, the positive impact of the boarding school on mathematics test scores emerges only two years after the lottery. The order of magnitude is large: overall, at the end of the second year, students who were exposed to the intervention for two years would have a test score in mathematics almost 40 percent (2×0.191) of a standard deviation larger than untreated students. Nevertheless, this arises entirely in the second year. This finding is in sharp contrast with most papers in the literature studying dynamic effects of educational interventions, which usually find effects linear or concave in the amount of exposure (see e.g. Krueger, 1999 or Abdulkadiroglu et al., 2011).

Table 6: Effect of the boarding school on test scores

Panel A: First stage estimates

	Control mean	FS after 1 year	SE	FS after 2 years	SE	FS 1 = 2	N
Years of treatment	0.053	0.773***	0.040	1.327***	0.084	0.000***	744

Panel B: Intention to treat estimates

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
French	0.022	-0.060	0.109	-0.112	0.121	0.626	744
Mathematics	0.023	-0.014	0.094	0.251**	0.110	0.015**	735

Panel C: Two stage least squares estimates

	$E(Y_0\|C)$	2SLS after 1 year	SE	2SLS after 2 years	SE	2SLS 1 = 2	N
French	0.014	-0.077	0.140	-0.085	0.091	0.948	744
Mathematics	-0.026	-0.019	0.120	0.191**	0.082	0.064*	735

Notes. Panel A reports coefficients from a regression of the number of years spent in the school on a dummy for year 1 (column 2), the interaction of this dummy with our lottery offer (column 3), a dummy for year 2, the interaction of this dummy with our lottery offer (column 5), and the statistical controls listed in Section 2.2, within the sample of students who took at least one cognitive test. Panel B reports coefficients from regressions of French and math test scores on the same explanatory variables, within the sample of students who took these tests. Panel C reports coefficients from 2SLS regressions of the French and math tests scores on a dummy for year 1, the interaction of this dummy with the number of years spent in the school after one year (column 3), a dummy for year 2, the interaction of this dummy with the number of years spent in the school after two years (column 5), and the statistical controls listed in Section 2.2, using our lottery offer interacted with the year 1 and year 2 dummies as instruments, within the sample of students who took these tests. The second column of this panel reports an estimate of the mean of French and math test scores for compliers not enrolled in the school. We use propensity score reweighting to control for lottery strata. Standard errors reported in columns 4 and 6 are clustered at the student's level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. *significant at 10%; **significant at 5%; ***significant at 1%.

4.2 Dynamic selection or increasing returns?

The effect of the school on math scores is more than twice as large after two years than after one year. There are two potential explanations for this pattern. First, students enrolled two years might benefit more from their second than from their first year. Second, students leaving between year one and two might benefit from lower effects than those who stay, leading to an increase of the average effect of the school over time. In this section, we try to tell apart these two potential explanations.

The effects we estimate arise from students who comply with our lottery offer. This population can be divided into two groups: students induced by our offer to enroll in year one but not in year two, and students induced to enroll in both years. Let us call movers (M) and stayers (S) the first and second group, respectively. We can estimate p_M (resp. p_S), the share of movers (resp. stayers) in the population, by regressing a dummy for being enrolled in year one only (resp. in year one and two) on our lottery offer. We find $p_M = 20.7$ percent and $p_S = 58.5$ percent. If no control group student had managed to enroll in the school, movers (resp. stayers) would merely be lottery winners enrolled in year one only (resp. in year one and two). In practice, as very few students from the control group managed to enroll, this distinction does not matter much. Students enrolled only in year one account for 22.2 percent of the treatment group, implying that only (22.2-20.7)/22.2=6.7 percent of them are actually not movers but always takers. Students enrolled both years account for 63.6 percent of the treatment group, implying that only (63.6-58.5)/63.6=8.0 percent of them are actually not stayers but always takers. In what follows, we omit this distinction, and we present our analysis as if treatment group students enrolled one and two years were the same as movers and stayers. This simplifies the exposition without affecting the substantive conclusions of the discussion.[7]

Let δ_1^M and δ_1^S denote the average effect produced by the boarding school on students' mathematics score after one year, among movers and stayers respectively. Accordingly, let δ_2^M and δ_2^S denote the corresponding effects after two years. δ_2^S is the *cumulated* effect of having been enrolled two years in the boarding school among stayers, while δ_2^M is the effect of

[7] In Table 7, we compare the mean of some covariates between lottery winners enrolled one year only and those enrolled both years. We could perform the same comparisons between movers and stayers. Let $1\{M\}$ denote a dummy for being enrolled only one year and let X be some covariate. The coefficient of $1\{M\}$ in a 2SLS regression of $X1\{M\}$ on $1\{M\}$ using the lottery offer as an instrument will estimate the mean of X among movers. The intuition is the same as that described in Angrist & Pischke (2008) to recover the mean of X among compliers in a standard instrumental variable setting with binary treatment. One can also estimate the mean of X among stayers. In practice, estimating these 2SLS regressions instead of what is currently shown in Table 7 hardly changes the results. Similarly, one can estimate average mathematics scores at baseline and one and two years after the lottery among movers and stayers, and use them to conduct a difference-in-differences analysis similar to that shown in Panel A of Table 8. Here as well, this hardly changes the results.

having been enrolled one year ago among movers. δ_1^M, δ_1^S, δ_2^M, and δ_2^S are the four structural parameters generating our ITT estimates. Indeed, it is easy to show that α_1 and α_3 in Equation (3) respectively satisfy

$$\alpha_1 = p_M\delta_1^M + p_S\delta_1^S \tag{5}$$

$$\alpha_3 = p_M\delta_2^M + p_S\delta_2^S. \tag{6}$$

This system of two equations with four unknowns is not identified. To be able to estimate, say δ_1^M, we would need to compare movers' scores to the scores of the control group students who would also have been enrolled only in year one if they had won the lottery. As we do not know who these students are, we cannot perform that comparison.

A whole range of parameter values are compatible with our estimates of α_1 and α_3 and with Equations (5) and (6), some of which do not imply increasing returns to the school for stayers. Figure 2 plots two interesting polar cases. In the left panel, δ_1^M is very negative, $\delta_2^M = 0$, and $\delta_2^S = 2\delta_1^S$. This corresponds to a scenario in which movers are hurt by their first year in the school but fully recover once they leave, while stayers' test scores are not convex, but increase linearly with exposure. In this scenario, our ITT effect after one year is a weighted average of a negative effect for movers and a positive effect for stayers; after two years, it reflects the positive effect on stayers. In the right panel, $\delta_1^M = \delta_2^M = \delta_1^S = 0$, and $\delta_2^S > 0$. This corresponds to a scenario where nobody benefits from their first year in the school, while stayers start benefiting in their second year. In this second scenario, stayers have larger returns to their first than to their second year ($\delta_2^S > 2\delta_1^S$).

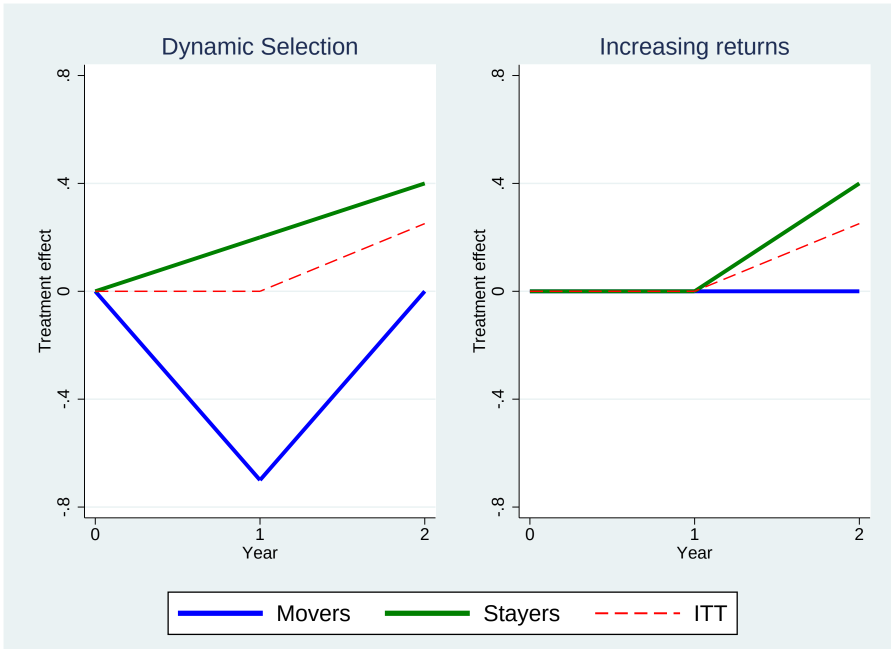

Figure 2: Two sets of parameter values compatible with our estimates.

In the first scenario, stayers and movers have highly heterogeneous returns to their first year in the boarding school. This is not entirely implausible as these two groups of students are quite heterogeneous in terms of baseline characteristics, as shown in Table 7. Movers have significantly weaker baseline academic results and they are more often disruptive students. Therefore, they might suffer from the strict discipline and high academic demands in the boarding school, which could eventually generate a large negative effect for them.

In order to further examine the possible interpretations of the increase in the average effect of the school over time, we conduct two suggestive tests which we report in Table 8. In Panel A, we estimate two differences-in-differences (DID). We compare the evolution of mathematics scores from baseline to year one and from year one to year two, between movers and the control group; we then run the same comparison between stayers and the control group. Specifically, we estimate the two following regressions, in the sample of movers and lottery losers, and in the sample of stayers and lottery losers respectively:

$$\begin{aligned} Y_{it} &= \lambda_0 + \lambda_1 1\{t \geq 1\} + \lambda_2 1\{t = 2\} + \lambda_3 1\{M\} + \lambda_4 1\{M\} 1\{t \geq 1\} + \lambda_5 1\{M\} 1\{t = 2\} + X_{it}'\zeta + \varepsilon_{it} \\ Y_{it} &= \gamma_0 + \gamma_1 1\{t \geq 1\} + \gamma_2 1\{t = 2\} + \gamma_3 1\{S\} + \gamma_4 1\{S\} 1\{t \geq 1\} + \gamma_5 1\{S\} 1\{t = 2\} + X_{it}'\zeta + \mu_{it}. \end{aligned}$$

where $1\{M\}$ (resp. $1\{S\}$) is a dummy for lottery winners enrolled in year one only (resp. in both years). ε_{it} and μ_{it} are residuals. Estimates of λ_3, λ_4, and λ_5 are reported in the second, third, and fifth columns of the first line of the panel. Estimates of γ_3, γ_4, and γ_5 are reported

Table 7: Comparison of treatment group students by enrolment status

	Movers	Stayers	P-value
Share in the population	0.222	0.636	
Baseline ability and disruptiveness			
Standardized mark in French, transcripts	-0.480	0.082	0.000***
Rank in French, transcripts	0.428	0.247	0.004***
Standardized mark in Mathematics, transcripts	-0.240	0.031	0.206
Rank in Mathematics, transcripts	0.347	0.259	0.126
Middle school exam, French	-0.384	0.162	0.005***
Middle school exam, Mathematics	-0.156	0.224	0.059*
Standardized school behavior grade, transcripts	-1.251	0.228	0.001***
Socio-economic background			
Means tested grant	0.397	0.481	0.272
Parent employee	0.231	0.252	0.761
Parent blue collar	0.293	0.280	0.856
Parent inactive	0.177	0.146	0.590
Parent has completed high school	0.301	0.225	0.276
Only French spoken at home	0.587	0.349	0.003***

Notes. This table compares lottery winners enrolled only in year one in the boarding school to those enrolled both years. Socio-economic variables come from the “Sconet” administrative data set. Transcripts come from application files. Grades for the end-of-middle-school exam come from the “Base Brevet” administrative data set.

in the second, third, and fifth columns of the second line. In the seventh column, we report p-values of tests of $\lambda_4 = \lambda_5$ and $\gamma_4 = \gamma_5$. For these DIDs to capture the causal effect of the boarding school among movers and stayers, a common trend assumption must be satisfied: in the absence of the treatment, lottery losers, movers, and stayers would all have experienced the same evolution of their scores from baseline to year one and from year one to year two.

The DID estimates resemble more the right than the left panel of Figure 2. From baseline to year one, movers' test scores increase slightly less than that of lottery losers, but this difference is small and insignificant. From year one to year two, their scores increase slightly more, but the difference is also small and insignificant. By contrast, the estimates for those who stayed two years display a marked convex pattern: between baseline and year one, their test scores do not follow a significantly different evolution from that in the control group; between year one and two they increase by 30.6 percent of a standard deviation more. Under the aforementioned common trends assumption, this implies that the increasing ITT effects of the school mostly come from the fact that stayers benefit more from their second than from their first year.

Panel B of Table 8 presents a second suggestive test. It builds upon the idea that if we could isolate a subpopulation in which all compliers were stayers, the ITTs in this subpopulation would only be driven by δ_1^S and δ_2^S. Then, a convex pattern in these ITTs would imply increasing returns for stayers in that subpopulation. As shown in Table 7, baseline characteristics can be used to predict whether a student will be enrolled one year only in the boarding school. Based on this observation, we estimate a probit model among lottery winners, with a dummy for being enrolled only in year one as our dependent variable, and with a dummy for whether French is the only language spoken at home, order two polynomials in baseline French, mathematics, and school behavior grades, and interaction terms between the polynomials and the dummy as explanatory variables.[8] Then, we estimate the predicted probability of being enrolled only one year for every student, and we divide each lottery stratum into two groups, depending on whether a student's predicted probability is above or below the median in her stratum. Finally, we estimate the effect of the lottery on enrollment in the school and mathematics scores as in Equation (3), separately in these two subgroups.

Our probit model predicts reasonably well who is going to leave: the exit rate during year one is 17 percent $(1 - 0.636/0.768)$ among students with a low probability to leave, against 33 percent $(1 - 0.572/0.853)$ among those with a high probability.[9] Despite these differences

[8] We do not include middle-school exam grades because they are available for only half of the students.

[9] Note that students with a low probability to leave have a higher probability to enroll when they lose the lottery than those with a high probability to leave (9.0 percent against 3.0 percent). This explains most of the difference in the effect of the lottery on enrollment in year one between these two groups (+76.8 and +85.3

in enrollment patterns, ITT effects on mathematics scores do not substantially differ in the two groups. In both groups, the point estimates are close to zero after one year, while they are large and significant after two years.[10] In the low probability group, p_M and p_S are respectively equal to 13.2 and 63.6%, so ITTs closely reflect δ_1^S and δ_2^S. The increase in the effect of the school in this group is very likely to be driven by increasing returns to the school among stayers.

Overall, the two suggestive tests presented in Table 8 point towards the same conclusion: the increasing ITT effects of the school on students' mathematics test scores seem to arise from the fact that students staying two years benefit much more from their second than from their first year, not from dynamic selection into the treatment.

4.3 Distributional and heterogeneous effects.

To conclude this section, we explore whether the average effects displayed in Table 6 hide heterogeneity along the distribution of the outcome. We focus on effects after two years in mathematics, as this is where average effects are statistically significant.[11] Figure 3 displays unconditional quantile treatment effects (QTE), following Firpo et al. (2009), and using the lottery indicator Z_i as the treatment variable. QTE estimates should therefore be compared to ITT estimates in Table 6, panel B (+0.251 of a standard deviation).[12]

Our lottery offer has a positive effect on the upper part of the distribution of the outcome, but has a negative effect on the lower part. Quantile treatment effects are: negative and significant in the lower decile, around -0.3 standard deviation of the outcome; positive and marginally significant in the middle of the distribution, around +0.3 standard deviation; large, positive, and significant in the upper quintile, around +0.7 standard deviation. Overall, the lottery offer produces a strong increase in the variance of the outcome.

---

percent, respectively).

[10] Both point estimates are slightly larger than that appearing in Table 6 for the entire population. When we divide lottery strata into the high- and low-probability subgroups as described above, we end up with some strata with no treatment or control group students. We therefore have to drop these strata from our estimation. One can see that the number of observations used in the two regressions of Panel B of Table 8 does not sum up to that in the regression of mathematics scores in Panel B of Table 6. Thus, this is the source of the slight discrepancy.

[11] Results on French and after one year are in the Appendix. Most quantile treatment effects for these outcomes are small and insignificant.

[12] As our treatment variable is not binary, we cannot use the instrumental variable quantile treatment effect estimator proposed in Abadie et al. (2002) or Froelich & Melly (2013).

Table 8: Increasing effects in mathematics: convexity or heterogeneity?

Panel A: Evolution of scores of movers/stayers compared to lottery losers

	Diff 0	DID 1 - 0	SE	DID 2 - 1	SE	1-0 = 2-1	N
Movers vs. lottery losers	-0.269	-0.036	0.164	0.033	0.161	0.794	550
Stayers vs. lottery losers	0.136	-0.090	0.125	0.306**	0.121	0.054*	885

Panel B: ITT effects according to probability of leaving during or after first year

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
Low proba to leave							
Enrolled	0.090	0.768***	0.062	0.636***	0.070	0.000***	387
Mathematics score	0.409	-0.039	0.137	0.287*	0.170	0.067*	387
High proba to leave							
Enrolled	0.030	0.853***	0.052	0.572***	0.065	0.000***	331
Mathematics score	-0.411	0.059	0.122	0.278**	0.125	0.102	331

Notes. The first line of panel A reports coefficients from a regression of students math test scores at baseline and in year 1 and 2, on a constant, a dummy for year 1 and 2, a dummy for year 2, a dummy for lottery winners enrolled in year one only (column 2), the interaction of this dummy and the dummy for year 1 and 2 (column 3), the interaction of this dummy and the dummy for year 2 (column 5), and the statistical controls listed in Section 2.2, within the sample of lottery losers and winners enrolled only in year 1. The second line reports coefficients from a regression of students math test scores at baseline and in year 1 and 2, on a constant, a dummy for year 1 and 2, a dummy for year 2, a dummy for lottery winners enrolled in both years (column 2), the interaction of this dummy and the dummy for year 1 and 2 (column 3), the interaction of this dummy and the dummy for year 2 (column 5), and the statistical controls listed in Section 2.2, within the sample of lottery losers and winners enrolled in both years. The upper part of panel B reports coefficients from regressions of students enrolment status and math test scores on a dummy for year 1 (column 2), the interaction of this dummy with our lottery offer (column 3), a dummy for year 2, the interaction of this dummy with our lottery offer (column 5), and the statistical controls listed in Section 2.2, within the sample of students with a low probability of leaving who took at least one cognitive test. The lower part of panel B reports the same coefficients from the same regressions, within the sample of students with a high probability of leaving who took at least one cognitive test. We use propensity score reweighting in both panels. Standard errors reported in columns 4 and 6 are clustered at the student's level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. *significant at 10%; **significant at 5%; ***significant at 1%.

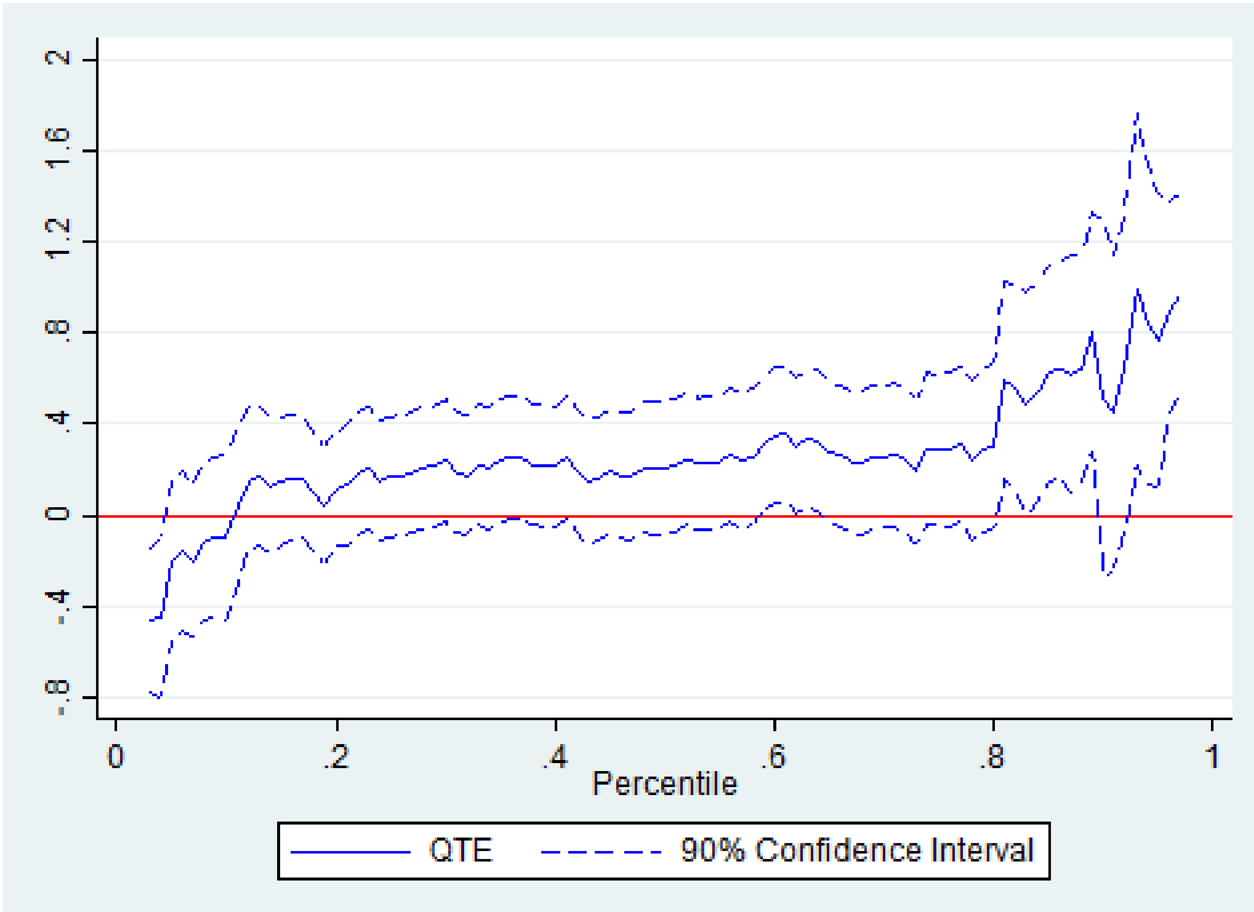

Figure 3: QTE in Mathematics after 2 years, intention-to-treat.

Under the assumption that the boarding school does not change the rank of a student in the distribution of mathematics scores, these findings imply that winning the lottery is mostly beneficial to the strongest students. To test the validity of this interpretation, we investigate heterogeneous treatment effects according to baseline ability in math. Given the sharp difference between quantile treatment effects in the upper part and in the rest of the distribution, we compare ITT estimates for students in the top tercile of baseline math scores and for those in the middle and bottom terciles.[13] Table 9 reproduces Table 6 for those two subgroups. Panel B shows that the 0.251 ITT effect of Table 6 is actually the average of a large, positive, and highly significant effect in the upper tercile (+0.605) and of a small and non significant impact in the other two terciles. As the effect of the lottery on actual enrollment is very similar in the two subgroups (Panel A), the 2SLS estimates are also very different in these two populations (Panel C).

[13]When we disaggregate the middle and bottom terciles, we do not find any significant difference between the ITT effects in these two terciles.

To sum up, assignment to the boarding school has a large positive impact on math scores after two years, whose magnitude is comparable to available estimates of charter school impacts in the United States. However, two surprising results emerge: the positive value-added of the boarding school only emerges after two years, and even at that time, it is mostly concentrated among students with higher initial ability. There is even evidence suggesting that a non-negligible share of lottery winners are actually harmed by the offer to enter the school.

Table 9: Heterogeneous effects, according to baseline mathematics scores

Panel A: First stage estimates

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
In upper tercile at baseline	0.049	0.781***	0.068	1.265***	0.140	0.000***	237
Out of upper tercile at baseline	0.057	0.785***	0.054	1.338***	0.134	0.000***	472
P-value In = Out		0.968		0.706			

Panel B: Intention to treat estimates

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
In upper tercile at baseline	0.763	-0.017	0.197	0.605***	0.196	0.006***	237
Out of upper tercile at baseline	-0.331	0.019	0.095	0.090	0.119	0.545	472
P-value In = Out		0.869		0.025**			

Panel C: Two stage least square estimates

	$E(Y_0\|C)$	2SLS after 1 year	SE	2SLS after 2 years	SE	2SLS 1 = 2	N
In upper tercile at baseline	0.835	-0.048	0.252	0.493***	0.150	0.027**	237
Out of upper tercile at baseline	-0.410	0.026	0.125	0.070	0.092	0.715	472
P-value In = Out		0.790		0.016**			

Notes. The first line of Panel A reports coefficients from the same regression as that in Panel A of Table 6, within the sample of students who took at least one math test and who were in the first tercile of math scores in their lottery stratum at baseline. The second line reports the same coefficients from the same regression, within the sample of students who took at least one cognitive test and who were not in the first tercile of math scores in their lottery stratum at baseline. In the third (reps. fifth) column of the third line of the panel, we report p-values of a test of equality of the coefficients reported in the third (resp. fifth) column of the first and second lines. Accordingly, Panel B and C reproduce results for math scores in Panel B and C of Table 6, separately for students in and out of the first tercile of math scores at baseline. We use propensity score reweighting to control for lottery strata. Standard errors reported in columns 4 and 6 are clustered at the student's level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. *significant at 10%; **significant at 5%; ***significant at 1%.

5 Ready for boarding?

Standard education production functions usually posit that a student's test score is an increasing function of the quality of the learning environment, of the ability of the student's peers, and of the student's own effort. The students' learning environment improves on a number of dimensions in the boarding school. As shown in Section 3, classes are 25 percent smaller, teachers are more engaged, and levels of classroom disruption are lower. Peer quality is also higher. As shown in Table 2, applicants to the boarding school scored 40 percent of a standard deviation higher than their classmates in the national middle-school exam before applying to the boarding school. Under the assumption that the peer quality of control group students did not change much after the lottery, this means that boarders benefit from substantially better peers than control group students. Students' effort should also be higher. As shown in Table 5, boarders must spend one hour and a half every evening in a study room to do their homework, while control group students have no similar requirement.

Under this simple production function, results from the previous section are puzzling. First, it is surprising that exposure to such an intensive treatment has no effect after one year. Then, even if positive effects after two years are more in-line with expectations, it is not clear why the second year of treatment has very different effects from the first one. Indeed, the supplementary inputs the school is providing to students have not changed between the two years. Tables 3, 4, and 5 described the treatment by comparing schooling conditions for boarders and control students two years after the lottery. In Tables 19, 20, and 21 shown in the Appendix, we reproduce similar tables, in which we also report the differences in schooling conditions for boarders and control students one year after the lottery, and the result of a test for whether the difference after one year significantly differs from that after two years.[14] There is no evidence that the nature or the intensity of the treatment changed between the two years. One year after the lottery, lottery winners already benefit from smaller classes and better studying conditions, and they already praise the quality of their teachers. Of the 35 tests we conduct to assess whether school conditions experienced by boarders and control students changed between the first and second year, only three have have a p-value lower than 0.10. If anything, studying conditions in the boarding school seem to have slightly deteriorated. Class size slightly increased, relationships among students deteriorated as per one of our four measures, and tensions between students and their supervisors increased as per one of our six measures.

This suggests that this simple production function omits an important input. Our pre-

[14]Unfortunately, one year after the lottery not all measures are available for the first cohort of students, and, as a result, the samples in the supplementary tables are sometimes smaller than in the baseline tables.

ferred candidate is students' well-being. We now review the evidence supporting this interpretation, before discussing other potential mechanisms.

When they arrive in the boarding school, students need to adjust to a number of negative changes. They have to cope with the separation from their friends and families; they relinquish a certain amount of freedom; and they face higher academic demands. This probably explains why one year after the lottery, levels of school well-being were significantly lower among boarders, as shown in Table 10.[15] At that date, as per our standardized score, lottery winners' well-being is reduced by 32 percent of a standard deviation. When we look separately at the eight items included in our score, we find two significant differences: boarders felt more uncomfortable in school, and they were more likely to think that other students did not like them. These two effects are unlikely to be false positives. Using a Bonferroni adjustment to account for the fact we have eight measures of well-being, we obtain p-values respectively equal to 0.04 and 0.08. Also, although they are not significant, all the other effects point to a reduction in well-being.

In the end of their second year, students seem to have adjusted to their new environment. At this point, the well-being score is slightly higher for boarders than for control students, and we can reject at the 5 percent level that the effect of the boarding school is the same in year one and year two. We also reject this test at the 10 percent level for two items of this score out of eight. For seven items out of eight, the point estimates indicate that students' well-being increased between the two years. We also measure the effect of the boarding school on students' academic, social, and general self-esteem, using the French translation of the Self-Perception Profile for Adolescents (see Bouffard et al., 2002). The effect of the boarding school on students academic self-esteem is insignificant both after one year and after two years, but it significantly increases over time (p-value =0.065).

At the same time that levels of well-being catch-up, students' motivation increases, and they start spending more time on their homework. To measure students' motivation for schooling, we use the "motivation for education" scale (see Vallerand et al., 1989). Whereas one year after the lottery there were no noticeable differences between boarders and control students on any of its three sub-scales (extrinsic and intrinsic motivation, and amotivation), after two years boarders have more intrinsic motivation for schooling as shown in Table 11. Moreover, the effect of the school on students' amotivation significantly decreases between year one and two.

Similarly, although after one year, boarders did not report spending more time per week

[15] As school well-being questions were not included in the questionnaires administered to the first cohort one year after the lottery, we only report results for the second cohort.

Table 10: Effects of the school on well-being and self-esteem

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
School well-being							
School well-being score	0.176	-0.322*	0.166	0.085	0.173	0.032**	364
In school, I feel like a stranger	-0.095	0.123	0.164	-0.025	0.184	0.482	398
I have few friends	0.077	-0.092	0.178	0.002	0.184	0.666	398
I feel home	0.148	-0.228	0.173	0.256*	0.150	0.028**	398
I feel uncomfortable	-0.116	0.514***	0.182	0.243	0.193	0.243	397
Other students like me	0.157	-0.468**	0.182	-0.054	0.175	0.098*	365
I feel lonely	-0.071	0.000	0.156	0.024	0.162	0.910	398
I do not want to go	-0.097	0.046	0.180	-0.017	0.168	0.754	398
I am often bored	-0.109	0.263	0.178	-0.065	0.166	0.136	398
Self-Esteem							
Academic Self-Esteem	0.077	-0.141	0.109	0.079	0.125	0.065*	735
Social Self-Esteem	0.052	-0.026	0.136	0.034	0.130	0.579	734
General Self-Esteem	0.081	0.010	0.117	0.087	0.146	0.538	734

Notes. This table reports coefficients from regressions of the outcomes listed in the first column on a dummy for year 1 (column 2), the interaction of this dummy with our lottery offer (column 3), a dummy for year 2, the interaction of this dummy with our lottery offer (column 5), and the statistical controls listed in Section 2.2, within the sample of students for whom these outcomes are available at least one year. For well-being, our estimation sample is the second cohort of students, as well-being measures are not available one year after the lottery for the first cohort. We use propensity score reweighting to control for lottery strata. Standard errors reported in columns 4 and 6 are clustered at the student's level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. All the variables come from students' questionnaires. The school well-being score is standardized; it is computed from the individual variables listed below. Self-esteem scores are also standardized and are based on Bouffard et al. (2002). *significant at 10%; **significant at 5%; ***significant at 1%.

on their homework, after two years lottery winners spend 25 percent more time on it than lottery losers. During school days, boarders spend more time on their homework and less time watching TV or playing video games. This effect is somewhat mechanical, merely reflecting the rules in the boarding school: differences are large and quite constant over time. The increase in total homework time during the second year seems to be driven by week-end behavior. Although we lack statistical power to make definitive conclusions, it seems that during the first year, treated students tend to compensate weekday effort by relaxing more during the week-end. After two years, this pattern has changed markedly: boarders now spend more time on their homework and less time watching TV or playing video games during the week-ends. This is consistent with the increase in their intrinsic motivation we observe between the first and the second year. None of these three evolutions between year one and two – time spent on homework, television and video games on Saturdays – are statistically significant, but the estimates all go in the same direction. To gain power, we compute the difference between homework and "screen-time", so as to concentrate this consistent information into one coefficient. Both the substitution between homework and screen time on Saturdays during the first year and the reversal after the second year are now significant.

Finally, we find some indication that the initial negative shock on well-being and motivation is more pronounced among weaker students, and that the recovery is faster for stronger students, although we lack statistical power to make definitive conclusions. This could explain why even after two years, only high-performing students seem to benefit from the school. In Table 12, we report ITT effects of the school on the outcomes of Tables 10 and 11 for which we found different effects after one and two years, distinguishing students in the upper tercile of math scores at baseline from those in the middle and bottom terciles. After one year, weaker students have more negative effects on each of these five outcomes, even though none of the differences is statistically significant. Between year one and year two, effects increase more for stronger than for weaker students on four outcomes out of five, even though once again these differences are not significant.

To sum up, we find that the school has a negative effect on students' well-being after one year, which reverses in the second year. This could explain why its positive effect on cognitive outcomes and on a number of measures of motivation and effort only appear in the second year, although from their first year onwards boarders experience a number of positive inputs. Results from other studies also point towards a positive link between well-being and learning. Ly et al. (2013) study the transition from middle school to high school in France, where students change schools and, as a result, part from most of their previous classmates. They find that being assigned to a high school class with more of one's previous classmates

Table 11: Effects of the school on students motivation and effort

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
Motivation for schooling							
Extrinsic motivation	-0.026	-0.124	0.134	-0.030	0.126	0.556	732
Intrinsic motivation	-0.010	0.034	0.124	0.330***	0.119	0.023**	732
Amotivation	0.011	0.179	0.181	-0.173	0.151	0.076*	732
Hours spent last week...							
Doing homework	6.090	0.042	0.491	1.581***	0.514	0.013**	720
Hours spent last Monday...							
Doing homework	1.302	0.357***	0.130	0.458***	0.127	0.477	722
Playing video games	0.500	-0.274**	0.131	-0.141	0.116	0.321	714
Watching TV	1.378	-0.835***	0.145	-0.670***	0.168	0.361	722
Homework -(video games+TV)	-0.579	1.465***	0.257	1.233***	0.287	0.434	703
Hours spent last Saturday...							
Doing homework	1.672	-0.138	0.199	0.266	0.196	0.118	721
Playing video games	1.175	0.333	0.236	-0.039	0.294	0.168	715
Watching TV	2.671	0.280	0.296	-0.144	0.281	0.225	720
Homework -(video games+TV)	-2.147	-0.736*	0.391	0.493	0.452	0.013**	696

Notes. This table reports coefficients from regressions of the outcomes listed in the first column on a dummy for year 1 (column 2), the interaction of this dummy with our lottery offer (column 3), a dummy for year 2, the interaction of this dummy with our lottery offer (column 5), and the statistical controls listed in Section 2.2, within the sample of students for whom these outcomes are available at least one year. We use propensity score reweighting to control for lottery strata. Standard errors reported in columns 4 and 6 are clustered at the student's level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. All the variables come from students' questionnaires. Motivation scores are standardized; they are computed from the "motivation for education" scale (see Vallerand et al. (1989)). *significant at 10%; **significant at 5%; ***significant at 1%.

Table 12: Effects on non-cognitive outcomes, according to baseline scores

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
School well-being							
In upper tercile at baseline	0.130	-0.235	0.360	0.307	0.297	0.126	121
Out of upper tercile at baseline	0.164	-0.256	0.187	-0.040	0.214	0.297	234
P-value In = Out		0.958		0.342			
Academic self-esteem							
In upper tercile at baseline	0.488	-0.004	0.169	0.354	0.216	0.122	239
Out of upper tercile at baseline	-0.115	-0.169	0.131	0.047	0.148	0.112	470
P-value In = Out		0.440		0.241			
Intrinsic motivation							
In upper tercile at baseline	0.045	0.201	0.223	0.549**	0.226	0.058*	236
Out of upper tercile at baseline	-0.063	0.060	0.173	0.315**	0.145	0.174	470
P-value In = Out		0.617		0.384			
Amotivation							
In upper tercile at baseline	-0.282	0.067	0.259	-0.311	0.236	0.160	236
Out of upper tercile at baseline	0.166	0.182	0.237	-0.182	0.192	0.179	470
P-value In = Out		0.744		0.672			
Hours spent on homework							
In upper tercile at baseline	5.664	1.259	0.977	2.495***	0.821	0.196	236
Out of upper tercile at baseline	6.024	-0.375	0.535	1.280**	0.612	0.039**	458
P-value In = Out		0.143		0.236			

Notes. The first line of the table reports coefficients from the same regression as that in the first line of Table 10, within the sample of students who took at least one math test and who were in the first tercile of math scores in their lottery stratum at baseline. The second line reports the same coefficients from the same regression, within the sample of students who took at least one cognitive test and who were not in the first tercile of math scores in their lottery stratum at baseline. In the third (reps. fifth) column of the third line of the panel, we report p-values of a test of equality of the coefficients reported in the third (resp. fifth) column of the first and second lines. Accordingly, the remaining lines of the table reproduce results for academic self-esteem, intrinsic motivation, amotivation, and weekly hours spent on homework shown in Tables 10 and 11, separately for students in and out of the first tercile of math scores at baseline. We use propensity score reweighting to control for lottery strata. Standard errors reported in columns 4 and 6 are clustered at the student's level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. All the variables come from students' questionnaires. All measures except hours spent on homework are standardized. *significant at 10%; **significant at 5%; ***significant at 1%.

from middle school significantly reduces subsequent grade repetition and drop-out rates. The effect is mostly driven by low socioeconomic-status students who move to high schools with more high socioeconomic-status students than had been present in their middle schools. This is evidence that maintaining earlier social ties, which presumably has a positive effect on well-being, also has positive effects on learning. The interactions between well-being and learning have also long been documented by educational and cognitive psychologists (see e.g. Boekaerts, 1993 or Williams et al., 1988).

The boarding school induces strong non-cognitive adjustments; interactions between non-cognitive and cognitive dimensions can account both for the time pattern we observe in this experiment, and for the larger effects we observe among initially stronger students.

But students' well-being is not the only missing input in the simple production function introduced above that could account for our findings. A first alternative candidate could be distance to teachers' target level of instruction, as in Duflo et al. (2011). If teachers in the boarding school tend to target their highest achieving students, this could explain why weaker students do not improve, even after two years. This interpretation is not entirely consistent with our data, however. First, we checked whether the increase in student's opinion about their teachers reported in Table 4 is larger for strong students than for weak students. Indeed, if boarding school teachers target strong students, the increase in students' satisfaction should be larger for them. Appendix Table 22 shows that, if anything, the increase in students' satisfaction is larger for weak students. Second, this mechanism cannot explain why strong students do not benefit from their first year in the boarding school.

A second alternative candidate could be students' rank in the classroom distribution. Recent research has indeed shown that higher within-class ordinal position has a positive effect on academic performance (see e.g. Murphy & Weinhardt (2013)). This can explain why weaker students do not improve in the boarding school, as they lose many ranks when they join. However, this still fails to explain why strong students do not improve during their first year: these students do not lose many ranks when they join, and accordingly their academic self-esteem does not seem affected at all in the end of their first year (cf. Table 12).

6 Conclusion

The boarding school we study produces surprising effects. It increases students' math test scores only two years after admission, even though we cannot find any evidence that the supplementary educational inputs provided by the school changed between the two years. We argue that an education production function in which students' well-being interacts with their studying conditions can account for this pattern. Indeed, we find that levels of well-being were lower among boarders one year after admission, probably due to the separation from their friends and families and to the strict discipline and high academic demands in the boarding school. By contrast, two years after admission boarders seemed to have adjusted to their new environment: levels of well-being had caught up with that in the control group, and they also started showing higher levels of motivation. We also find that effects after two years mostly come from the strongest students at baseline. The boarding school does not seem well-suited to weaker students: even after two years they do not experience any strong increase in their test scores. In future work, we will investigate the effects of this boarding school on students' higher education and labor market outcomes.

References

Abadie, A. (2003), 'Semiparametric instrumental variable estimation of treatment response models', *Journal of Econometrics* **113**(2), 231–263.

Abadie, A., Angrist, J. & Imbens, G. (2002), 'Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings', *Econometrica* **70**(1), 91–117.

Abdi, H. (2007), 'Bonferroni and šidák corrections for multiple comparisons. in nj salkind (ed.). encyclopedia of measurement and statistics. thousand oaks, ca: Sage'.

Abdulkadiroglu, A., Angrist, J., Dynarski, S. M., Kane, T. J. & Pathak, P. A. (2011), 'Accountability and flexibility in public schools: Evidence from boston's charters and pilots', *The Quarterly Journal of Economics* **126**(2), 699–748.

Abdulkadiroğlu, A., Angrist, J. & Pathak, P. (2014), 'The elite illusion: Achievement effects at boston and new york exam schools', *Econometrica* **82**(1), 137–196.

Adams, D. W. (1995), *Education for Extinction: American Indians and the Boarding School Experience, 1875-1928.*, ERIC.

Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A. & Walters, C. R. (2010), 'Inputs and impacts in charter schools: Kipp lynn', *American Economic Review* **100**(2), 239–43.

Angrist, J. D. & Imbens, G. W. (1995), 'Two-stage least squares estimation of average causal effects in models with variable treatment intensity', *Journal of the American statistical Association* **90**(430), 431–442.

Angrist, J. D. & Pischke, J.-S. (2008), *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Behaghel, L., de Chaisemartin, C., Charpentier, A. & Gurgand, M. (2013), 'Les effets de l'internat d'excellence de sourdun sur les éleves bénéficiaires: résultats d'une expérience contrôlée'.

Boekaerts, M. (1993), 'Being concerned with well-being and with learning', *Educational Psychologist* **28**(2), 149–167.

Bouffard, T., Seidah, A., Mcintyre, M., Boivin, M., Vezeau, C. & Cantin, S. (2002), 'Mesure de l'estime de soi à l'adolescence : version canadienne-française du self-perception profile for adolescents de harter', *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* **34**(3), 158–162.

Cookson, W., Cookson Jr, P. W. & Persell, C. (2008), *Preparing for power: America's elite boarding schools*, Basic Books.

Curto, V. E. & Fryer, R. G. (2014), 'The potential of urban boarding schools for the poor: Evidence from seed', *Journal of Labor Economics* **32**(1), 65–93.

de Janvry, A., Dustan, A. & Sadoulet, E. (2012), 'The benefits and hazards of elite high school admission: Academic opportunity and dropout risk in mexico city', *Unpublished Manuscript, University of California at Berkeley* .

DGESCO (2010), *Éducation prioritaire, Les réseaux "ambition réussite" et les réseaux de réussite scolaire, Académie de Créteil Données 2009-2010*, Direction générale de l'enseignement scolaire.

Dobbie, W. & Fryer Jr, R. G. (2013), 'The impact of attending a school with high-achieving peers: Evidence from the new york city exam schools', *American Economic Journal: Applied Economics* .

Dobbie, W. & Fryer, R. G. (2011), 'Are high-quality schools enough to increase achievement among the poor? evidence from the harlem children's zone', *American Economic Journal: Applied Economics* **3**(3), 158–87.

Duflo, E., Dupas, P. & Kremer, M. (2011), 'Peer effects, teacher incentives, and the impact of tracking', *American Economic Review* **101**(5 (August 2011)).

Firpo, S., Fortin, N. M. & Lemieux, T. (2009), 'Unconditional quantile regressions', *Econometrica* **77**(3), pp. 953–973.
URL: *http://www.jstor.org/stable/40263848*

Froelich, M. & Melly, B. (2013), 'Unconditional quantile treatment effects under endogeneity', *Journal of Business & Economic Statistics* **31**(3), 346–357.

Frölich, M. (2007), 'Nonparametric iv estimation of local average treatment effects with covariates', *Journal of Econometrics* **139**(1), 35–75.

Hanushek, E. & Rivkin, S. (2006), Teacher quality, Vol. 2, Elsevier: North-Holland, chapter 18.

Hopkins, K. D. & Bracht, G. H. (1975), 'Ten-year stability of verbal and nonverbal iq scores', *American Educational Research Journal* **12**(4), 469–477.

Krueger, A. (1999), 'Experimental estimates of education production functions', *Quarterly Journal of Economics* **114**(2), 497–532.

Lucas, A. M. & Mbiti, I. (2012), 'Effects of school quality on student achievement: Discontinuity evidence from kenya', *Southern Methodist University* .

Ly, S. T., Riegert, A. et al. (2013), 'Persistent classmates: How familiarity with peers protects from disruptive school transitions'.

Murphy, R. & Weinhardt, F. (2013), 'Top of class: The importance of ordinal rank position'.

Piketty, T. & Valdenaire, M. (2006), *L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français. Estimations à partir du panel primaire 1997 et du panel secondaire 1995*, Vol. 173 of *Les Dossiers*, Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche Direction de l'Évaluation et de la prospective.

Rosenbaum, P. R. & Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Vallerand, R. J., Blais, M. R., Brière, N. M. & Pelletier, L. G. (1989), 'Construction et validation de l'échelle de motivation en éducation (eme).', *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* **21**(3), 323.

Williams, J. M. G., Watts, F. N., MacLeod, C. & Mathews, A. (1988), *Cognitive psychology and emotional disorders.*, John Wiley & Sons.

A Appendix, for online publication

Table 13: Balancing checks

	Control Mean	T-C	SE	N
Ability and disruptiveness				
Grade in French	12.209	-0.125	0.343	392
Grade in Maths	12.437	0.189	0.393	392
Studies latin or greek	0.286	-0.097*	0.059	375
Studies german	0.266	-0.051	0.056	375
School behavior grade	15.304	0.390	0.508	343
Times missed school last term	5.596	0.600	0.687	350
Socio-economic background				
Parent blue collar or employee	0.448	-0.035	0.061	393
Recipient of means tested grant	0.404	0.013	0.062	393
Number of children in the family	2.827	-0.100	0.196	393
Parents divorced	0.264	-0.020	0.063	351
Single-parent family	0.376	-0.075	0.064	353
Parent has no degree	0.106	-0.015	0.052	347
Parent completed high school	0.217	0.011	0.058	347
Only French spoken at home	0.366	0.064	0.062	353

Notes. This table reports results from regressions of the outcomes in the first column on a constant and a dummy for our lottery offer. The second column reports the coefficient of the constant, while the third reports the coefficient of the dummy. Standard errors in column 4 are robust. We use propensity score reweighting to control for lottery strata. Measures of baseline ability and disruptiveness come from application files. Socio-economic variables come from the "Sconet" administrative data set. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 14: ITT effects on the share of students spending more time than allowed on the tests.

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
French	0.108	0.007	0.037	-0.012	0.046	0.722	722
Maths	0.000	0.005	0.006	0.011	0.016	0.731	712

Notes. This table reports coefficients from regressions of dummies for whether a student spent more time than allowed on the French and Maths test on a dummy for year 1 (column 2), the interaction of this dummy with our lottery offer (column 3), a dummy for year 2, the interaction of this dummy with our lottery offer (column 5), and the statistical controls listed in Section 2.2, within the sample of students for whom these outcomes are available at least one year. We use propensity score reweighting to control for lottery strata. Standard errors reported in columns 4 and 6 are clustered at the student's level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 15: ITT effects on test scores, excluding tests taken at home

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
French	-0.001	-0.050	0.109	-0.104	0.122	0.643	712
Mathematics	0.029	-0.018	0.094	0.333**	0.138	0.010***	704

Notes. This table reports coefficients from the same regressions as those presented in Panel B of Table 6, excluding tests which were taken at home by the student. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 16: Response rates to surveys

	Control Mean	T-C	SE	N
One year after the lottery				
Took the French test	0.928	-0.024	0.019	395
Took the maths test	0.922	-0.029	0.020	395
Two years after the lottery				
Took the French test	0.904	-0.021	0.022	395
Took the maths test	0.887	-0.013	0.029	395

Notes. This table reports results from regressions of the outcomes in the first column on a constant and a dummy for our lottery offer. The second column reports the coefficient of the constant, while the third reports the coefficient of the dummy. Standard errors in column 4 are robust. We use propensity score reweighting to control for lottery strata. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 17: ITT effects on test scores, without controls

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
French	0.022	-0.076	0.120	-0.118	0.143	0.699	744
Mathematics	0.023	0.016	0.132	0.277**	0.134	0.021**	735

Notes. This table reports coefficients from the same regressions as those presented in Panel B of Table 6, without statistical controls. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 18: ITT effects on test scores, clustering standard errors at the class level

	Control mean	ITT after 1 year	SE	ITT after 2 years	SE	ITT 1 = 2	N
French	0.022	-0.060	0.117	-0.112	0.137	0.761	744
Mathematics	0.023	-0.014	0.098	0.251**	0.101	0.058*	735

Notes. This table reports coefficients from the same regressions as those presented in Panel B of Table 6, clustering standard errors at the class level. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 19: Ressources allocated to the school, after 1 and 2 years

	$E(Y_0\|C)$	LATE year 1	SE	LATE year 2	SE	LATE 1 = 2	N
Class size	24.987	-7.476***	0.864	-5.953***	1.059	0.035**	396

Notes. This table reports coefficients from a 2SLS regression of class size on a dummy for year 1, the interaction of this dummy with the number of years spent in the school after one year (column 3), a dummy for year 2, the interaction of this dummy with the number of years spent in the school after two years (column 5), and the statistical controls listed in Section 2.2, using our lottery offer interacted with the year 1 and year 2 dummies as instruments. Our estimation sample is the second cohort of students, as class size is not available one year after the lottery for the first cohort. The second column of this panel reports an estimate of the mean of French and maths test scores for compliers not enrolled in the school. We use propensity score reweighting to control for lottery strata. Standard errors reported in columns 4 and 6 are clustered at the class level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. Measures of class size come from students' questionnaires. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 20: Students' experience in the classroom, after 1 and 2 years

	$E(Y_0\|C)$	LATE year 1	SE	LATE year 2	SE	LATE 1 = 2	N
Attendance over the last two weeks							
Attendance score	-0.067	0.200	0.282	-0.179	0.342	0.315	398
Missed school	0.106	-0.194	0.285	0.272	0.392	0.259	400
Skipped classes	0.123	-0.284	0.266	0.133	0.352	0.307	398
Arrived late	0.036	-0.100	0.245	-0.142	0.380	0.918	400
Disruption							
Disruption score	0.011	-0.634**	0.265	-1.067***	0.337	0.287	399
Teacher often waits students calm down	-0.064	-0.466*	0.252	-0.622*	0.335	0.700	400
Students start working long after class begins	0.148	-0.430*	0.222	-0.598**	0.280	0.627	400
Students cannot work well	-0.010	-0.531**	0.232	-0.464	0.315	0.860	399
There is noise and disruption in the classroom	0.093	-0.459**	0.226	-0.851***	0.324	0.308	400
Students do not listen to the teacher	0.090	-0.670***	0.246	-1.208***	0.434	0.264	400
Relationships between students							
Students relationships score	0.046	0.696**	0.279	0.492	0.311	0.571	366
Students are ashamed when they have good grades	-0.137	0.222	0.260	-0.091	0.352	0.453	367
Weak students make fun of strong ones	0.326	-0.629**	0.250	0.488	0.343	0.008***	400
Students do their homework in group	-0.048	0.673**	0.266	0.284	0.386	0.347	400
Strong students help weak ones	0.170	0.837***	0.271	1.179***	0.325	0.301	399
Teachers' engagement							
Teachers' engagement score	-0.322	1.424***	0.314	1.311***	0.424	0.828	400
She cares for students progress	-0.174	0.749***	0.218	0.439	0.307	0.400	400
She explains until students understand	-0.331	1.199***	0.250	1.319***	0.373	0.783	400
She listens to students opinions	-0.261	0.688***	0.225	0.577*	0.335	0.780	400
Teacher-students relationships							
Teacher-students relationships score	-0.065	0.703***	0.231	0.731**	0.365	0.947	365
Students get along well with their teachers	-0.006	0.489**	0.210	0.657**	0.313	0.648	400
Teachers care for students	-0.065	0.526**	0.235	0.494	0.334	0.937	367
Teachers listen to students	-0.033	0.321	0.234	0.320	0.370	0.999	398
Teachers give supplementary help if needed	0.052	0.393*	0.227	0.416	0.382	0.956	398
Teachers are fair to students	0.055	0.356*	0.215	0.811*	0.415	0.311	398

Notes. This table reports coefficients from 2SLS regressions of the outcomes listed in the first column on a dummy for year 1, the interaction of this dummy with the number of years spent in the school after one year (column 3), a dummy for year 2, the interaction of this dummy with the number of years spent in the school after two years (column 5), and the statistical controls listed in Section 2.2, using our lottery offer interacted with the year 1 and year 2 dummies as instruments. Our estimation sample is the second cohort of students, as the outcomes studied here are not available one year after the lottery for the first cohort. For other outcomes, we use both cohorts of students. The second column of this panel reports an estimate of the mean of French and maths test scores for compliers not enrolled in the school. We use propensity score reweighting to control for lottery strata. Standard errors reported in columns 4 and 6 are clustered at the class level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. All variables come from students' questionnaires. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 21: Students' experience outside the classroom, after 1 and 2 years

	$E(Y_0 \mid C)$	LATE year 1	SE	LATE year 2	SE	LATE 1 = 2	N
Students' schedule after the school day							
Hours spent last week in study room	2.672	2.557***	0.517	4.403***	0.945	0.104	717
Hours spent last Monday playing video games	0.452	-0.369**	0.173	-0.251	0.198	0.562	714
Hours spent last Monday watching TV	1.459	-1.119***	0.176	-1.188***	0.263	0.803	722
Supervisor-students relationships							
Supervisor-students relationships score	-0.128	0.061	0.283	-0.356	0.386	0.307	364
Students get along well with their supervisors	-0.060	-0.323	0.222	-0.689*	0.369	0.359	365
Supervisors care for students	-0.012	0.188	0.273	0.064	0.366	0.785	398
Supervisors listen to students	-0.022	0.239	0.271	0.076	0.362	0.680	397
Supervisors give supplementary help if needed	-0.189	0.518*	0.267	-0.286	0.394	0.075*	397
Supervisors are fair to students	-0.180	-0.082	0.276	-0.800**	0.357	0.034**	398

Notes. This table reports coefficients from 2SLS regressions of the outcomes listed in the first column on a dummy for year 1, the interaction of this dummy with the number of years spent in the school after one year (column 3), a dummy for year 2, the interaction of this dummy with the number of years spent in the school after two years (column 5), and the statistical controls listed in Section 2.2, using our lottery offer interacted with the year 1 and year 2 dummies as instruments. For some outcomes, our estimation sample is the second cohort of students, as these outcomes are not available one year after the lottery for the first cohort. For other outcomes, we use both cohorts of students. The second column of this panel reports an estimate of the mean of French and maths test scores for compliers not enrolled in the school. We use propensity score reweighting to control for lottery strata. Standard errors reported in columns 4 and 6 are clustered at the class level. In column 7, we report the p-value of a test of equality of the coefficients in the third and fifth columns. All variables come from students' questionnaires. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 22: Students' opinion on teachers: heterogeneous effects according to maths baseline score.

	$E(Y_0\|C)$	LATE	SE	N
Teachers engagement score				
In upper tercile at baseline	-0.163	0.854***	0.317	129
Out of upper tercile at baseline	-0.209	1.314***	0.275	232
Teachers-students relationships score				
In upper tercile at baseline	-0.012	0.666**	0.284	123
Out of upper tercile at baseline	0.018	0.914***	0.221	223

Notes. The first line of the table reports coefficients from the same regression as that in Table 4 for teachers' engagement score, within the sample of students who took at least one maths test and who where in the first tercile of maths scores in their lottery stratum at baseline. The second line reports the same coefficients from the same regression, within the sample of students who took at least one cognitive test and who where not in the first tercile of maths scores in their lottery stratum at baseline. Accordingly, the following lines of the table reproduce results for teachers-students relationships score shown in Table 4, separately for students in and out of the first tercile of maths scores at baseline. We use propensity score reweighting to control for lottery strata. Standard errors reported in column 3 are clustered at the class level. All variables come from students' questionnaires. *significant at 10%; **significant at 5%; ***significant at 1%.

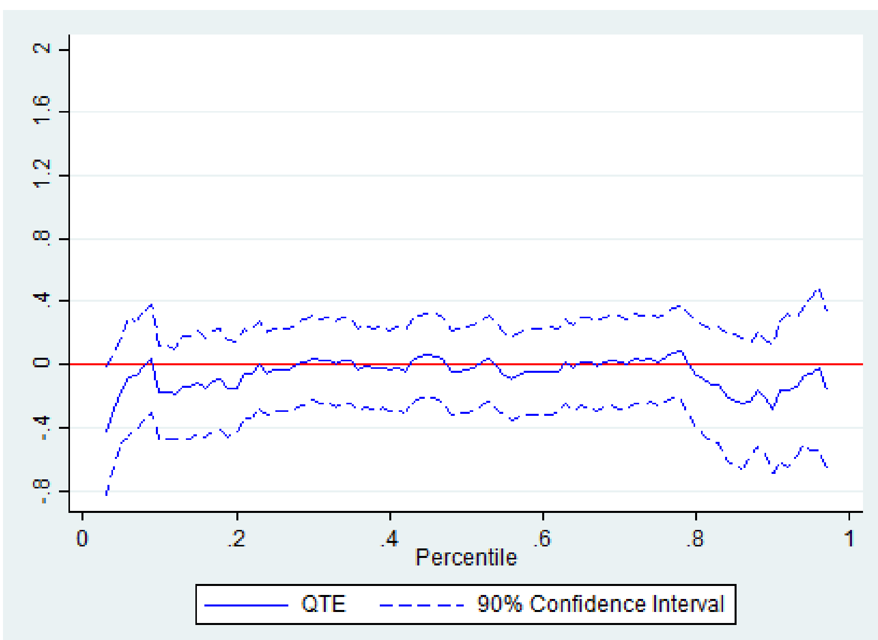

Figure 4: QTE in French after 1 year, intention-to-treat.

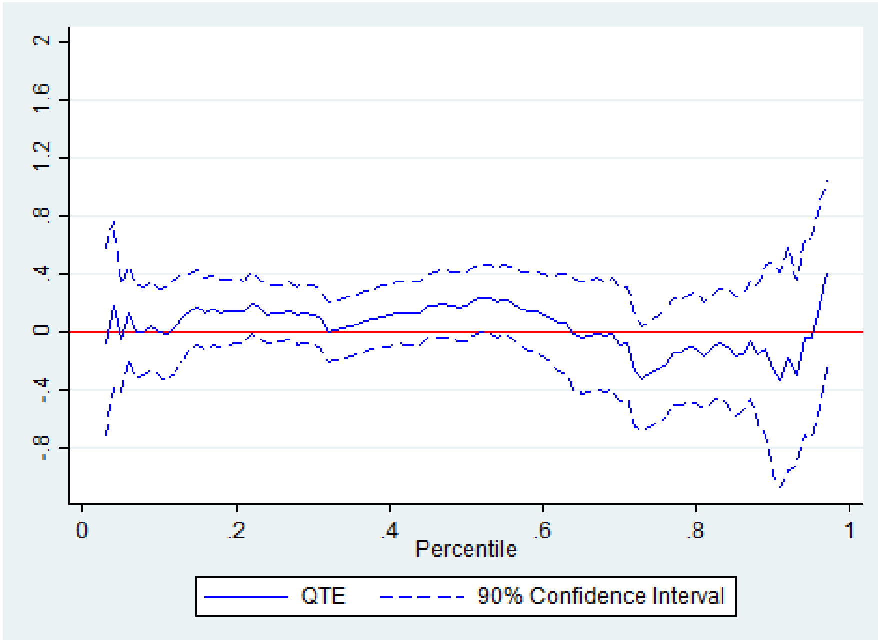

Figure 5: QTE in Mathematics after 1 year, intention-to-treat.

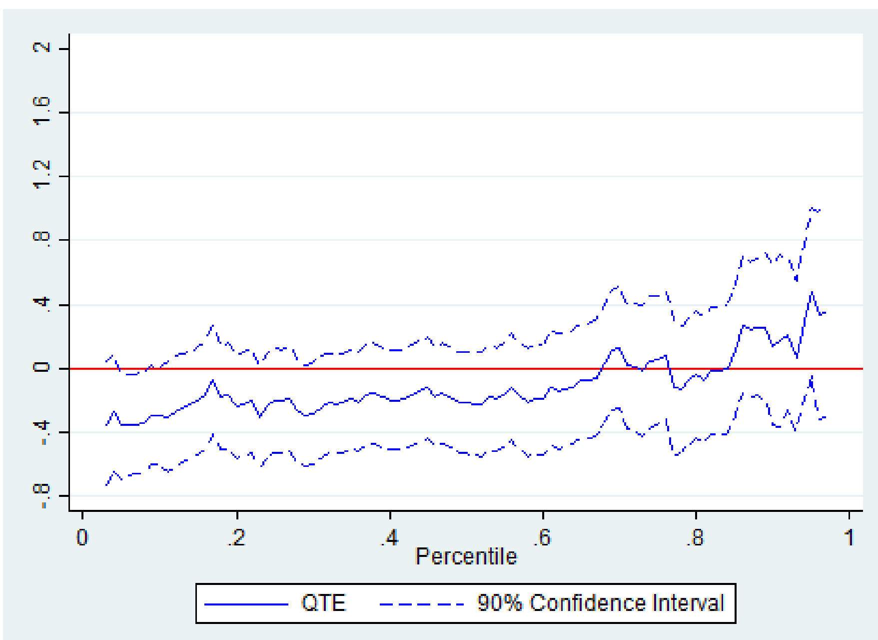

Figure 6: QTE in French after 2 years, intention-to-treat.