

# Goals and Gaps: Educational Careers of Immigrant Children \*

Michela Carlana <sup>†</sup>, Eliana La Ferrara <sup>‡</sup>, Paolo Pinotti <sup>§</sup>

July 2020

## Abstract

We study the educational choices of children of immigrants in a tracked school system. We first show that immigrants in Italy enroll disproportionately into vocational high schools, as opposed to technical and academically-oriented ones, compared to natives of similar ability. The gap is greater for male students and it mirrors an analogous differential in grade retention. We then estimate the impact of a large-scale, randomized intervention providing tutoring and career counseling to high-ability immigrant students. Male treated students increase their probability of enrolling into the high track to the same level of natives, also closing the gap in grade retention. There are no significant effects on immigrant girls, who exhibit similar choices and performance as native ones in absence of the intervention. Increases in academic motivation and changes in teachers' recommendation regarding high school choice explain a sizable portion of the effect. Finally, we find positive spillovers on immigrant classmates of treated students, while there is no effect on native classmates.

**Keywords:** tracking, career choice, immigrants, aspirations, mentoring

---

\*We thank Guido Imbens, three anonymous referees, Josh Angrist, David Autor, Erich Battistin, Pietro Biroli, Daniele Checchi, Kaivan Munshi, Imran Rasul, Enrico Rettore, Miguel Urquiola and seminar participants at several universities for helpful comments. We are particularly grateful to Gianpaolo Barbetta, Paolo Canino, Stefano Cima, and Andrea Trisoglio for continuous support during the project. We thank Gianna Barbieri and Lucia De Fabrizio from MIUR and Patrizia Falzetti and Michele Cardone from INVALSI for giving us access to the administrative data used in this paper. We are grateful to the schools that took part into the intervention for their collaboration. Serena Cocciolo, Rosa De Vivo and Cristina Clerici provided excellent research assistance. This project was partly funded by Fondazione CARIPLO, Compagnia di San Paolo and Fondazione Cassa di Risparmio di Padova e Rovigo. Carlana acknowledges financial support from the "Policy Design and Evaluation Research in Developing Countries" Initial Training Network (PODER), which is financed under the Marie Curie Actions of the EU's Seventh Framework Programme (Contract no. 608109). La Ferrara acknowledges financial support from the ERC Advanced Grant "Aspirations, Social Norms and Development" (ASNODEV, Contract no. 694882). AER RCT Registry ID: AEARCTR-0002148.

<sup>†</sup>Harvard Kennedy School and IZA (e-mail: [michela\\_carlana@hks.harvard.edu](mailto:michela_carlana@hks.harvard.edu)).

<sup>‡</sup>Department of Economics, IGIER and LEAP, Bocconi University, and CEPR (e-mail: [eliana.laferrara@unibocconi.it](mailto:eliana.laferrara@unibocconi.it)).

<sup>§</sup>Department of Social and Political Sciences and DONDENA, Bocconi University, CReAM Centre, fRDB, IR-VAPP, and CEPR (e-mail: [paolo.pinotti@unibocconi.it](mailto:paolo.pinotti@unibocconi.it)).

# 1 Introduction

Migrant flows have grown considerably over the past decades, increasingly involving families with children. In 2015, almost one in four 15-year-old students in OECD countries had an immigrant background (OECD, 2018). This has profoundly changed the challenges that schooling systems face in order to ensure skill development in a diverse student population and promote social cohesion. The ethnic gap in achievement test scores and socio-emotional abilities increases substantially during childhood (Fryer Jr and Levitt, 2004; Cunha and Heckman, 2007). The problem is exacerbated in schooling systems characterized by stratification in high school tracks, as early tracking may lead to the educational segregation of children from disadvantaged backgrounds in schools of lower quality. This could ultimately have long term effects on the skills and occupational careers of children from immigrant families, reducing social mobility and creating unequal opportunities (Guyon et al., 2012; Brunello and Checchi, 2007).

This paper documents the extent of educational segregation of immigrant students and evaluates the effectiveness of an innovative program aimed at steering high-achieving immigrants towards high schools that fit their academic potential. Throughout the paper, we use the term “immigrant” to indicate children without Italian citizenship: around half of them are first generation immigrants (i.e., they are born outside Italy) and the other half are second generation immigrants (i.e., they are born in Italy from non-native parents). We perform our analysis in the context of Italy, where the schooling system is characterized by tracking in the transition from middle to high school. While uncommon in the Anglo-Saxon world, this type of stratification is the norm among OECD countries, with the age of selection varying from 10 to 16 and an average of three high school tracks per country (OECD, 2013).

We start by documenting that immigrants are more likely to enroll in vocational over technical or academic-oriented curricula, relative to native students with similar ability, as measured by a standardized test administered at the beginning of middle school. We denote this phenomenon as “educational segregation”. The gap in track choice for boys persists along the entire distribution of ability, while for girls it is found at the low end of the distribution but not at the high end. The presence of a bigger gap for boys, on average, is consistent with evidence from various countries that boys increasingly lag behind in educational attainment and that the female-male educational advantage is larger for low-socioeconomic status (SES) families (Chetty et al., 2016; Autor et al., 2019; Bertrand and Pan, 2013).

We then estimate the impact of an innovative program called “Equality of Opportunity for Immigrant Students” (EOP henceforth) that provided tutoring and career counseling to immigrant children displaying high academic potential. The curriculum of EOP included a number of meetings

that helped students reflect on their aspirations and their potential through a series of psychological exercises based on Social Cognitive Career Theory (Lent et al., 1994).

We evaluate the effects of EOP leveraging random assignment of the program across schools as well as unique data on students' careers, cognitive and soft skills, and parental background. The program was offered in 70 middle schools (grades 6 to 8) randomly chosen from a sample of 145 in Northern Italy. We selected the 10 immigrant students with the highest standardized test scores in grade 6: in treatment schools these students were invited to participate in EOP through grades 7 and 8; in control schools they were not offered such program.<sup>1</sup>

We find that EOP was remarkably successful in reducing educational segregation. Treated males have a 44 percent lower probability to be retained and a 12 percent higher probability of attending an academic or technical high school (as opposed to a vocational one) relative to males in the control group. Indeed, treated immigrant males chose demanding tracks in the same proportion as native males of comparable ability. In other words, EOP completely closed the immigrant-native track choice gap at the end of grade 8. The effects are in the same direction but smaller and not significant for girls (for whom no educational segregation had been identified in the first place). Although multiple factors could explain the difference in impact across genders, it is suggestive that EOP increases immigrants' enrollment into technical or academic schools only when counterfactual enrollment rates lie below those of comparable native students.

To shed light on mechanisms, we collected first-hand data on psychological traits. We find that male treated students have higher aspirations and more confidence in their own abilities, and they perceive that environmental barriers will play a smaller role in their future choices, relative to the control group. They also display an improvement in cognitive skills, as measured by their standardized test score at the end of grade 8. For girls we detect a reduction in perceived barriers but no relevant effect on aspirations or cognitive skills.

Program effects seem to have been internalized by teachers, who recommend treated boys for more demanding high schools. Importantly, no statistically significant effect is found on teachers' recommendations for treated girls, suggesting that teachers do not automatically recommend the high track to all students involved in EOP, but selectively update their beliefs based on actual changes in motivation and performance – greater for males, lower for females.

We estimate heterogeneous treatment effects using a causal forest following [Athey and Imbens \(2016\)](#) and [Wager and Athey \(2018\)](#) to understand who benefits most from the treatment. We find that, although on average EOP did not have a statistically significant impact on girls' track choice,

---

<sup>1</sup>For the sake of exposition, in the rest of the paper we refer to these two groups as “treated” and “control”, though we really mean “assigned to treatment” and “assigned to control”. In other words, we present intention-to-treat estimates. We assess compliance with treatment assignment in Appendix B.

all male and female students in the middle of the initial test score distribution and those from a low socio-economic background benefited from the intervention. This pattern of heterogeneity could be used to target future scaled-up interventions.

Following the students in our sample through the first two years of high school, we assess the longer term effects of EOP. We find that treated students did equally well as control ones, despite attending (on average) more demanding high schools. Finally, we find evidence of positive spillovers of the intervention on immigrant classmates of treated students, while there is no effect on native classmates.

Our work is related to several strands of literature. The first comprises evaluations of interventions aimed at reducing inequality in educational achievement and opportunities. Several interventions have targeted low achieving students and provided a combination of information on school options and mentorship on soft skills. Some of these programs were successful in reducing grade retention and high school dropout rates (e.g., Goux et al., 2017; Martins, 2010); others had zero or negative effects (Rodriguez-Planas, 2012). Our program can be seen as complementary to the ones cited above, as it targets a different population – high achieving students – with the aim of aligning their potential to more ambitious goals.<sup>2</sup>

A second strand of literature looks at the role of soft skills (e.g., Heckman et al., 2006; Heckman and Kautz, 2012). Low aspirations and high perceived socioeconomic barriers can lead students to choose less demanding educational paths, perpetuating a negative cycle (Dalton et al., 2016; Genicot and Ray, 2017; Mookherjee et al., 2010). Recent contributions show that students from disadvantaged backgrounds often lack ambition and suffer from negative stereotypes (e.g., Hoxby and Avery, 2013; Guyon and Huillery, 2016). Our paper shows that it is possible to modify aspirations and soft skills through a program that combines academically relevant information with psychological tools.

Finally, our work speaks to the literature on tracking within education systems.<sup>3</sup> For example, some authors find that postponing school stratification or increasing the proportion of seats in aca-

---

<sup>2</sup>Card and Giuliano (2016a) study an intervention for high-achieving students in 4th grade that has a significant positive effect on minority students but not on whites. Their treatment closes an achievement gap with whites of similar baseline ability and makes minority students more likely to participate in an accelerated math track in middle school. Different types of interventions have targeted students' inaccurate beliefs about their chances of succeeding in competitive academic environments. Bobba and Frisncho (2016) provided students with feedback on their performance in a (mock) exam for admission to selective high schools in Mexico, while Goodman (2016) estimates the effect of mandated exams on selective college admission in the US.

<sup>3</sup>For a review, see Betts (2011). Differently from systems where tracking takes the form of sorting higher ability students within the same type of education into specialized instruction (e.g., gifted programs, Card and Giuliano, 2016a) or into magnet schools (e.g., Pop-Eleches and Urquiola, 2013), our context is one where tracking involves sorting into high schools with very different curricula and differential access to college. Brunello and Checchi (2007) show that parental background has stronger effects on labor market outcomes when tracking starts earlier.

demographic tracks leads to improved educational outcomes (Malamud and Pop-Eleches, 2011; Guyon et al., 2012). Our goal is not to assess what would be the effects of postponing or modifying high school tracking, given that tracking is held constant in our experiment. However, one interpretation for the gap in choice that we see between immigrant boys and native students with similar performance is that there is a mismatch. Under that interpretation, our findings suggest that – within an existing tracking system – it is possible to align ability and high school choice for a population that may be particularly misinformed and disadvantaged: immigrant students.

The remainder of the paper is organized as follows. Section 2 describes the institutional setting and provides evidence of the educational segregation of immigrant children in Italy. Section 3 illustrates the components of the intervention and our evaluation design. Section 4 describes the data and Section 5 the results. Finally, Section 6 concludes by discussing the policy implications of our findings.

## 2 Institutional background

### 2.1 Immigrants in Italian schools

Immigration is a relatively recent phenomenon in Italy. The number of (legal) foreign residents increased from 781,000 to 5 million between 1990 and 2015 – 1.4 and 8.3 percent of total residents, respectively. The majority of immigrants in Italy come from low and middle income countries, and are characterized on average by lower socioeconomic background than native households.<sup>4</sup>

At the beginning of the 2016/17 school year, immigrant children represented 10.8 percent of students in primary school, 9.7 percent in middle school, and 7.2 percent in high school (see Appendix Figure A.3). The decline in immigrants' presence from middle to high school reflects higher dropout rates of immigrants in later grades compared to natives. Importantly, the share of immigrant students also differs between different types of high schools. Immigrants represent 12.5 percent of the student population in vocational schools, 8.5 percent in technical schools, and only 4.1 percent in academic schools. As we detail in the next section, these three types of high schools offer very different educational and employment opportunities.

---

<sup>4</sup>Figure A.1 in the Appendix shows the number of immigrants by nationality in 2015 (first 20 nationalities). Figure A.2 compares the income distribution across immigrant and native families, respectively.

## 2.2 Secondary education in Italy

Italian pupils normally enter formal schooling the year they turn 6 and the compulsory schooling age is 16. Pre-university education comprises five grades in elementary school, three grades in middle school, and five grades in high school. Within each type of school (elementary, middle and high school), pupils take all their subjects within the same class and with the same set of peers. In elementary and middle school, they typically also retain the same teacher for a given subject throughout the relevant education cycle (e.g., a pupil’s maths and science teacher will be the same throughout middle school). At the end of middle school, students must choose among three different types of high schools: vocational schools (*istituto professionale* and *formazione professionale*), technical schools (*istituto tecnico*), and academically-oriented schools (*liceo*). Students are free to enroll in any track, as there is no tracking by ability.

The three tracks have the same duration, 5 years, but differ widely in terms of curriculum, difficulty, and prestige.<sup>5</sup> Vocational schools focus on practical training in specialized manual, low-skilled jobs (e.g., plumber or hairdresser), and are meant to prepare students for immediate employment at the end of high school. Technical and academic schools offer instead a comprehensive curriculum in math, humanities, and science. In principle, academic schools are primarily intended for students who want to pursue a university degree, whereas technical schools complement theory with practical training in specific non-manual jobs (e.g., accountant or graphic designer). Enrollment in college is always possible from the academic-oriented and the technical tracks; in principle it is also possible from some schools within the vocational track, although very few students who attended vocational education typically obtain further education.<sup>6</sup> In practice, academic and technical schools offer much better educational and employment prospects than vocational schools. Therefore, in what follows we refer to vocational schools as the “low-track”, and we group technical and academic schools together into the “high track”. We will also run some robustness checks on the three tracks separately.

Appendix Table A.I compares average outcomes by track four years after graduation, separately by gender and for native vs. immigrant students.<sup>7</sup> Panel A shows that only 14.5 percent of Italians graduated from vocational schools, with no relevant differences by gender. They exhibit a much

---

<sup>5</sup>Regarding duration, the only exception to the 5-year rule is a sub-track of the vocational track (*formazione professionale*) that lasts 3 years.

<sup>6</sup>Enrollment in college is not possible for students in the sub-track of the vocational track that lasts 3 years (*formazione professionale*). In the 2015-16 school year, less than 4 percent of students enrolled in college had attended a vocational track.

<sup>7</sup>The source of these data is the “Survey on Educational and Professional Paths of Upper Secondary School Graduates”, conducted in 2015 by the Italian National Statistical Institute (ISTAT) on a representative sample of about 26,000 students graduating from high school in 2011.

lower probability of pursuing tertiary education compared to high-track graduates: 20.5 percent, as opposed to 70.4 percent. College dropout rates in university also differ dramatically between the two groups, at 30.6 and 11.8 percent, respectively. In light of these figures, employment rates and salaries four years after graduation are not really informative about labor market prospects, as only a selected group of high-track graduates has already entered the labor market. However, we can compare the share of those “Not in Education, Employment or Training”(NEET), which reaches 29 percent among low-track graduates, 10 percentage points higher than among other graduates. This is particularly surprising, as in principle vocational schools should prepare students for immediate employment. Finally, graduates from vocational schools also have a higher probability – about one third – of regretting their choice, as elicited in the ISTAT survey. These figures are fairly similar by gender.

Panel B of Appendix Table [A.I](#) shows comparable statistics for immigrant students. Conditional on completing the same high school track, educational and occupational outcomes are remarkably similar to those experienced by natives. Also among immigrants, graduates from vocational schools exhibit lower enrollment into (and higher dropout rates from) tertiary education, as well as a higher prevalence of NEETs. A stark difference emerges, however, when we compare high school choice: 37 percent of immigrants (42 percent among males) graduate from the low track, compared to the aforementioned 14.5 (15.6 for males) of natives. The fact that the share of NEETs is not lower among immigrants graduating from the low track helps dispel a commonly held belief, i.e., that vocational tracks may be better for immigrants who may be seeking immediate employment. In light of the (worse) outcomes experienced on average by low-track graduates, the over-representation of immigrants in this group raises concerns about immigrants’ future career opportunities and, eventually, their prospects for successful integration and upward social mobility. The spirit of the intervention we evaluate was to improve the future of talented immigrant students, by making them more aware of a career path that –based on the above data– on average leads to better subjective and objective outcomes.

Of course, enrollment rates and outcomes in Appendix Table [A.I](#) reflect endogenous sorting by ability and socioeconomic background. Below we provide a more informative comparison of transitions to high school between immigrants and natives, exploiting a unique dataset that matches administrative data on educational careers with standardized test scores and information on parental background.



## 2.3 Educational segregation

As we discuss in detail in Section 4.1, Italian students take a series of standardized tests of proficiency in reading and math at various points of their careers. These tests are known as INVALSI, from the name of the agency that administers them. The tests are identical for all students in a given grade and are blindly scored, so results are fully comparable across schools. Throughout the paper, we use the standardized test score obtained in grade 6 (INVALSI6) as a measure of students' ability at the beginning of middle school.

Using a unique dataset that matches the above scores with students' educational careers, we can compare the average probability of enrolling into the high track for native and immigrant students conditioning on their initial ability.<sup>8</sup> Figure 1 plots the probability of enrolling in a “high track” (academic or technical) by quintile of INVALSI6, separately for male and female students (left and right panel, respectively).<sup>9</sup>

[Insert Figure 1]

Not surprisingly, the probability of choosing the high track is increasing in INVALSI6 for all groups. However, such probability remains significantly lower for immigrant males than for native ones. The gap is larger in the upper part of the ability distribution, reaching 16 percentage points in the top quintile. By contrast, the gap between immigrant and native females is much smaller and it is negligible in the upper part of the ability distribution. Appendix Table A.II shows that the gap in educational choices between native and immigrant males persists when conditioning in addition on family background, as measured by parents' education, employment status, and occupation. This suggests that migration status plays a role, on top of general socioeconomic disadvantage.

Why is the native-immigrant gap larger for boys than for girls? The gender gap we uncover is consistent with a recent literature that has documented significant gender differences in the effect of family background on child development and outcomes during adulthood. Unlike girls, boys' development and behavior appears to be extremely responsive to worse parental inputs. Low-income boys have lower performance in standardized test scores, lower high-school graduation rates, and more disruptive behavior than their sisters – or girls who grew up in similar contexts (see among others, Autor et al., 2019; Almås et al., 2016; Bertrand and Pan, 2013; DiPrete and Jennings, 2012; Conger and Long, 2010; Jackson and Moore III, 2008). These effects translate into long-term

---

<sup>8</sup>As we explain in Section 4, this dataset was constructed for the evaluation of the EOP program, which took place in five large cities of Northern Italy. The data used in Figure 1 are those from schools in the control group, which were unaffected by our intervention.

<sup>9</sup>Appendix Figure A.4 reports corresponding graphs for other outcomes that we consider in our analysis, namely: grade retention, teachers' recommendations and test scores at the end of grade 8 (INVALSI8).



negative consequences: minority boys are less likely to attend university, work disproportionately less and have higher crime rates than girls (Chetty et al., 2018, 2016; Brenøe and Lundberg, 2019). An additional reason behind the gender gap in favor of girls among immigrant children may be related to gender roles and norms imposing social control of daughters and more lax regulations for sons (Lopez, 2003), which may lead to better outcomes for high-achieving immigrant girls compared to boys.

Although the focus of our paper is not on the gender gap among minority students – but, rather, on the gap between natives and immigrants – the fact that this gap differs between boys and girls will help us to understand the differential impact of our intervention across genders.

### 3 The intervention

The intervention we evaluate was developed in collaboration with the Italian Ministry of Education and three bank foundations.<sup>10</sup> The program was called “Equality of Opportunity for Immigrant Students” (EOP) and aimed at aligning the goals and aspirations of high-achieving immigrant students with their ability, in order to favor congruous educational choices at the end of middle school. The intervention took place during the last two years of middle school (grades 7 and 8) and was administered in a randomized fashion in five large cities of Northern Italy: Milan, Turin, Genoa, Brescia, and Padua.

The first dimension of targeting involved the definition of the school sample: schools were eligible to receive the program if they had at least 20 immigrant students, where ‘immigrant’ was defined as being a citizen of a country with lower GDP than Italy.<sup>11</sup> In the five cities there were 145 such schools: 70 were randomized into the treatment group and 75 into control. To enhance comparability between the two groups of schools we stratified randomization by province and school size.

The second step was the definition of the target students. Because the goal of EOP was to reduce mismatch in track choice for high-achieving immigrants, within each school we defined as ‘high-achievers’ the 10 immigrant students with the highest standardized test score in grade 6

---

<sup>10</sup>The program was financed by three philanthropic institutions operating in Northern Italy, namely Fondazione CARIPLO, Compagnia di San Paolo, and Fondazione Cassa di Risparmio di Padova e Rovigo.

<sup>11</sup>This excludes from the program immigrant children from high income European countries, for whom no educational segregation exists. When we use the term “immigrant” we include both first generation immigrants (i.e., children born outside Italy and without Italian citizenship) and second generation immigrants (i.e., children born in Italy from parents without Italian citizenship). In Italy, citizenship is mainly acquired through *ius sanguinis*, i.e., if an individual’s parents have Italian citizenship. It is possible to obtain the Italian citizenship when first generation immigrants turn 18 years old.

(INVALSI6). In the treatment schools these 10 students took part in EOP while in the control schools they did not. In both sets of schools, these top 10 immigrant students were surveyed and their academic performance and school choices were followed through administrative records. In our empirical analysis, we will thus compare outcomes between the 10 immigrant students with the highest INVALSI6 scores in treated and control schools.

The EOP program consisted in two components: (i) a psychologically grounded career choice consultancy, and (ii) an academic tutoring one. The career choice consultancy was developed based on Social Cognitive Career Theory, a paradigm that views career development as a choice subject to contextual influences and constraints (Lent et al., 1994). Under this view, goals and self-efficacy are as important as cognitive skills in shaping individual careers.<sup>12</sup> Specifically, “persons with adequate skills but weak self-efficacy beliefs in a particular performance domain may prematurely rule out that domain from further occupational or academic choice consideration” (Brown, 2002). The goal of EOP was to help high-achieving immigrant students to identify educational and occupational goals congruous with their talents. The approach was not to unconditionally push students towards high tracks, but to make them aware of existing opportunities and of their own skills, so they could make more informed choices.

The protocol of the career choice consultancy involved a total of 13 meetings and the program guidelines required participants to attend at least 75 percent of the meetings. All meetings were administered by career counselors with graduate degrees in psychology and significant experience in career choice guidance for secondary school, especially with immigrant children. Some of the meetings were one-to-one, while others were in groups.

Students had five (one-to-one) meetings with a counselor and worked on tasks that prompted them to reflect on their goals, the personal resources needed to achieve such goals, and whether they already had or they needed to develop such resources. Examples of the tasks that students worked on include: (i) “Think about your past life, indicate five study experiences and five other experiences that you have completed successfully. Consider now such experiences one by one and briefly indicate where and with whom it happened, what you did and which personal resources helped you doing well in that thing - your knowledge, skills, personality traits, motivations and everything you believe it was important to have”; (ii) “Choose a number of professions that interest you. For each of them, indicate which resources are needed (knowledge, skills, personality traits, motivations, ...) and divide them into ‘I have it’ and ‘I need to develop it’;” (iii) “List the results you would like to achieve with your job, from the most to the least important.”

Five other meetings were held in groups, where counselors provided students with information

---

<sup>12</sup>Self-efficacy can be defined as people’s beliefs about their capabilities “to organize and execute courses of action required to attain designated types of performances” see Bandura (1986), p. 391.

about the Italian education system and showed videos with success stories of older immigrant students.

Two further meetings, respectively at the beginning and at the end of the intervention, were intended for parents. In the first meeting counselors described the content of the program, while in the second they shared with parents aspirations and barriers perceived by students. Parents also received a brochure, translated into their mother tongue, summarizing the main options for secondary education in Italy.<sup>13</sup> Finally, towards the end of the intervention career counselors met with teachers and discussed the educational path and high school track chosen by the students involved in the intervention.

[Insert Figure 2]

Figure 2 shows the timeline of the intervention and the realization of the main outcomes of interest. In grade 6 we selected the eligible students on the basis of the standardized test score they got that year. EOP meetings started at the beginning of grade 7 and continued through grade 8, until the month of March. By the month of January of grade 8, students receive a formal ‘recommendation’ by their teachers about the high school track that teachers deem most appropriate for them. This recommendation is not binding but it serves as a signal to the students and their families. In February, all students have to pre-register for the high school they wish to attend through a web portal of the Ministry of Education. This choice can later be modified (though this is not very common), so we use the high school track in which students actually enroll at the end of grade 8 as our variable of interest.

The second component of the intervention, the academic tutoring part, was motivated by the concern that participants who enrolled in the high track as a result of the program may subsequently experience difficulties in completing this (more demanding) high school. For this reason, the counseling and career choice module was accompanied by a module on Cognitive Academic Language Proficiency (CALP). CALP was not aimed at improving students’ knowledge or their cognitive skills but, rather, at teaching them a method for studying several subjects – Italian grammar, geography, algebra, and geometry. Since the main motivation of the CALP module was to decrease the risk of subsequent failure for immigrant students enrolling in demanding tracks, students with a lower INVALSI6 were offered a higher number of CALP meetings. Specifically, students scoring below 65 out of 100 in INVALSI6 were invited to 29 meetings (55 hours tutoring). This group constituted 66 percent of all treated students. Students scoring between 65 and 80 (30

---

<sup>13</sup>While it would be very interesting to disentangle the role that parents play in high school choice (e.g., Dustmann, 2004; Giustinelli, 2016), it was not possible to survey parents within our experiment.

percent of the treatment group) were invited to 17 meetings (32 hours tutoring); finally, students scoring above 80 (4 percent of the treatment group) were not invited to CALP sessions.

Due to ethical concerns we could not implement a fully factorial design, as our sponsors were concerned that encouraging students to pursue ambitious goals without endowing them with the tools for succeeding in demanding high schools could have created a risk of harm. However, we can assess the relative effectiveness of career counseling and CALP at the intensive margin by exploiting the different cutoffs for the number of CALP meetings and employing a regression discontinuity design (RDD). We do this in Section 5.3.

## 4 Data

We build a unique dataset including survey data and administrative information from the Italian Ministry of Education (MIUR) and the Institute for the Evaluation of the Italian Schooling System (INVALSI), an independent public agency that monitors students' performance.

### 4.1 School choice and academic performance

**Administrative data from school registry.** From the MIUR administrative registry we take the following variables that we use as outcomes for each student: (i) track choice at the end of grade 8; (ii) track recommended by teachers half-way through grade 8; (iii) grade retention for all grades between 6 and 9 (included); (iv) number of retakes after the summer of grade 9;<sup>14</sup> and (v) teachers' assessment of student's behavioral conduct during grades 8 and 9. The registry also contains some information on students' background, in particular: citizenship, country of origin, date of birth and school and class attended throughout their careers.

**INVALSI tests.** Since 2010, INVALSI administers standardized reading and math proficiency tests to all students at the end of grades 2, 5, 6, 8, and 10. Such tests resemble those administered by the OECD Programme for International Student Assessment (PISA) to representative samples of 15-year old students. They consist of a series of questions including multiple choice as well as open ended questions, the exact structure of the test varying by grade.<sup>15</sup> Importantly, the test is identical for all students in a given grade, it is administered to everyone on the same day (towards the end of

---

<sup>14</sup>While through middle school pupils are either admitted to the next grade or retained, in high school they can also be admitted to the next grade conditional on re-taking (and passing) after the summer an examination in one or more subjects in which they were deficient during the year.

<sup>15</sup>In general, math questions are related to calculus, geometry, probability and algebra, while reading questions are related to text comprehension and grammar.

school year), and it is blindly scored, so results are fully comparable across schools in Italy. This is crucial for the purposes of our analysis, because it allows us to compare the educational choices of immigrant and native students holding constant their academic proficiency.

We use two test scores for the cohort of students who were in grade 6 in 2012. The first is the standardized test score for grade 6 (INVALSI6), which we include in all specifications to control for students' initial ability. The second is the test score for grade 8 (INVALSI8), which is one of our outcomes.

## 4.2 Soft skills

We complement the above dataset with original survey data on soft skills collected at the end of grade 8. The goal of this survey is to allow us to better understand what mechanisms shape career-related interests and high school track choice. The survey was administered to all treated students and to a random 50 percent sample of control schools.<sup>16</sup> The questionnaire was developed with a team of psychologists and includes three main sections:

(i) *Goals*. This section comprises both educational (e.g., university degree, diploma or less) and occupational targets (e.g., blue collar, white collar, managerial or entrepreneurial jobs) that the student aims to achieve. For example, questions in this module include: "Thinking about your future, education-wise, what objectives do you intend to achieve?" (Options: Work as soon as I finish this school; Study until age 16 then work; Study until age 18 and get a high school diploma; Enroll in university); and "Thinking about your future, work-wise, what objectives do you intend to achieve?" (Options: Blue collar job; White collar job; Managerial job; Entrepreneur).

(ii) *Self-efficacy*. This section includes a student's own assessment of the extent to which he or she possesses the skills and resources required to achieve the goals stated above, as well as broad notions of self esteem. Questions in this module are worded as: "Independently of your educational aim but thinking about your abilities, do you think you could get a ... (university degree/ white collar job/ managerial job, as worded above)?" Answers in this module are on a scale of 1 to 4, with 1 meaning 'not at all', and 4 meaning 'very much'.

(iii) *Barriers*. A series of questions elicit students' perceptions of environmental barriers, be they related to economic constraints, racial prejudice, or family preferences that differ from a student's own plans. Typical questions are worded as: "Do you think the following barriers could

---

<sup>16</sup>We chose not to administer the survey to half of the control schools because we wanted to be able to test if filling in a questionnaire on goals and perceived barriers may constitute a 'treatment' in itself. In Appendix Table A.V we show that students in control schools involved and not involved in the soft skills questionnaire do not systematically differ in terms of high-school choice, grade retention, and test scores in grade 8.

be an obstacle in the achievement of your educational aims?” (Options: Economic resources; The needs and ideas of your family; Racial prejudice; Family plans related to children or marriage; Not feeling up to the standards). Answers in this module also range from 1 to 4.

Following [Thompson \(2004\)](#), we summarize the individual variables described above into interpretable aggregates using first exploratory factor analysis and then confirmatory factor analysis. This method extracts latent factors from subsets of psychological measures by maximizing (minimizing) the correlation across measures within (between) subsets. The measures associated with each factor and their respective loadings are reported in Appendix Table [A.VI](#).

### 4.3 Sample and randomization check

Our working sample at the inception of EOP comprises 1,217 students: 597 in treated schools and 620 in control ones.<sup>17</sup>

[Insert Table I]

Table [I](#) reports average characteristics of the treatment and control group at the start of our intervention. We distinguish between individual student characteristics (Panel A) and family background (Panel B). Half of the students in our sample are girls, 56 percent are first generation immigrants, and 26 percent were born before 1999 (the typical birth year of the cohort in our study). The mean of the standardized test score INVALSI6 is 60.9 in the treatment group and 60.7 in the control group (not significantly different).

[Insert Figure 3]

To get a more complete picture, Figure [3](#) plots the distribution of INVALSI6 across three groups of students in our 145 schools: native students, all immigrant students, and the 10 immigrant students with the highest score (i.e., our treated and control groups). Although immigrants generally exhibit lower schooling performance than native students, the top 10 immigrant students in each school are comparable to natives in the medium-upper part of the distribution.

INVALSI also provides information on parents’ education and occupation which, however, is missing for more than a third of our sample. In Appendix Table [A.III](#), we show that the share of

---

<sup>17</sup>By construction, the sample should have comprised 1,451 students, that is,  $10 \times (70 + 75) = 1,450$  but in one school the 10th and 11th students obtained the same INVALSI6 score and were both eligible for the program. For 70 students it was impossible to match the MIUR and INVALSI identifiers, which reduced the sample to 1,381 students. In addition, some students were retained in grade 6, moved to another school, or dropped out between the moment when they took the INVALSI6 test and the beginning of grade 7, leading to the above sample of 1,217. Appendix Table [A.VII](#) shows that missingness is not selective across treatment and control schools.

records with missing information on either or both parents' education correlates negatively with socio-economic conditions across Census tracts. This finding suggests that missing information on parents background may be a proxy for low socio-economic status. When information on education is provided, 22 (20) percent of mothers (fathers) have not completed high school, 29 (30) percent have a high school diploma, and 12 (10) percent have post-secondary education. Unemployment rates are about 6 percent for mothers and 6 percent for fathers. The share of mothers only working at home is 26 percent, while that of fathers is negligible. Among those working outside the home, 24 (36) percent of mothers (fathers) have a blue collar job, and 12 (20) percent a white collar job. Importantly, none of the student or family characteristics differ significantly between the treatment and control sample, indicating that our randomization was successful.<sup>18</sup>

## 5 Results

In this section, we estimate the impact of EOP on educational choices and grade retention, separately for males and females, and we investigate heterogenous effects using causal forest. In addition, we decompose treatment effects into several mediating factors, estimate impacts on longer term outcomes, and spillover effects on non-eligible students in treatment schools.

### 5.1 Educational choices and grade retention

In Table II we estimate the impact of EOP on high school track choice. The dependent variable is a dummy equal to 1 for students who choose the high track and 0 for those who choose the low track. The explanatory variable of interest is EOP, an indicator for whether a student attends a school that was (randomly) selected to receive our intervention. The coefficient of this dummy should thus be interpreted as the intention-to-treat (ITT) effect. Odd-numbered columns condition on treatment only, while the specifications in even-numbered columns include a squared polynomial in INVALSI6, a dummy for first generation immigrants, and province fixed effects. In all cases we cluster standard errors at the school level, the unit of randomization.

[Insert Table II]

In column 1, assignment to EOP increases the probability of choosing the high track by 5 percentage points, on a baseline rate of 75 percent in the control group. Given random assignment,

---

<sup>18</sup>Children in the treatment and control group are from 72 different nationalities, out of which the most represented countries are Romania, Philippines, Albania, Morocco, Peru, Ecuador, and China. The balance table for these characteristics is reported in the Appendix Table A.VIII.



this estimate is largely unaffected when controlling for student characteristics and province fixed effects (column 2).

This average effect, however, masks important differences by gender. EOP increases males' enrollment into the high track by 8 to 9 percentage points, up from a baseline rate of 67.4 percent (columns 3 and 4). This is a 12 to 13 percent increase over the mean. By contrast, there is no effect on female students, who start from a higher baseline enrollment rate of 82.4 percent (columns 5 and 6).

[Insert Figure 4 ]

Figure 4 compares the educational choices of students eligible for EOP, namely the ten immigrant students obtaining the highest score in INVALSI6 within each school, with the educational choices of other students down in the ranking. The left graph confirms that (eligibility to) EOP increases the probability of choosing the high track for eligible male students compared to non-eligible students near the cutoff for eligibility. The "difference-in-discontinuities" between treated and control schools in the probability of choosing the high track at the eligibility cutoff amounts to 18 percentage points (estimates are reported in Appendix Table A.IV, columns 1-3). By contrast, there is no significant discontinuity in schooling choices between eligible and non-eligible female students near the cutoff.

While the above analysis aggregates academic and technical tracks into a single outcome category, in Appendix Table A.IX we maintain the distinction among tracks and investigate the impacts of EOP using an ordered probit model.<sup>19</sup> We find that the program had an effect for boys throughout the distribution of high-school track choices, most likely by moving students from vocational to the middle track and from the middle track to the top tier schools. The coefficients for girls are very small and never significant.

[Insert Figure 5 ]

Panel A of Figure 5 compares enrollment in the high track for high-achieving immigrant students randomized into the control group (leftmost bar), treatment group (middle bar) and for a group of native students with comparable ability in the first year of middle school (rightmost bar). Specifically, we match each immigrant student in our sample with one native student of the same

---

<sup>19</sup>In the Italian schooling context, as suggested by [Carlana \(2019\)](#) and supported in our context by appendix Figure A.5, high schools can be clearly divided into three groups using the average test score of students in grade 8 who self-select into the different tracks: top-tier schools (scientific and classical academic tracks), middle-tier schools (including linguistic, artistic, pedagogical academic tracks as well as technical tracks), bottom-tier schools (vocational).

gender who obtained an identical score in INVALSI6. By construction, these three groups of students had the same standardized test score in the first year of middle school. The figure shows that two years later the immigrant boys who received EOP make similar choices as natives who started off like them, while the untreated immigrants have a significantly lower probability of choosing the high track. EOP thus prevented the type of educational segregation that we documented in section 2.3. Interestingly, Figure 5 confirms that immigrant girls make similar choices as native girls even in the absence of intervention.

Taken together, the evidence in Figure 1, Table II, and Figure 5 highlights a remarkable feature of the intervention: EOP influences educational choices only when counterfactual enrollment rates into the high track lie below those of comparable native students. Therefore, EOP seems to align immigrant students' goals and aspirations to those of native students when there is an initial misalignment, as opposed to just pushing all immigrant students towards the high track. In Section 5.4 we provide direct evidence in this respect.

[Insert Table III]

Table III and Panel B of Figure 5 convey similar evidence for grade retention in grade 7 or 8. Grade retention is surprisingly high among male immigrant students in our sample: absent the intervention, it reaches 8.5 percent, as compared to only 4.2 percent for native males with a similar INVALSI6 score. This gap disappears in EOP schools, whereas there is neither a significant gap nor an effect for female students. In both respects, the effect on grade retention across genders is very similar to that on high school choice.

Since there is one-sided non-compliance with treatment assignment, we can also estimate the average treatment-on-the-treated (ATT) effect on the subset of compliers. We do so in Appendix B. In particular, Appendix Table B.I shows that, if we classify as treated all students attending at least 75 percent of the meetings (in accordance with the program guidelines discussed in Section 3), the ATT effects on males are a 12.5 percentage point increase in enrollment in the high track, and a 5.7 percentage point decrease in grade retention.

## 5.2 Heterogeneous effects: Multiple Hypothesis Testing and Causal Forest

Given the strong gender differences in educational segregation, we focused our main analysis on the impact of EOP on track choice by gender and showed that the effect is not significantly different from zero for girls. However, estimating heterogeneous effects by splitting the sample may increase the probability of Type I errors. We take a number of steps to ensure that the results are not driven by our ex-ante choice of the relevant subgroups.

First, we adjust the p-values following the bootstrap-based procedure, as recommended by List et al. (2019).<sup>20</sup> The results in Table II retain similar levels of statistical significance, with an adjusted p-value of 0.024 for boys and 0.723 for girls.

Second, in Appendix Table A.X, we examine heterogeneity of the ITT effect along additional dimensions beyond gender, reporting the coefficient for each subgroup in column 1, the p-value in column 2, and the p-value adjusted for multiple hypothesis testing in column 3. Each panel of the table focuses on a different, additional dimension of heterogeneity (on top of gender): mother’s education in Panel A, terciles of INVALSI6 in Panel B, and EU citizenship in Panel C. While most p-values increase above standard levels once we adjust for multiple hypothesis testing, the pattern of coefficients suggests that boys and girls from low socio-economic background benefited the most from EOP.

Finally, to capture high-dimensional combination of covariates that the researcher-specified interactions may miss, we explore heterogeneity using machine learning tools, following the recent literature on heterogeneous treatment effects (Athey and Imbens, 2016; Davis and Heller, 2017; Bertrand et al., 2017). Online Appendix C describes our methodology in detail.

In a nutshell, we estimate the Conditional Average Treatment Effect (CATE), including in the causal forest the following baseline characteristics: gender, INVALSI6 (squared polynomial), generation of immigration, mother’s and father’s education and occupation, school province, and region of citizenship. We use the predictions on the expected treatment effect for each individual, given the covariates, to investigate treatment heterogeneity. We divide the sample in two groups, i.e., top and bottom half of the predictions. Appendix Table C.I reports the balance test for the CATE and p-values adjusted for multiple hypothesis testing. Figure 6 shows a heatmap of the CATE by gender, mother’s education and INVALSI6.

[Insert Figure 6]

Overall the results on gender differences are consistent with our main analysis: girls are over-represented among students with lower CATE. However, additional interesting results emerge. Boys and girls with low or missing levels of mother’s education and students in the central part of the ability distribution benefit the most from participating in EOP.<sup>21</sup> This suggests that targeting only boys in future scaled-up interventions may miss girls that could substantially benefit from EOP, such as girls in the middle of the initial test score distribution and those with parents with low-socio economic status.

---

<sup>20</sup>As shown in List et al. (2019), under weak assumptions, this procedure asymptotically controls for Familywise Error Rate (FWER), i.e. for the probability of getting one or more false rejections.

<sup>21</sup>As shown in Appendix Table A.III and discussed in Section 4.3, missing information on parents’ education is a proxy for low socio-economic status.

### 5.3 Mechanisms

In this section we try to disentangle the effects of different components of the program and explore the mechanisms through which they impacted educational choices.

#### **Career counseling vs. academic tutoring**

As explained in Section 3, the EOP intervention comprised two components: motivational and career counseling (the Social Cognitive Career Theory module), and academic tutoring to provide students with a method for studying (the CALP module). Ideally, one would want to disentangle the contribution of each of these two components using a factorial treatment design. However, ethical constraints prevented us from delivering the motivational treatment without also delivering help for studying. This means we cannot disentangle the contribution of CALP on the extensive margin, but we can assess the effectiveness of CALP on the intensive margin by comparing students who scored below and above the cutoff used to determine the number of CALP meetings. In particular, while virtually every treated student was invited to at least 17 CALP meetings in grade 8, students who had scored below 65 in the entry level test (INVALSI6) were invited to an additional 12 CALP meetings in grade 7. Achieving a score of 65 does not impact students' careers in any other way, so this rule provides a suitable regression discontinuity design.<sup>22</sup>

[Insert Figure 7]

The first two graphs in Figure 7 show that students with a score below 65 attended on average 5.5 more CALP meetings than students with a score above 65, whereas there are no differences in the number of career counseling meetings. Therefore, being invited to additional tutoring meetings increased the 'dosage' of CALP. The two graphs at the bottom of Figure 7 show that there is no significant discontinuity in the probability of choosing the high track or in grade retention between students on one side or the other of the cutoff. This implies that neither one of the main outcomes was significantly affected by an increase in the number of CALP meetings.<sup>23</sup> Therefore, at the intensive margin, EOP seems to operate mostly through motivational and career counseling, as opposed to specialized help for studying. This of course does not exclude that academic tutoring may have played a role at the extensive margin.

---

<sup>22</sup>Appendix Figure A.7 confirms that the density of INVALSI6 does not change discontinuously around the cutoff, using a formal McCrary test.

<sup>23</sup>The estimated changes at the cutoff are reported in Panel A of Appendix Table A.XI.

## Personality and cognitive skills

[Insert Table IV ]

In Table IV we report the effects of EOP on the cognitive and personality skills described in Section 4.2.<sup>24</sup> The dependent variables include the indexes of aspirations and perceived barriers whose components and loading factors are reported in Appendix Table A.VI, as well as measures of academic performance.

Starting with personality skills, the intervention substantially increased students' aspirations, especially for males (+0.31 standard deviations), whereas the effect is weaker and not statistically significant for females. EOP also reduced students' perceptions that their choices would be limited by barriers such as financial constraints, prejudice, or family plans. The effect is sizable (a reduction in perceived barriers of about 0.4 standard deviations) and virtually identical between males and females. In the last row of each panel, we also report the p-values adjusted for multiple hypothesis testing (List et al., 2019), considering jointly the four outcomes and the two subgroups by gender. Appendix Table A.XII shows that the effect is generally statistically significant for the individual psychological measures that enter the two indexes, even after we account for multiple hypothesis testing (last column).

Columns 1 and 2 of the bottom panel of Table IV show that EOP increased the standardized test score in grade 8 (INVALSI8) for male students but not for females. The effect is smaller than that on personality traits (+0.16 standard deviations). Appendix Figure A.8 compares the entire distribution of aspirations, perceived barriers, and INVALSI8 across treated (solid red line), control (solid blue line), and native (dashed black line) students who had the same standardized test score in grade 6. The pattern in the figure confirms that EOP caused a shift in the distribution of all three outcomes for males, and only of perceived barriers for females.

In the last two columns of the bottom panel of Table IV we estimate the effect of EOP on teachers' recommendations. Recall that these recommendations are made about half-way through grade 8, and before students make their high school track choice. We find that, on average, the probability that teachers recommend the high track is 17.1 percentage points higher for male immigrant students in EOP schools, on a baseline of 37.1 percent in control schools. Importantly, teachers do *not* significantly revise their recommendations for female students in treated schools: the point estimate for females is less than half of that for males and is not significantly different from zero. This is an important result to gain insights into what determines teachers' recommendations. teachers

---

<sup>24</sup>The sample differs across columns because questionnaires measuring personality skills were administered only to a 50 percent random sample of control students, see Section 4.2.

do not mechanically adjust their track recommendations simply because they know that a student is enrolled in EOP

If teachers were simply reacting to knowledge about the program, mechanically adjusting their track recommendations upwards for all students in EOP, we would find similar effects for boys and girls because they were enrolled in EOP in equal proportions. Instead, we find that teachers only revise their recommendations for boys, consistent with teachers reacting to detectable changes in motivation and test scores (recall that aspirations and cognitive skills only improve for boys).<sup>25</sup>

### **Decomposing the treatment effect**

Following Heckman et al. (2013), we decompose the treatment effect on educational choices into experimentally induced changes in the mediating factors listed in Table IV and changes in other (unmeasured) factors. This decomposition method requires strong assumptions that do not leverage the experimental variation. In particular, one needs to assume that the observable mediating factors are independent of the unobservables in the control group. The results in this section should thus be interpreted in light of such assumptions. Details on the methodology can be found in Online Appendix D; here we report the main findings.

[Insert Table V]

Table V shows the decomposition of the effect of EOP on the educational choices of boys.<sup>26</sup> Changes in personality skills explain about one third of the overall effect (column 1). However, this effect is entirely driven by aspirations, whereas perceptions of barriers do not seem an important mediating factor. This is consistent with the fact that males and females experience a similar decrease in perceived barriers, but educational choices only change for males. In column 2 we add school performance (measured by the standardized test score in grade 8) as an additional mediating factor, and in column 3 we further add teachers' recommendation.

We find that aspirations and teachers' recommendation are the most important factors, jointly explaining about two thirds of the treatment effect on educational choices. The importance of the first channel is in line with experimental evidence showing that raising aspirations through role models increases educational investment (Tanguy et al., 2014). Our result on teachers' recommendations aligns with the literature showing that teachers' expectations on students' performance play

---

<sup>25</sup>This does not imply that teachers did not contribute to changes in students' aspirations and performance. For example, EOP may have led teachers to pay more attention to immigrants who had misaligned aspirations (e.g., boys), and this may have contributed to improving students' aspirations and test scores.

<sup>26</sup>We only present results for males because treatment effects on females are not significantly different from zero.

a crucial role in affecting educational choices (Papageorge et al., 2018).<sup>27</sup> We obtain very similar results when we employ the alternative decomposition method by Gelbach (2016), see Appendix Table A.XIV.

To sum up, although EOP did improve school performance and reduce perceived barriers, these two factors play a less crucial role as a direct channel of influence in track choice of high-achieving immigrants.

## 5.4 Longer term effects

The results presented so far show that EOP increased the probability of enrolling into academic and technical schools (high track) after grade 8. A potential concern is that such schools may prove too demanding for the students who were affected by the treatment – a version of the so called ‘mismatch hypothesis’ (e.g., Arcidiacono and Lovenheim, 2016). One may worry that immigrant students may face additional constraints compared to Italian students of similar ability, for example, they may be less embedded in social networks that could help with studying challenging subjects, or they may face financial constraints in paying for private tutoring.

[Insert Table VI]

In Table VI we address this concern by estimating the effect of (assignment to) EOP on performance during the first two years of high school. We consider four different outcomes: (i) the probability of being admitted to grade 10, the second year of high school (columns 1-2); (ii) the number of make-up exams students need to take during the summer in order to avoid repeating the grade (typically no more than three exams)<sup>28</sup>; (iii) the probability that a student drops out before completing grade 10 (columns 5-6); and (iv) the probability of changing school between grade 9 and 11 (columns 7-8).

For all these outcomes, we find that treated students are no more likely to experience difficulties compared to the control group. If anything, they are more likely to be admitted to grade 10, less

---

<sup>27</sup>This phenomenon has been documented since the seminal work of Rosenthal and Jacobson (1968), whereby high and low expectations bring improvements and worsening in performance, respectively. Card and Giuliano (2016b) show that teachers tend to overlook the potential of high-ability minority students, especially English language learners. Relatedly, children of minority groups benefit the most from positive expectations of teachers and may be negatively affect by teacher stereotypes (Carlana, 2019; Alesina et al., 2018).

<sup>28</sup>The sample in columns 3-4 is smaller than in columns 1-2 because it only comprises students who did not fail outright and who did not drop out of high school. If a student is below the Pass level in too many subjects, the decision of the school is typically to fail this student rather than give make-up exams after the summer.



likely to fail courses and less likely to drop out, although the estimated coefficients are not significantly different from zero.<sup>29</sup> It is worth stressing that the lack of a significant effect should not be seen as a shortcoming: given that our treated students were more likely to enroll in demanding high schools, the fact that they are doing as well as the control group (and if anything better, given the pattern of coefficients in Table VI) is actually a positive result.

## 5.5 Spillover effects

Our last piece of evidence relates to spillover effects. Any change in the achievement and educational choices of treated students may influence their peers – particularly those sharing the same immigrant background (Sacerdote et al., 2011). For instance, treated students could serve as role models to other students in their social network (Patacchini and Zenou, 2016). This may be particularly relevant in our context, because we showed that EOP had a strong impact on aspirations and barriers perceived by treated students. We estimate the effect of EOP on treated students’ classmates exploiting random assignment of the intervention across schools. Specifically, we include in the sample only classmates of treated and control students and estimate the following equation:

$$Y_{ics} = \alpha + \beta EOPinClass_{cs} + \gamma X_{ics} + \delta Z_{cs} + u_{ics} \quad (1)$$

where  $i$  denotes the student,  $c$  the class, and  $s$  the school.  $EOPinClass_{cs}$  is a dummy equal to 1 if the student belongs to a class with at least one treated student, and 0 if he/she belongs to a class with at least one control student (i.e., a class in a control school that contains at least one immigrant student whose INVALSI6 score was among the top 10 of the school). In this way, the sample used in the regression contains students from classes that are comparable in terms of having high-achieving immigrant students among them, some of which received EOP and others not.  $X_{ics}$  includes our baseline controls (gender, dummy for first generation immigrants, and second degree polynomial in INVALSI6); and  $Z_{cs}$  includes school size, class size, and percentage of immigrants in the class.

[Insert Table VII]

Table VII shows the effect of being in class with a treated student on native (columns 1-3) and immigrant (columns 4-6) classmates. Immigrant males who are in class with a treated student do not change their track choice (Panel A) but they experience a decrease in grade retention (Panel B). On the other hand, immigrant females in the same class with a treated student are more likely to

<sup>29</sup>Ideally, we would also like to estimate the longer term effects on cognitive skills, as measured by the standardized test score in grade 10 (INVALSI10). Unfortunately, it is not possible to match MIUR registries for middle school students with their INVALSI tests in high school.

enroll in the high track (Panel A), with no detectable impact on grade retention. Interestingly, the effects on male grade retention and female track choice are comparable in size to the direct effects observed for treated boys. This can be rationalized by observing that, by construction, non-treated immigrant classmates are selected to have lower academic performance at the beginning of middle school (they were not among the top 10 immigrants in terms of INVALSI6). It is thus possible that immigrant students down in the ability distribution are very responsive to (even indirect) treatment effects. This interpretation is reinforced when observing that immigrant females in the mid-to-lower part of the distribution make different choices from comparable native females in the absence of the intervention (see Figure 1). Therefore it should not be surprising that, while EOP did not affect track choice of treated females (who were selected to be high performing and already made comparable choices to natives), it had an impact on non-treated females, for whom a ‘choice gap’ existed.

Why are male immigrant classmates not affected in terms of track choice (Panel A, column 5)? Our interpretation is that they are too far from the margin at which a demanding high school would become a preferable choice. This can be seen for example when comparing soft skills of immigrant boys and girls who did not qualify for the program. Female immigrant classmates of students eligible for EOP have aspiration levels that are closer to those of treated male students, compared to other male immigrant classmates.<sup>30</sup> At the same time, the reduction in the likelihood of grade retention for untreated male immigrant classmates (Panel B, column 5) suggests that spillover effects did operate at that (lower) margin.

Finally, columns 1-3 of Table VII show that there are no significant differences in outcomes for native classmates of treated students. This is consistent with previous evidence that peer effects are particularly strong within groups with a similar background, where students follow similar norms (see, e.g. Sacerdote et al., 2011; Bursztyin and Jensen, 2015). Overall, the positive spillovers on immigrant classmates uncovered by our analysis have significant implications when assessing the success and cost-effectiveness of the intervention. We briefly discuss the latter in the next section.

## 5.6 Cost-benefit analysis

Having established that EOP had sizeable and significant impacts, it is important to know if it is also cost-effective. A fully fledged cost benefit analysis is not possible at this stage: on top of

---

<sup>30</sup>The average index of aspirations for boys and girls in the control group is  $-0.11$  and  $0.13$ , respectively, while the average score of their immigrant classmates (i.e., immigrant students outside the best 10 in terms of INVALSI6) is  $-0.26$  and  $-0.13$ , respectively. Hence, in terms of average academic motivation, female immigrant classmates are closer than male classmates to students selected for the program. Elective affinity in soft skills may ease the diffusion of positive spillover effects of the program through friendship networks (Rapallini and Rustichini, 2016).

the challenge of quantifying non-pecuniary benefits and costs, students involved in our experiment have not yet completed secondary school and therefore lifetime earnings profiles are not observed (Heckman et al., 2018). The computation of the lifetime rate of return of EOP is therefore naturally based on assumptions about long-term outcomes such as college enrollment, earnings and unemployment. Although EOP may potentially have strong effects on health and criminal behavior, we present conservative estimates focusing our cost-benefit analysis only on social benefits coming from higher income taxes and public savings on unemployment insurance (Heckman et al., 2010; Eisenhauer et al., 2015).

We examine the sensitivity of social rates of returns to a plausible range of assumptions. Appendix Table A.XV reports our calculations, which should be considered tentative for the reasons stated above. If we extrapolate the long-term benefits only on those treated individuals who were directly affected through a reduction in grade retention or a change in their high-school choice, we estimate social rates of return between 3 and 5 percent. However, when we include the positive spillovers on the immigrant classmates of treated students, we estimate that the lifetime rate of return of EOP ranges between 6.6 and 8.8 percent, close to the historical return on equity.

## 6 Conclusions

Educational segregation is a significant risk in societies where school tracking occurs at an early age. This risk disproportionately affects students whose parents have less information about or are less integrated in the local education system, such as children of immigrants. We show that it is possible to reduce the mismatch created by early tracking through an innovative program that provides a mix of career counseling and academic tutoring. The program, known as EOP and implemented in a random sample of middle schools in Northern Italy, targeted high-achieving immigrant students selected on the basis of their test performance in grade 6. Two years later, immigrant boys assigned to treatment were 12 percent more likely than control ones to enroll in academic or technical high schools (as opposed to vocational ones), virtually closing the gap with native boys. No effect was found for treated girls, for whom no mismatch was detected in the first place.

The significance and magnitude of our effects is noteworthy when compared to existing evidence from randomized field experiments that study partial derivatives of the human capital production function, recently summarized by Kautz et al. (2014) and Fryer Jr (2016). For instance, the latter shows that, among school interventions, only ‘high-dosage’ tutoring (defined as “being

tutored in groups of 6 or fewer students for 4 or more days per week”) has significant effects.<sup>31</sup> Our EOP treatment can be considered ‘low dosage’, given the above definition, yet its impact on treated boys is 0.19 standard deviations for math achievement and 0.14 standard deviations for reading.

Although data constraints prevent us from performing a fully fledged cost benefit analysis, tentative estimates suggest internal rate of returns between 7 and 9 percent, after accounting for spillovers on non-treated students. Our finding that soft skills played a more important role than improved test performance in determining high school choice suggests that scaled down versions of the program may be even more cost effective. For example, one could target the program to the students with the highest potential benefit or reduce the number of meetings with academic tutors and explore forms of delivering the information and motivational components of the program through teachers as opposed to dedicated counselors, to reduce costs. We leave this to future work.

## References

- Alesina, A., Carlana, M., La Ferrara, E., and Pinotti, P. (2018). Revealing stereotypes: Evidence from immigrants in schools. *NBER Working Paper No. 25333*.
- Almås, I., Cappelen, A. W., Salvanes, K. G., Sørensen, E. Ø., and Tungodden, B. (2016). What explains the gender gap in college track dropout? Experimental and administrative evidence. *American Economic Review*, 106(5):296–302.
- Angrist, J. D., Cohodes, S. R., Dynarski, S. M., Pathak, P. A., and Walters, C. R. (2016). Stand and deliver: Effects of Boston’s charter high schools on college preparation, entry, and choice. *Journal of Labor Economics*, 34(2):275–318.
- Arcidiacono, P. and Lovenheim, M. (2016). Affirmative action and the quality-fit trade-off. *Journal of Economic Literature*, 54(1):3–51.
- Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360.
- Athey, S., Tibshirani, J., Wager, S., et al. (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

---

<sup>31</sup>Fryer Jr (2016) finds that high-dosage tutoring leads to a meta-coefficient of 0.309 standard deviations for math achievement (with a standard error of 0.106) and 0.229 standard deviations for reading achievement (with a standard error of 0.033).

- Autor, D., Figlio, D., Karbownik, K., Roth, J., and Wasserman, M. (2019). Family disadvantage and the gender gap in behavioral and educational outcomes. *American Economic Journal: Applied Economics*, 11(3):338–81.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ Prentice Hall.
- Bertrand, M., Crépon, B., Marguerie, A., and Premand, P. (2017). Contemporaneous and post-program impacts of a public works program: Evidence from côte d’ivoire. *Working Paper*.
- Bertrand, M. and Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics*, 5(1):32–64.
- Betts, J. (2011). The economics of tracking in education. *Handbook of the Economics of Education*, 3:341–381.
- Bobba, M. and Frisancho, V. (2016). Perceived ability and school choices. *TSE Working Paper No. 660, Toulouse School of Economics*.
- Brenøe, A. A. and Lundberg, S. (2019). Gender gaps in the effects of childhood family environment: Do they persist into adulthood? *European Economic Review*, 109:42–62.
- Brown, D. (2002). *Career choice and development*. John Wiley & Sons.
- Brunello, G. and Checchi, D. (2007). Does school tracking affect equality of opportunity? new international evidence. *Economic Policy*, 22(52):781–861.
- Bursztyn, L. and Jensen, R. (2015). How does peer pressure affect educational investments? *The quarterly journal of economics*, 130(3):1329–1367.
- Card, D. and Giuliano, L. (2016a). Can tracking raise the test scores of high-ability minority students? *American Economic Review*, 106(10):2783–2816.
- Card, D. and Giuliano, L. (2016b). Universal screening increases the representation of low-income and minority students in gifted education. *Proceedings of the National Academy of Sciences*, 113(48):13678–13683.
- Carlana, M. (2019). Implicit stereotypes: Evidence from teachers’ gender bias. *The Quarterly Journal of Economics*, 134(3):1163–1224.

- Chetty, R., Hendren, N., Jones, M. R., and Porter, S. R. (2018). Race and economic opportunity in the united states: An intergenerational perspective. *NBER Working Paper No. 24441*.
- Chetty, R., Hendren, N., Lin, F., Majerovitz, J., and Scuderi, B. (2016). Childhood environment and gender gaps in adulthood. *American Economic Review*, 106(5):282–288.
- Conger, D. and Long, M. C. (2010). Why are men falling behind? gender gaps in college performance and persistence. *The Annals of the American Academy of Political and Social Science*, 627(1):184–214.
- Cunha, F. and Heckman, J. (2007). The technology of skill formation. *American Economic Review*, 97(2):31–47.
- Dalton, P. S., Ghosal, S., and Mani, A. (2016). Poverty and aspirations failure. *The Economic Journal*, 126(590):165–188.
- Davis, J. and Heller, S. B. (2017). Using causal forests to predict treatment heterogeneity: An application to summer jobs. *American Economic Review: Papers and Proceedings*, 107(5):546–50.
- DiPrete, T. A. and Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research*, 41(1):1–15.
- Dustmann, C. (2004). Parental background, secondary school track choice, and wages. *Oxford Economic Papers*, 56(2):209–230.
- Eisenhauer, P., Heckman, J. J., and Vytlačil, E. (2015). The generalized Roy model and the cost-benefit analysis of social programs. *Journal of Political Economy*, 123(2):413–443.
- Fryer Jr, R. G. (2016). The production of human capital in developed countries: Evidence from 196 randomized field experiments. Technical report, National Bureau of Economic Research.
- Fryer Jr, R. G. and Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *The Review of Economics and Statistics*, 86(2):447–464.
- Gelbach, J. B. (2016). When do covariates matter? and which ones, and how much? *Journal of Labor Economics*, 34(2):509–543.
- Genicot, G. and Ray, D. (2017). Aspirations and inequality. *Econometrica*, 85(2):489–519.

- Giustinelli, P. (2016). Group decision making with uncertain outcomes: Unpacking child–parent choice of the high school track. *International Economic Review*, 57(2):573–602.
- Goodman, S. (2016). Learning from the test: Raising selective college enrollment by providing information. *The Review of Economics and Statistics*, 98(4):671–684.
- Goux, D., Gurgand, M., and Maurin, E. (2017). Adjusting your dreams? the effect of school and peers on dropout behaviour. *Economic Journal*, 127(602):1025–1046.
- Guyon, N. and Huillery, E. (2016). Biased aspirations and social inequality at school: Evidence from french teenagers. *LIEPP Working Paper No. 44*.
- Guyon, N., Maurin, E., and McNally, S. (2012). The effect of tracking students by ability into different schools a natural experiment. *Journal of Human Resources*, 47(3):684–721.
- Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6):2052–86.
- Heckman, J. J., Humphries, J. E., and Veramendi, G. (2018). Returns to education: The causal effects of education on earnings, health and smoking. *Journal of Political Economy*, 126(S1):S197–S246.
- Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4):451–464.
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., and Yavitz, A. (2010). The rate of return to the Highscope Perry Preschool Program. *Journal of public Economics*, 94(1–2):114–128.
- Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.
- Hoxby, C. M. and Avery, C. (2013). The missing “one-offs”: The hidden supply of high-achieving, low income students. *Brookings Papers on Economic Activity*, 2013(1):1–65.
- Jackson, J. F. and Moore III, J. L. (2008). The African American male crisis in education: A popular media infatuation or needed public policy response. *American Behavioral Scientist*, 51(7):847–853.

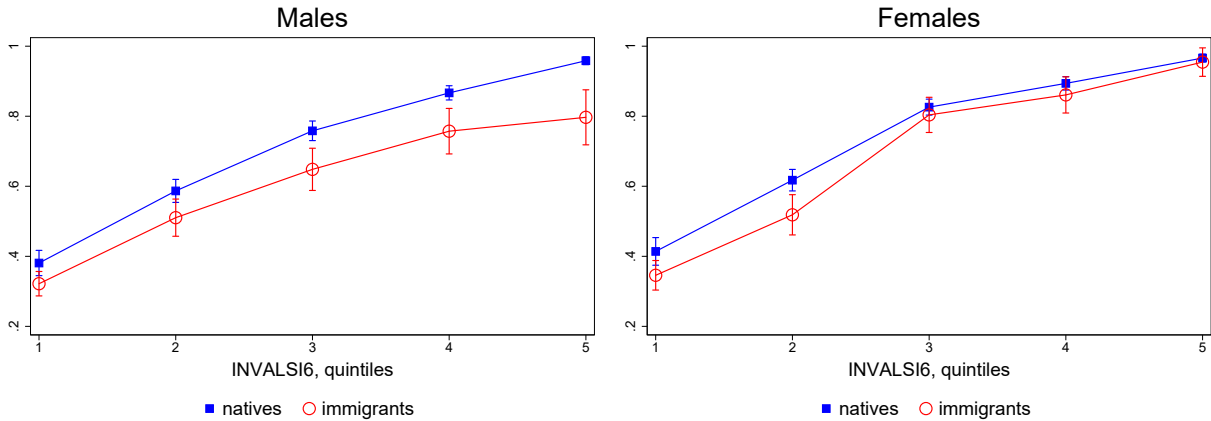


- Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., and Borghans, L. (2014). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. *Report prepared for the Organisation of Economic Co-operation and Development, Paris*, <http://www.oecd.org/edu/cei/Fostering-and-Measuring-Skills-Improving-Cognitive-and-Non-Cognitive-Skills-to-Promote-Lifetime-Success.pdf>.
- Lent, R. W., Brown, S. D., and Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45(1):79–122.
- List, J. A., Shaikh, A. M., and Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4):773–793.
- Lopez, N. (2003). *Hopeful girls, troubled boys: Race and gender disparity in urban education*. Psychology Press.
- Malamud, O. and Pop-Eleches, C. (2011). School tracking and access to higher education among disadvantaged groups. *Journal of Public Economics*, 95(11):1538–1549.
- Martins, P. S. (2010). Can targeted, non-cognitive skills programs improve achievement? Evidence from EPIS. *CEG-IST and IZA Discussion Paper No. 5266*.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.
- Mookherjee, D., Ray, D., and Napel, S. (2010). Aspirations, segregation, and occupational choice. *Journal of the European Economic Association*, 8(1):139–168.
- OECD (2013). PISA 2012 Results: What makes schools successful? Resources, policies and practices, (Volume IV). *OECD Publishing*, <http://dx.doi.org/10.1787/9789264201156-en>.
- OECD (2018). The resilience of students with an immigrant background: Factors that shape well-being. *OECD Publishing, Paris*, <https://dx.doi.org/10.1787/9789264292093-en>.
- Papageorge, N. W., Gershenson, S., and Kang, K. M. (2018). Teacher expectations matter. *NBER Working Paper No. 25255*.
- Patacchini, E. and Zenou, Y. (2016). Racial identity and education in social networks. *Social Networks*, 44:85–94.

- Pop-Eleches, C. and Urquiola, M. (2013). Going to a better school: Effects and behavioral responses. *American Economic Review*, 103(4):1289–1324.
- Rapallini, C. and Rustichini, A. (2016). Elective affinities matter as much as ethnicity in multi-ethnic schools. *Journal of Economic Behavior and Organization*, 131(20):243–262.
- Rodriguez-Planas, N. (2012). Longer-term impacts of mentoring, educational services, and learning incentives: Evidence from a randomized trial in the United States. *American Economic Journal: Applied Economics*, 4(4):121–139.
- Rosenthal, R. and Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review*, 3(1):16–20.
- Sacerdote, B. et al. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? *Handbook of the Economics of Education*, 3:249–277.
- Tanguy, B., Dercon, S., Orkin, K., and Taffesse, A. (2014). The future in mind: Aspirations and forward-looking behaviour in rural Ethiopia. *CEPR Discussion Paper No. DP10224*.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association.
- Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.
- Westfall, P. H. and Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*. John Wiley & Sons.

# Figures

Figure 1: Probability of enrolling in the high track at the end of middle school, by quintile of standardized test score in grade 6 (INVALS6)



Notes: This figure compares the probability of enrolling in the high-track between immigrant and native students, by quintiles of performance in the standardized test in grade 6 (INVALS6). The sample includes all students in the 75 control schools.

Figure 2: Time Line

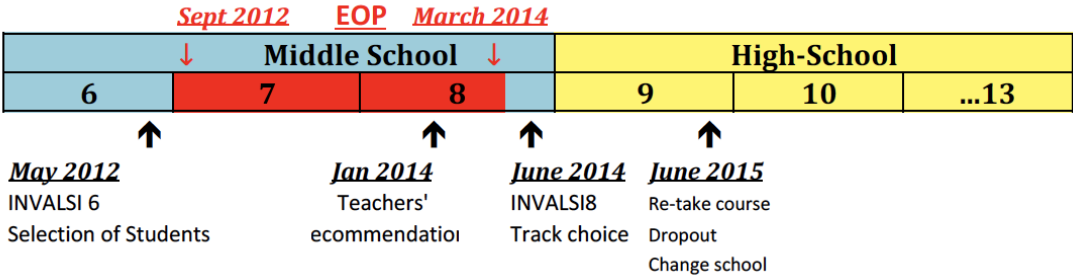
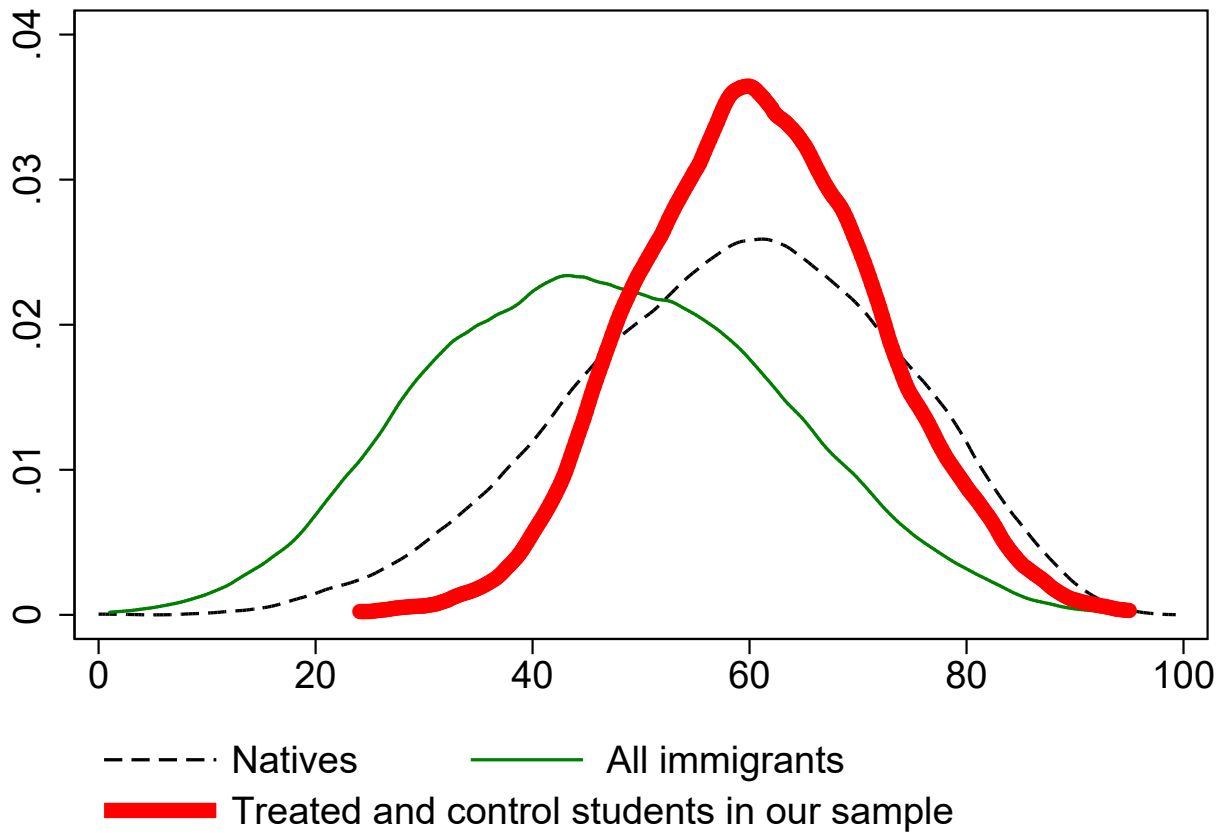
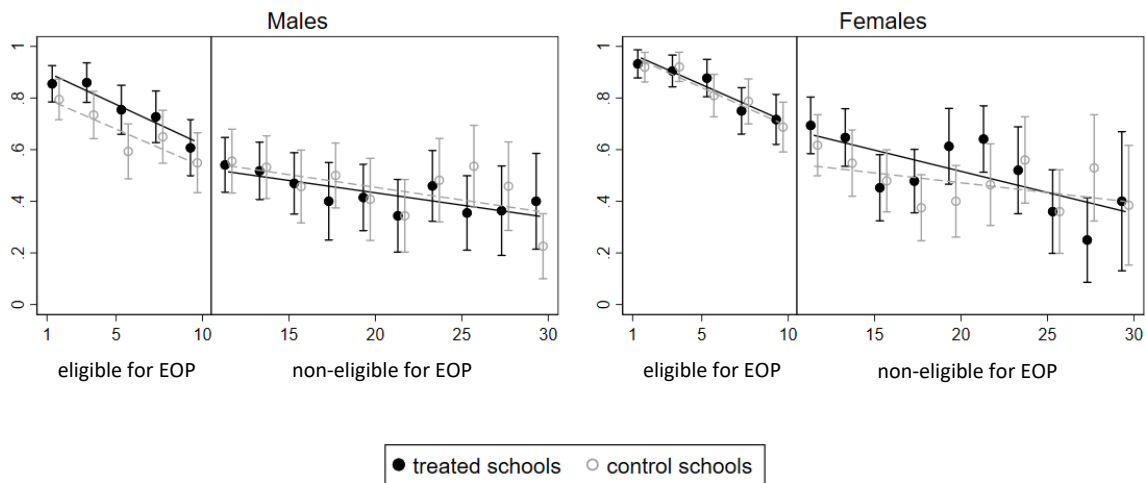


Figure 3: Distribution of standardized test score in grade 6 (INVALS6)



*Notes:* This figure compares the distribution of standardized test score in grade 6 (INVALS6) across native students, immigrant students, and treated and control students in our sample.

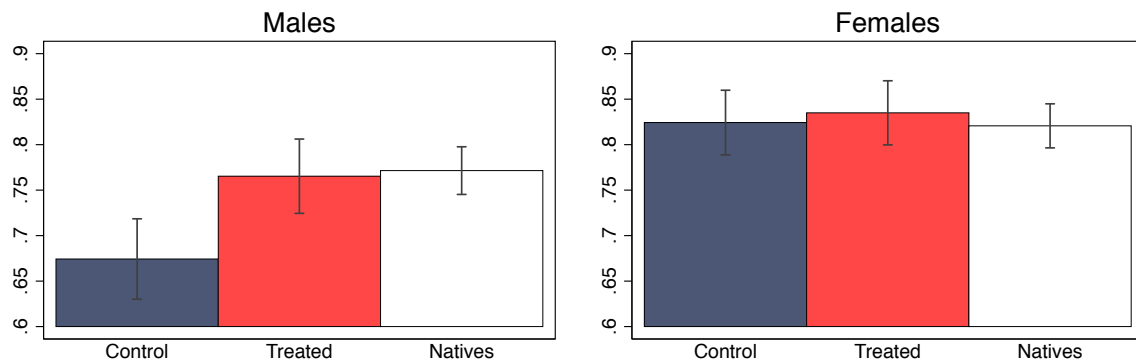
Figure 4: The effect of EOP on educational choices, difference-in-discontinuities at the eligibility cutoff



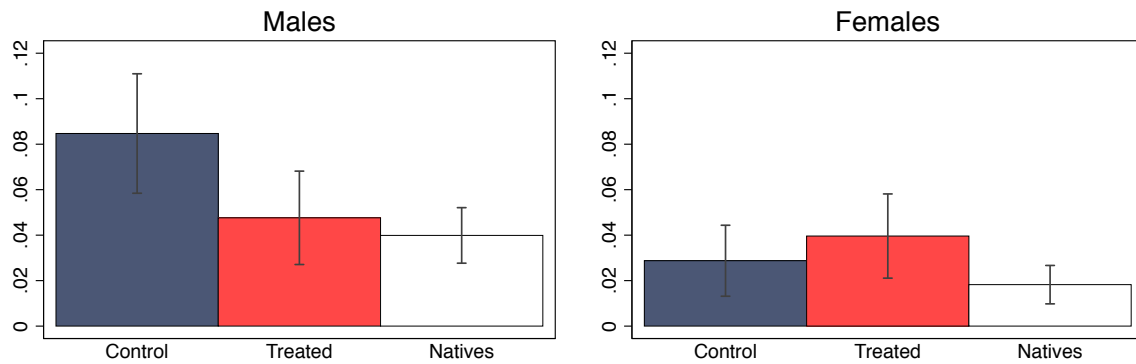
Notes: These graphs shows the average probability (and associated confidence intervals) of choosing the high track for male and female immigrant students in treated and control schools, by ranking in INVALSI6 within each school. The top 10 students in each school were eligible for EOP, while the others down in the ranking are not. The vertical line represents the eligibility cutoff.

Figure 5: Track choice and grade retention of immigrants and comparable natives

### Panel A: Probability of choosing the high track

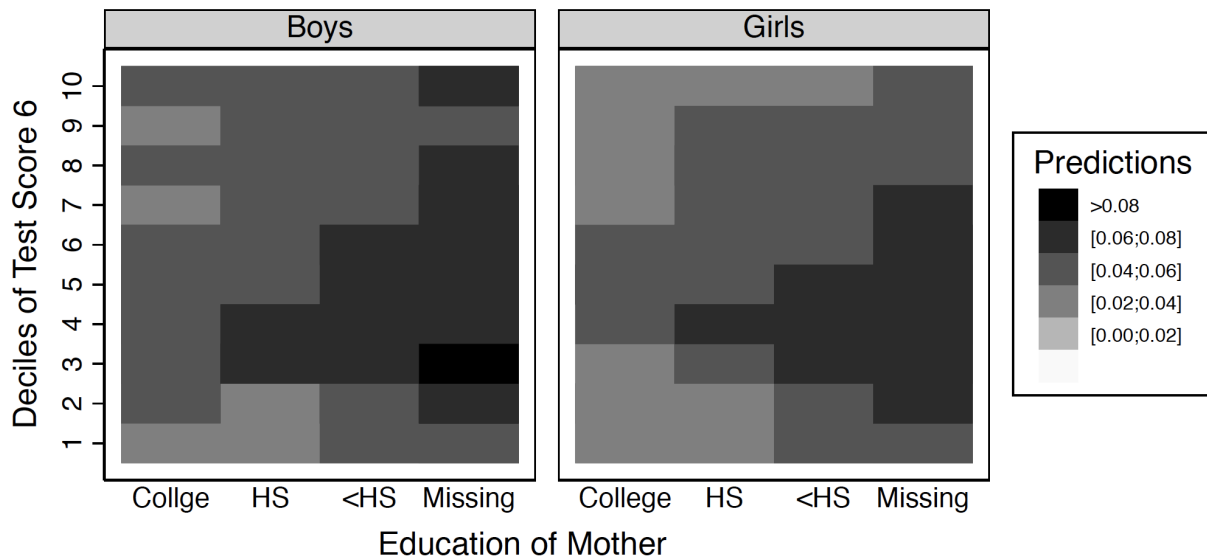


### Panel B: Retention rate in grade 7 or 8



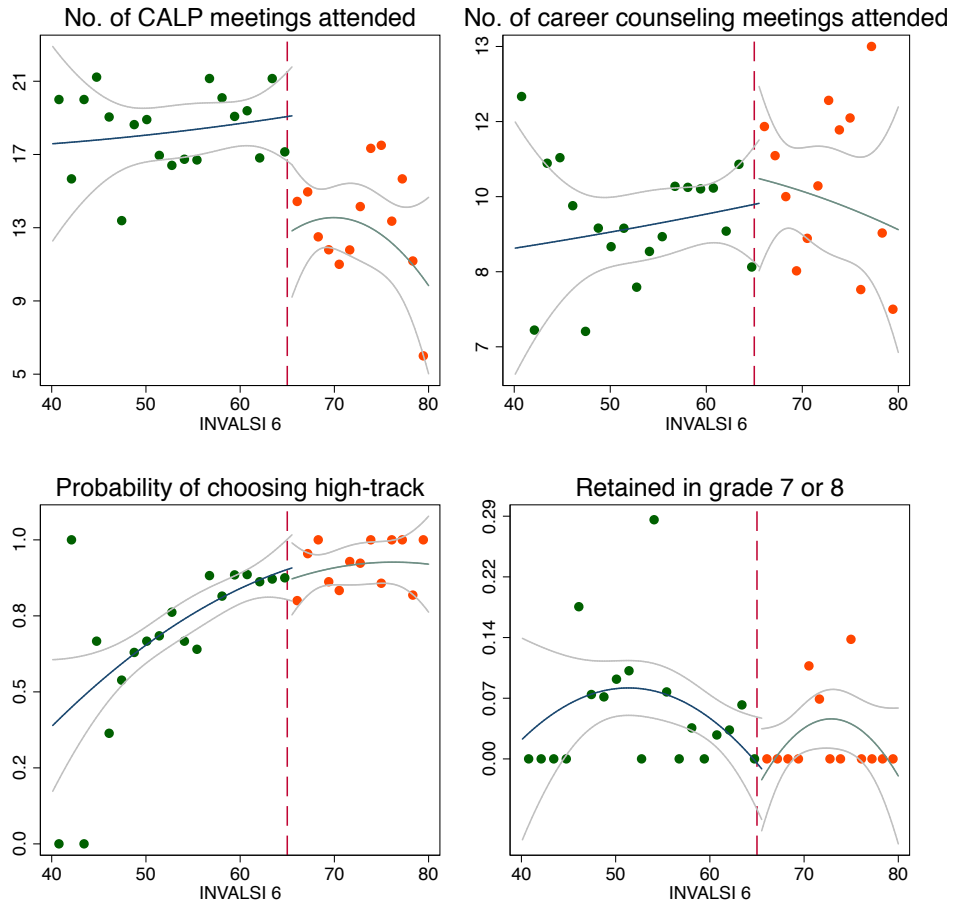
*Notes:* These graphs shows the average probability (and associated confidence intervals) of choosing the high-track (top graphs) and being retained in grade 7 or 8 (bottom graphs) for treated students, control students, and a group of Italian students that are comparable in terms of schooling ability. Specifically, we match each immigrant student with a native student of the same gender obtaining exactly the same score in INVALSI6.

Figure 6: Heatmap of the Conditional Average Treatment Effect by gender, mother education, and INVALSI6



Notes: This figure shows the value of the Conditional Average Treatment Effect in bins identified by gender, mother education, and INVALSI6. In the figure, HS refers to High-School Diploma and <HS to an education level lower than high-school diploma.

Figure 7: Effect of additional CALP meetings, regression discontinuity estimates



*Notes:* These graphs plot the number of meetings attended – distinguishing between career counselling and CALP modules – and treated students’ outcomes against standardized test scores in grade 6 (INVALSI6). The vertical line indicates the cutoff score below which treated students are offered additional CALP meetings.



# Tables

Table I: Treated and control students, balance test

		Full Sample	Treated	Controls	Difference	P-value	Std. Difference
Panel A: Student characteristics							
	Female	0.506	0.508	0.505	0.003	[0.93]	-0.006
	Test score in grade 6 (INVALSI6)	60.82	60.93	60.71	0.224	[0.86]	0.005
	First generation immigrant	0.555	0.547	0.561	-0.014	[0.73]	0.028
	Born before 1999	0.257	0.242	0.273	-0.031	[0.25]	0.071
	Brescia	0.179	0.165	0.194	-0.029	[0.67]	0.076
	Genova	0.067	0.074	0.06	0.014	[0.75]	-0.056
	Milan	0.496	0.476	0.516	-0.04	[0.65]	0.080
	Padua	0.055	0.064	0.047	0.017	[0.68]	-0.074
	Turin	0.203	0.223	0.184	0.039	[0.58]	-0.097
Panel B: Family characteristics							
Mother	Less than high school	0.224	0.228	0.221	0.007	[0.87]	-0.017
	High-school	0.464	0.315	0.281	0.034	[0.36]	-0.074
	Some post-secondary education	0.119	0.112	0.126	-0.014	[0.63]	0.043
	Missing Education	0.359	0.345	0.373	-0.028	[0.65]	0.058
	Blue collar	0.237	0.243	0.231	0.012	[0.76]	-0.028
	White collar	0.124	0.124	0.124	-0.000	[1.00]	0.000
	Unemployed	0.055	0.062	0.048	0.014	[0.59]	-0.061
	At home	0.259	0.256	0.261	-0.005	[0.91]	0.011
Father	Missing Occupation	0.325	0.314	0.335	-0.021	[0.72]	0.045
	Less than high school	0.197	0.198	0.197	0.001	[0.98]	-0.003
	High-school	0.296	0.301	0.290	0.011	[0.79]	-0.024
	Some post-secondary education	0.099	0.099	0.10	-0.001	[0.98]	0.003
	Missing Education	0.407	0.402	0.413	-0.011	[0.86]	0.022
	Blue collar	0.357	0.351	0.361	-0.01	[0.80]	0.021
	White collar	0.198	0.211	0.185	0.026	[0.47]	-0.065
	Unemployed	0.062	0.058	0.066	-0.008	[0.70]	0.033
At home	0.012	0.012	0.013	-0.001	[0.93]	0.009	
	Missing Occupation	0.371	0.367	0.374	-0.007	[0.90]	0.014

*Notes:* This table shows the number and characteristics of treated and control students in our sample. P-values for difference in means are reported in square brackets. The last column also reports the standardized difference between group averages.

Table II: The effect of EOP on educational choices

Dependent variable: Choosing the high-track (= 1 if choose high track)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All immigrants		Male immigrants		Female immigrants		All immigrants	
EOP	0.051 (0.027)	0.043 (0.024)	0.091 (0.037)	0.080 (0.035)	0.011 (0.034)	0.009 (0.031)	0.091 (0.037)	0.077 (0.034)
Female X EOP							-0.080 (0.046)	-0.067 (0.043)
Constant	0.750 (0.019)	0.683 (0.038)	0.674 (0.028)	0.651 (0.048)	0.824 (0.022)	0.720 (0.049)	0.674 (0.028)	0.621 (0.041)
Mean dep. var. control	0.750	0.750	0.674	0.674	0.824	0.824	0.750	0.750
Observations	1,217	1,217	601	601	616	616	1,217	1,217
Controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.004	0.087	0.010	0.086	0.000	0.100	0.024	0.105

*Notes:* This table shows the effect of EOP on immigrant students' educational choices by the end of middle school. The dependent variable is a dummy equal to 1 for students choosing the high-track (academic or technical schools) and equal to zero otherwise. EOP is a dummy equal to 1 for students in schools assigned to the treatment group and equal to zero for schools assigned to the control group. Specifications in columns (2), (4), (6), and (8) control in addition for a squared polynomial in INVALSI6, a dummy equal to 1 for first generation immigrants, and province fixed effects. Standard errors clustered by school are reported in parentheses.

Table III: The effect of EOP on grade retention

Dependent variable: Grade retention (= 1 if repeat a grade)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All immigrants		Male immigrants		Female immigrants		All immigrants	
EOP	-0.013 (0.016)	-0.013 (0.016)	-0.037 (0.021)	-0.037 (0.021)	0.011 (0.019)	0.009 (0.019)	-0.037 (0.021)	-0.035 (0.020)
Female X EOP							0.048 (0.024)	0.044 (0.023)
Constant	0.056 (0.010)	0.073 (0.028)	0.085 (0.016)	0.097 (0.028)	0.029 (0.010)	0.048 (0.043)	0.085 (0.016)	0.097 (0.030)
Mean dep. var. control	0.056	0.056	0.085	0.085	0.029	0.029	0.056	0.056
Observations	1,217	1,217	601	601	616	616	1,217	1,217
Controls	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.001	0.017	0.006	0.022	0.001	0.023	0.009	0.024

*Notes:* This table shows the effect of EOP on immigrant students' grade retention during middle school. The dependent variable is a dummy equal to 1 for students retained in grade 7 or 8, and equal to zero otherwise. EOP is a dummy equal to 1 for students in middle schools assigned to the treatment group and equal to zero for schools assigned to the control group. Specifications in columns (2), (4), (6), and (8) control in addition for a squared polynomial in INVALSI6, a dummy equal to 1 for first generation immigrants, and province fixed effects. Standard errors clustered by school are reported in parentheses.

Table IV: The effect of EOP on mediating factors

	(1)	(2)	(3)	(4)
Dep. var.	Aspirations		Perception of barriers	
	Males	Females	Males	Females
EOP	0.312 (0.102)	0.105 (0.090)	-0.385 (0.121)	-0.422 (0.101)
Mean dep.	-0.209	0.248	0.252	0.420
Observations	321	366	321	366
R-squared	0.156	0.110	0.096	0.086
MHT p-value	0.063	0.330	0.004	0.000
Dep. var.	INVALSI8		Teachers' Recomm.	
	Males	Females	Males	Females
EOP	0.163 (0.051)	0.001 (0.050)	0.171 (0.053)	0.074 (0.056)
Mean dep.	-0.129	0.057	0.371	0.578
Observations	520	574	601	616
R-squared	0.472	0.423	0.166	0.103
MHT p-value	0.044	0.757	0.000	0.138

*Notes:* This table shows the effect of EOP on several mediating factors. Aspirations and perception of barriers are the two principal components extracted from the psychological measures collected through students' questionnaires. The individual variables included in each index and their loading factors are reported in Appendix Table A.VI. INVALSI8 is the score obtained in the standardized test at the end of middle school (grade 8). Teachers' recommendation is a dummy equal to 1 when the teacher recommends to enroll in the high-track and equal to zero otherwise. All specifications control for a squared polynomial in INVALSI6, a dummy equal to 1 for first generation immigrants, and province fixed effects. Standard errors clustered by school are reported in parentheses. In the last row of each panel, we report the p-values adjusted for multiple hypothesis testing, considering jointly the four outcomes and the heterogeneity by gender.

Table V: Decomposition of the effect of EOP on high-school choice, male students

	(1)		(2)		(3)	
	Explained	p-value	Explained	p-value	Explained	p-value
Aspirations	0.0413	[0.0005]	0.0359	[0.002]	0.0357	[0.001]
Barriers	-0.0095	[0.5756]	-0.006	[0.722]	-0.0024	[0.882]
INVALSI8			0.013	[0.103]	0.0101	[0.178]
Teachers' recommendation					0.0333	[0.046]
Total explained	0.0318		0.0429		0.0767	
EOP effect on choosing high-track	0.091		0.091		0.091	

*Notes:* This table decomposes the effect of EOP between changes in personality skills (aspirations and perception of barriers), increased schooling achievement (as measured by INVALSI8) and teachers' recommendations. As explained in Appendix D, the decomposition follows the method devised by Heckman et al. (2013). All specifications control in addition for a squared polynomial in INVALSI6, a dummy equal to 1 for first generation immigrants, and province fixed effects. Bootstrap standard errors clustered at school level generated from 1,000 iterations.

Table VI: Effect of EOP on long-term outcomes

Dep. var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Outcomes in grade 9				Outcomes in grade 10			
	Admitted to grade 10	Re-take courses			Dropout	Change school		
EOP	0.014 (0.032)	0.034 (0.042)	-0.099 (0.072)	-0.146 (0.115)	-0.034 (0.027)	-0.057 (0.039)	0.000 (0.022)	-0.005 (0.034)
Female X EOP		-0.033 (0.057)		0.084 (0.146)		0.045 (0.050)		0.009 (0.043)
Constant	0.380 (0.044)	0.300 (0.048)	0.346 (0.126)	0.429 (0.150)	0.343 (0.038)	0.406 (0.041)	0.161 (0.043)	0.169 (0.047)
Mean dep. var. control males	0.411	0.411	0.251	0.251	0.365	0.365	0.119	0.119
Mean dep. var. control females	0.578	0.578	0.075	0.075	0.211	0.211	0.100	0.100
Observations	933	933	918	918	1,157	1,157	881	881
R-squared	0.054	0.077	0.025	0.028	0.063	0.083	0.011	0.011

*Notes:* This table shows the effect of EOP on immigrant students' outcomes in the first two years of high school, indicated above each column. EOP is a dummy equal to 1 for students in middle schools assigned to the treatment group and equal to zero for schools assigned to the control group. All regressions control in addition for a squared polynomial in INVALSI6, a dummy equal to 1 for first generation immigrants, and province fixed effects. Standard errors clustered by school are reported in parentheses.

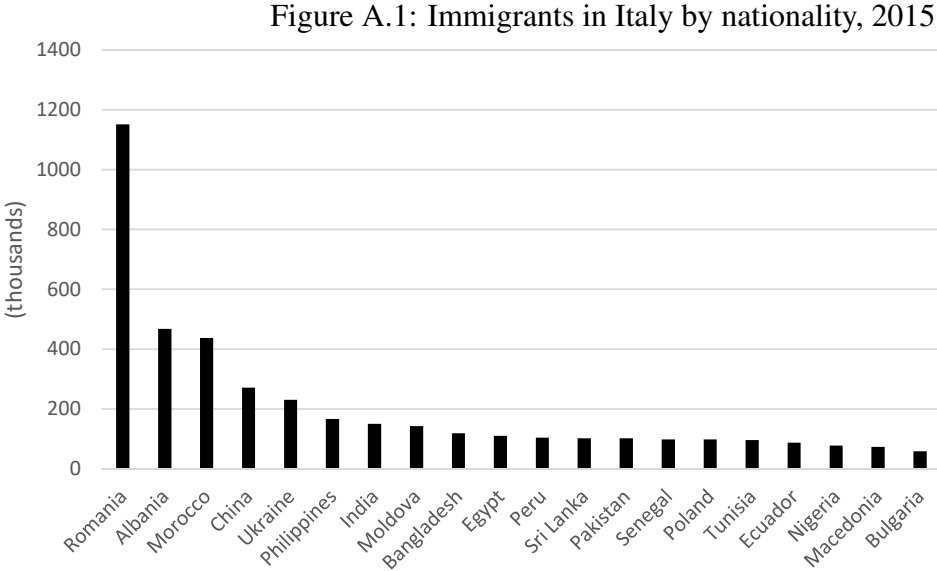
Table VII: Spillover effects in EOP schools

	(1)	(2)	(3)	(4)	(5)	(6)
	Native classmates			Immigrant classmates		
	All	Males	Females	All	Males	Females
<i>Panel A - Dependent Variable: Choosing the high-track</i>						
EOP class	0.001 (0.013)	0.008 (0.015)	-0.006 (0.017)	0.038 (0.029)	-0.009 (0.041)	0.091 (0.038)
Mean dep. var.	0.767	0.755	0.780	0.421	0.423	0.419
Observations	8,429	4,247	4,182	1,308	686	622
R-squared	0.197	0.205	0.193	0.097	0.096	0.125
<i>Panel B - Dependent Variable: Retained in grade 7 or 8</i>						
EOP class	-0.002 (0.007)	-0.005 (0.009)	0.002 (0.007)	-0.042 (0.023)	-0.064 (0.032)	-0.019 (0.027)
Mean dep. var.	0.044	0.053	0.035	0.148	0.192	0.099
Observations	8,429	4,247	4,182	1,308	686	622
R-squared	0.047	0.040	0.066	0.046	0.049	0.053
<i>Panel C - Dependent Variable: INVALSI8</i>						
EOP class	-0.017 (0.030)	-0.012 (0.034)	-0.021 (0.032)	-0.026 (0.062)	0.023 (0.077)	-0.069 (0.081)
Mean dep. var.	0.129	0.139	0.119	-0.961	-0.960	-0.965
Observations	7,533	3,736	3,797	1,007	500	507
R-squared	0.591	0.603	0.580	0.330	0.349	0.319
<i>Panel D - Dependent Variable: Teachers' recommendation</i>						
EOP class	-0.011 (0.037)	-0.012 (0.038)	-0.010 (0.039)	0.031 (0.038)	0.003 (0.039)	0.062 (0.052)
Mean dep. var.	0.597	0.574	0.621	0.207	0.210	0.205
Observations	8,429	4,247	4,182	1,308	686	622
R-squared	0.223	0.235	0.218	0.115	0.130	0.112

*Notes:* This table shows the effect of being in the same class with an immigrant student randomized into the intervention on several outcomes of interest, indicated in the title of each panel. The sample includes only classmates of treated and control students. The main explanatory variable, EOP class, is a dummy equal to 1 for the classmates of treated students. All regressions control in addition for all the individual characteristics in our baseline specification (dummy for first generation immigrants, second degree polynomial of test score in grade 6, and province fixed effects) as well as for class size, percentage of immigrants in the same class, and school size. Standard errors clustered by school are reported in parentheses.

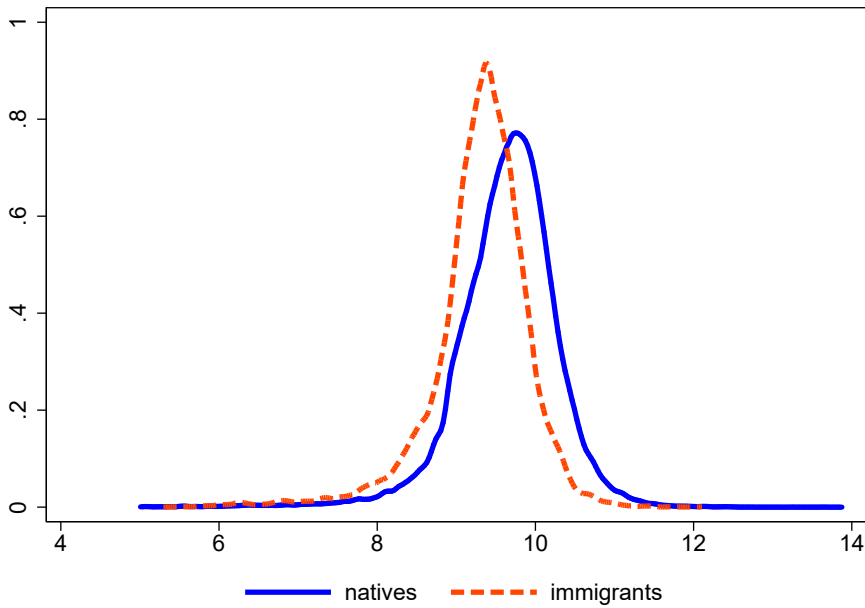
# Online Appendix - Not for publication

## A Appendix figures and tables



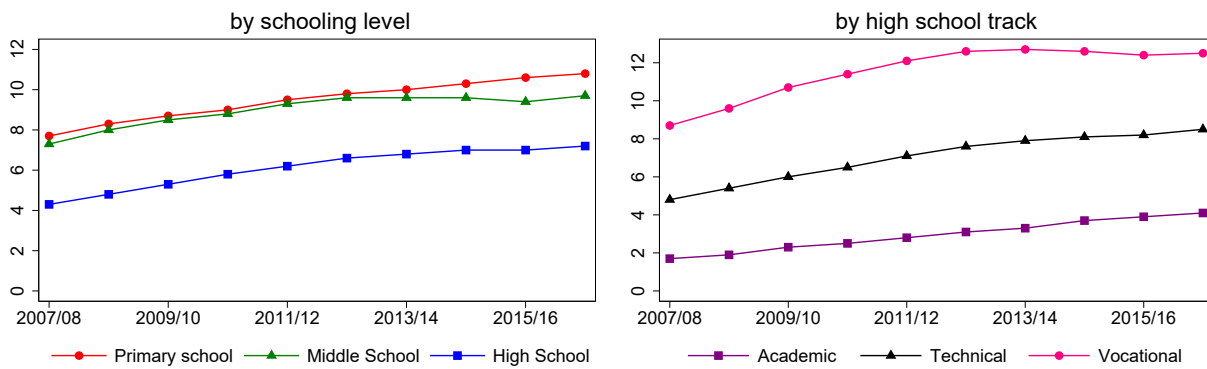
Source: ISTAT, “Demografia in Cifre”, several years ([www.demo.istat.it](http://www.demo.istat.it)).

Figure A.2: Distribution of (log) income across native and immigrant families in Italy



Notes: This graph shows the distribution of (log) disposable income per equivalent adult at constant 2010 prices.  
 Source: European Union Statistics on Income and Living Conditions (EU-SILC), 2007-2014.

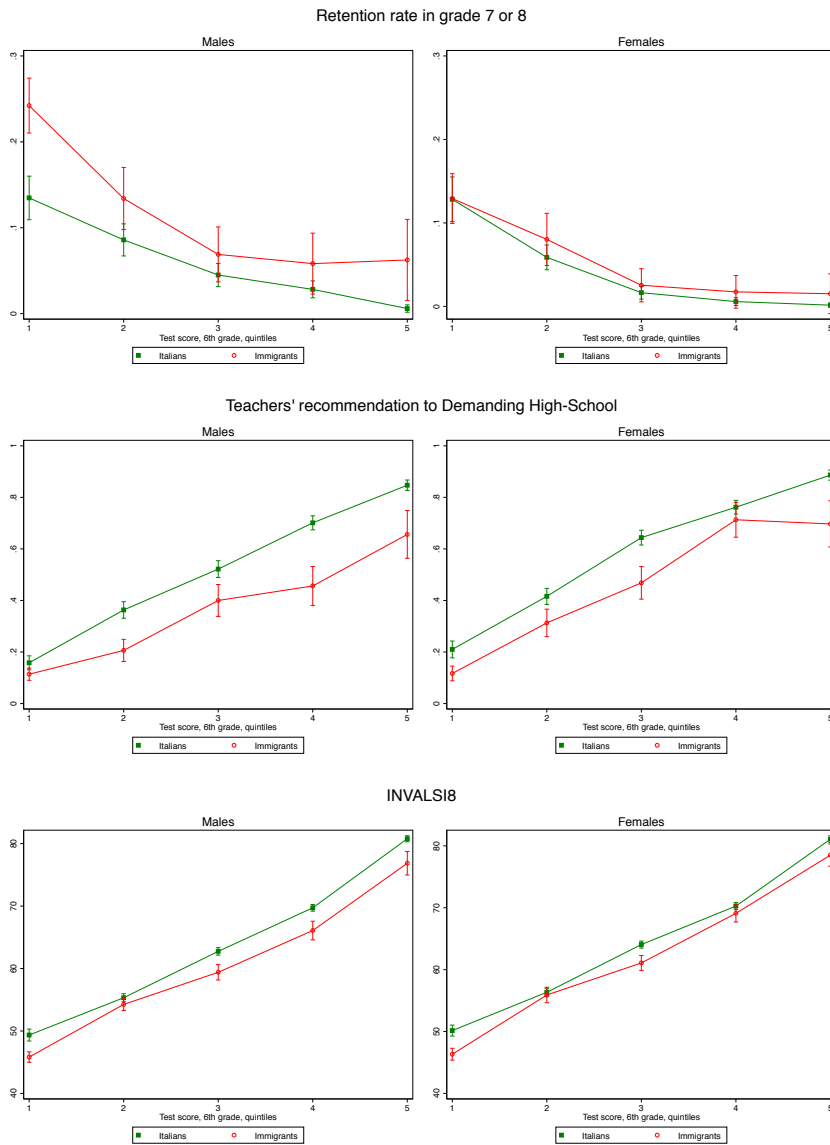
Figure A.3: Percentage of immigrants over total students in Italy, by schooling level and high school track



Source: MIUR, "Portale dei dati sulla scuola" (dati.istruzione.it), several years.

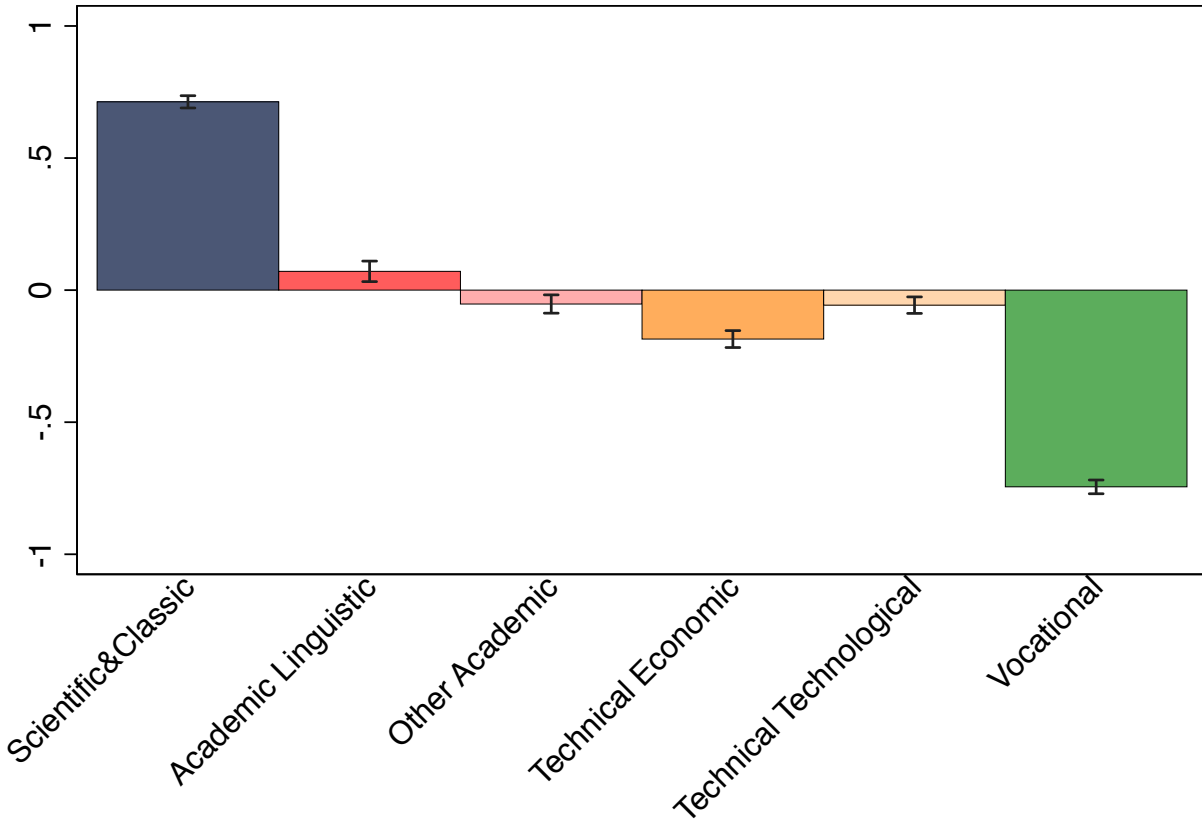


Figure A.4: Outcomes during middle school, by quintile of standardized test score in grade 6 (INVALS6)



Notes: This figure compares the probability of failing in grade 7 or 8 (Panel A), being recommended by teachers to the high-track (Panel B), and the test score in grade 8 (Panel C) between immigrant and native students, by quintiles of performance in the standardized test in grade 6 (INVALS6). The sample includes all students in the 75 control schools.

Figure A.5: Average test score in math in grade 8, by high school track

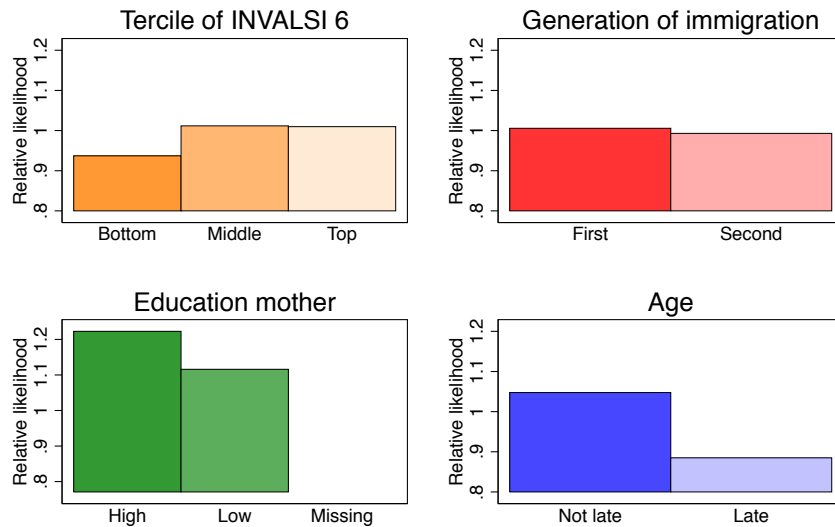


*Source:* This figure shows the average math test score in grade 8 of students who self-selected into different high-school tracks. Other academic includes social-psychology and arts. It shows a clear ranking in terms of ability between students in the scientific and classical academic track, other academic and technical tracks, and vocational track.

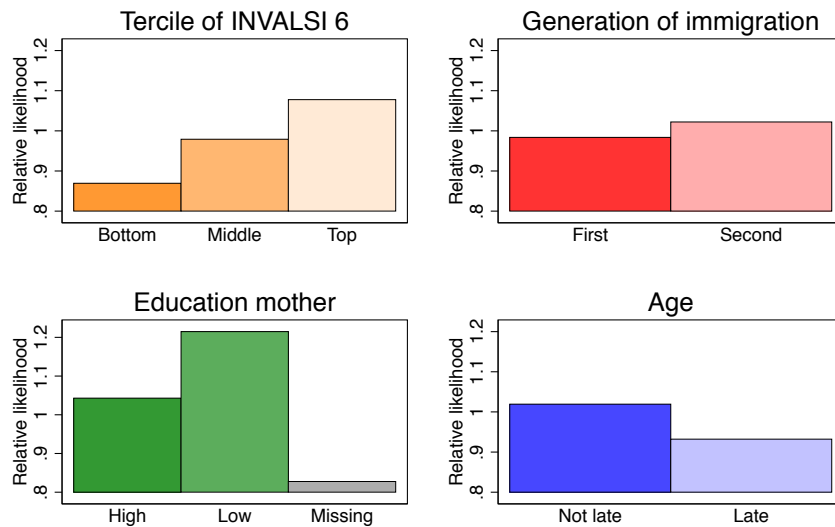
Figure A.6: Compliers' characteristics

## Compliers characteristics ratios (attending more than 75% of meeting)

### Male

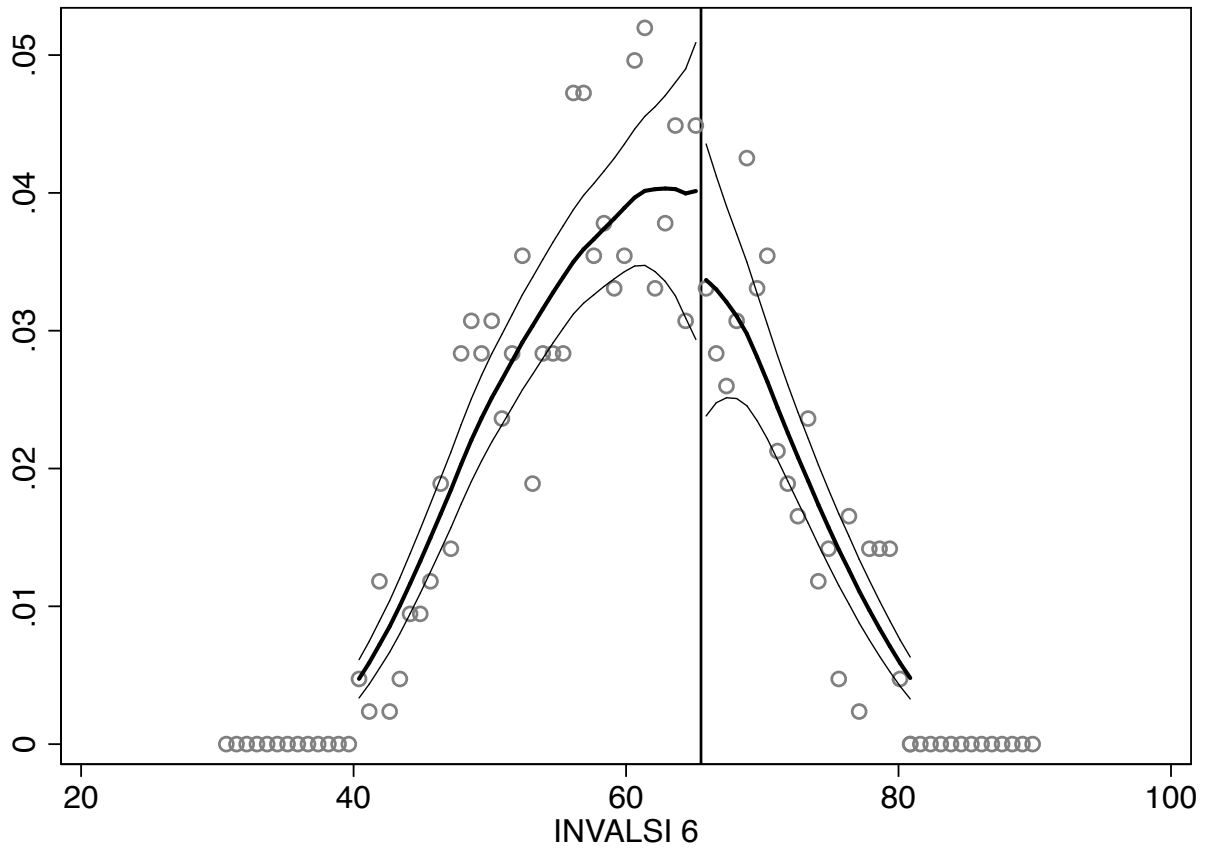


### Female



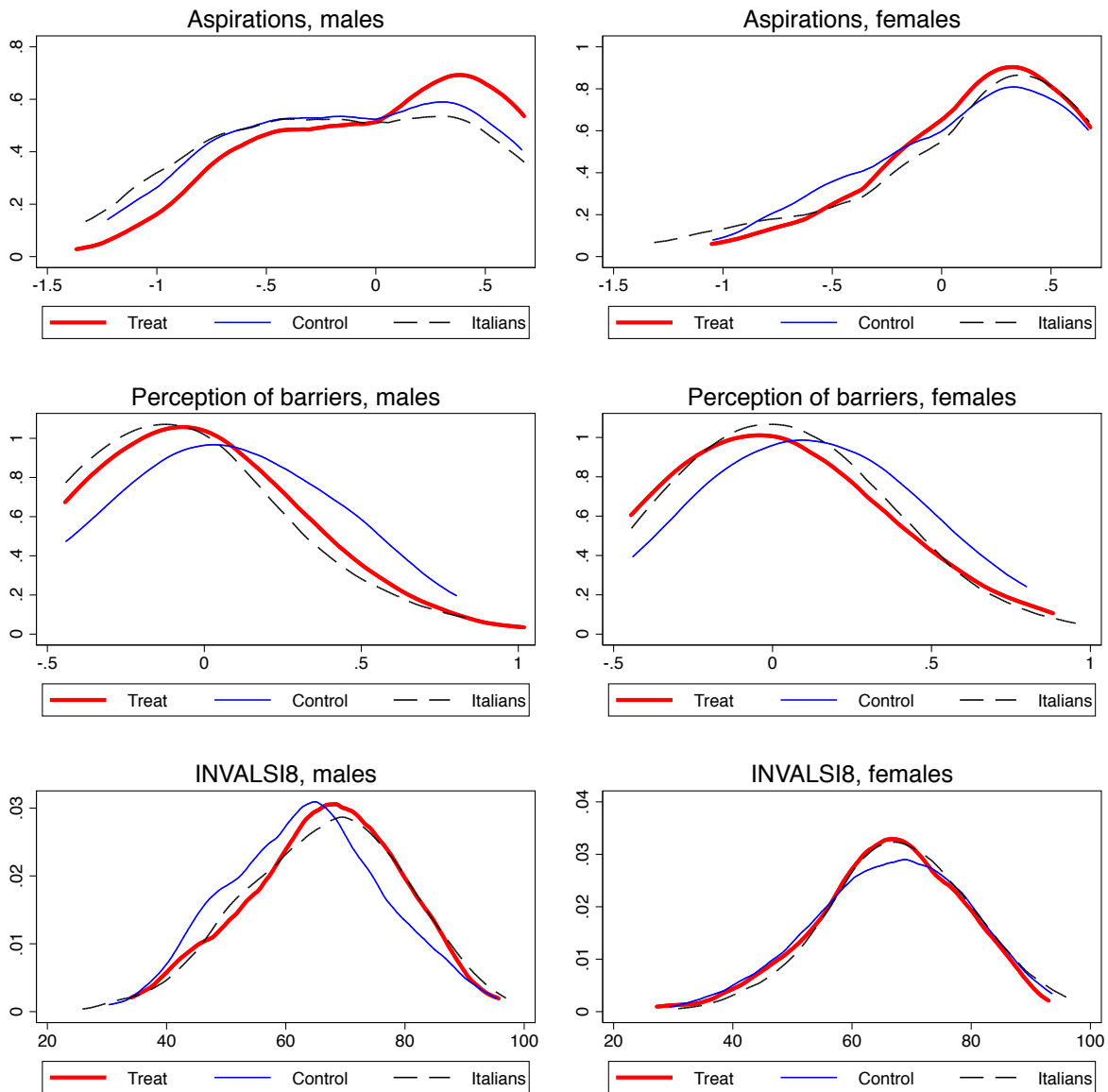
*Notes:* This figure shows compliers' characteristics ratios, i.e. the ratio of the first stage for student of a specific type (e.g., female/male) to the overall first stage. The instrument is the assignment to EOP and the endogenous variable is the probability of attending at least 75% of meetings. The figure illustrates the relative likelihood of compliers' gender, generation of immigration, tercile of INVALSI 6, and age.

Figure A.7: McCrary Test for the manipulation of the running variable in the RDD



*Notes:* This figure shows the McCrary density test to check the potential manipulation of the running variable in the regression discontinuity design. (McCrary, 2008).

Figure A.8: Distribution of cognitive and personality skills across treated, controls, and comparable native students



*Notes:* These graphs shows the distribution of aspirations, perception of barriers, and INVALSI8 across treated students, control students, and a group of Italian students that are comparable in terms of schooling ability. Specifically, we match each immigrant student with a native student obtaining exactly the same score in INVALSI6.

Table A.I: Educational and occupational outcomes 4 years after graduation, by high-school track

	All students		Males		Females	
	High Track	Low Track	High Track	Low Track	High Track	Low Track
<b>Panel A: Native students</b>						
Percentage of graduates by track	85.5	14.5	84.4	15.6	85.5	13.5
Ever enrolled into university	0.704 (0.003)	0.205 (0.004)	0.650 (0.006)	0.158 (0.006)	0.754 (0.004)	0.256 (0.007)
Dropout rate in university	0.118 (0.003)	0.306 (0.011)	0.145 (0.005)	0.353 (0.018)	0.097 (0.003)	0.274 (0.014)
Not in Employment, Education or Training (NEET)	0.199 (0.003)	0.291 (0.005)	0.189 (0.005)	0.264 (0.007)	0.208 (0.004)	0.320 (0.007)
Regretting high school choice	0.267 (0.003)	0.318 (0.005)	0.266 (0.005)	0.304 (0.007)	0.269 (0.004)	0.333 (0.008)
<b>Panel B: Immigrant students</b>						
Percentage of graduates by track	62.8	37.2	57.4	42.6	66.3	33.7
Ever enrolled into university	0.655 (0.03)	0.291 (0.022)	0.686 (0.049)	0.172 (0.027)	0.637 (0.037)	0.390 (0.032)
Dropout rate in university	0.150 (0.028)	0.257 (0.037)	0.231 (0.06)	0.307 (0.074)	0.100 (0.028)	0.238 (0.043)
Not in Employment, Education or Training (NEET)	0.264 (0.027)	0.294 (0.022)	0.237 (0.045)	0.238 (0.03)	0.280 (0.035)	0.340 (0.031)
Regretting about high school choice	0.269 (0.028)	0.331 (0.023)	0.325 (0.049)	0.290 (0.032)	0.238 (0.033)	0.365 (0.032)

*Notes:* This table shows average educational and occupational outcomes of students graduating from high school in year 2011 by gender and high school track; separate figures for native and immigrant students are presented in panel A and B, respectively. Standard errors are reported in parentheses.

Table A.II: Immigrant students' probability of choosing the high-track, controlling for socio-economic background

Dependent variable: Choosing the high-track						
	(1)	(2)	(3)	(4)	(5)	(6)
	Males			Females		
Immigrant	-0.087 (0.029)	-0.069 (0.029)	-0.059 (0.029)	-0.046 (0.029)	-0.041 (0.028)	-0.023 (0.029)
Low-educated Mother		-0.131 (0.029)	-0.119 (0.032)		-0.149 (0.026)	-0.123 (0.026)
Mid-educated Mother		-0.016 (0.018)	-0.011 (0.019)		-0.051 (0.016)	-0.042 (0.017)
Low-educated Father		-0.113 (0.029)	-0.094 (0.032)		-0.089 (0.026)	-0.069 (0.025)
Mid-educated Father		-0.018 (0.018)	-0.012 (0.018)		-0.007 (0.019)	-0.004 (0.019)
Mother bluecollar			-0.038 (0.022)			-0.051 (0.025)
Mother home/unemployed			-0.012 (0.017)			-0.026 (0.018)
Father bluecollar			-0.036 (0.022)			-0.051 (0.024)
Father home/unemployed			-0.044 (0.032)			-0.110 (0.038)
Constant	0.660 (0.023)	0.798 (0.029)	0.806 (0.029)	0.722 (0.022)	0.868 (0.024)	0.887 (0.025)
Observations	3,923	3,923	3,923	3,809	3,809	3,809
R-squared	0.220	0.255	0.264	0.214	0.240	0.247

*Notes:* This table shows how immigrant status influences the probability of choosing the high-track. The dependent variable is a dummy equal to 1 for students choosing the high-track. The main explanatory variable is a dummy equal to 1 for immigrant students. The sample includes all students in control schools. All regressions control in addition for a second degree polynomial of test score in grade 6 (INVALSI6), a dummy for first generation immigrants, and province fixed effects. Standard errors clustered by school are reported in parentheses.

Table A.III: Missing information on parents' education and local socio-economic conditions

	(1)	(2)	(3)	(4)
	Mother's education missing		Father's education missing	
<i>Panel A: Individual-level estimates</i>				
Employment rate	-3.526 (1.512)		-2.874 (1.417)	
Unemployment rate		5.160 (2.391)		4.347 (2.241)
Constant	3.428 (1.391)	-0.116 (0.146)	2.844 (1.304)	-0.054 (0.136)
Observations	1,217	1,217	1,217	1,217
R-squared	0.050	0.049	0.054	0.053
<i>Panel B: Area-level estimates, weighted by area population</i>				
Employment rate	-5.899 (1.657)		-5.151 (1.506)	
Unemployment rate		8.871 (2.646)		7.833 (2.383)
Constant	5.634 (1.526)	-0.315 (0.161)	4.965 (1.387)	-0.235 (0.144)
Observations	87	87	87	87
R-squared	0.183	0.179	0.194	0.192

*Notes:* This table shows the relationship between accuracy of information on parents' education in INVALSI registries and local socio-economic conditions. The dependent variables are binary indicators equal to one when mother's and/or father's education (columns 1-2 and 3-4, respectively) are not reported in INVALSI registries. The main explanatory variables of interest are the employment and unemployment rates across 87 Census tracts. Regressions in Panel A are estimated across individuals, whereas regressions in Panel B are estimated across Census tracts weighted by population. Province fixed effects are included in all regressions. In the individual-level regressions in Panel A, standard errors are clustered by Census tract



Table A.IV: The effect of EOP on educational choices, difference-in-discontinuities at the eligibility cutoff between treated and control schools

	(1)	(2)	(3)	(4)	(5)	(6)
	male students			female students		
EOP X top 10	0.144 (0.075)	0.186 (0.114)	0.184 (0.081)	-0.203 (0.127)	-0.202 (0.148)	-0.173 (0.125)
EOP	-0.041 (0.066)	-0.110 (0.093)		0.213 (0.124)	0.204 (0.137)	
top 10	0.151 (0.068)	-0.009 (0.107)	-0.022 (0.096)	0.359 (0.142)	0.218 (0.122)	0.261 (0.133)
rank	-0.011 (0.004)	-0.003 (0.004)	-0.002 (0.005)	-0.008 (0.006)	0.001 (0.005)	0.001 (0.007)
top 10 X rank	-0.017 (0.007)	-0.002 (0.011)	-0.002 (0.011)	-0.023 (0.007)	-0.009 (0.010)	-0.014 (0.008)
EOP X rank	0.001 (0.004)	0.007 (0.005)	0.007 (0.004)	-0.008 (0.006)	-0.007 (0.007)	-0.006 (0.006)
EOP X top 10 X rank	-0.003 (0.009)	-0.006 (0.015)	-0.008 (0.010)	0.008 (0.008)	0.009 (0.013)	0.009 (0.008)
Observations	1,320	1,290	1,318	1,274	1,254	1,270
Additional covariates	NO	YES	YES	NO	YES	YES
School FE	NO	NO	YES	NO	NO	YES
R-squared	0.097	0.157	0.255	0.144	0.205	0.286

*Notes:* This table shows the effect of EOP on educational choices exploiting the fact that only the 10 immigrant students with the highest INVALSI6 within each school were eligible for the program. The dependent variable is a dummy equal to 1 for students choosing the high track (academic or technical schools) and equal to zero otherwise. Top 10 is a dummy for students potentially eligible within each school, EOP is a dummy for schools actually participating to the program, and rank is a position in the rank. Therefore, the coefficient of top 10 estimates the average discontinuity in the probability of choosing the high track between eligible and non-eligible students near the cutoff across all schools, and the coefficient of the interaction EOP X top 10 estimates the differential discontinuity in treated schools (i.e., the "difference-in-discontinuities"). Columns (2)-(3) and (5)-(6) control in addition for a squared polynomial in INVALSI6, a dummy for first generation immigrants, and province fixed effects, and columns (3) and (6) also include school fixed effects. Standard errors clustered by rank in all specifications and, in addition, by school in columns (3) and (6) are reported in parentheses.

Table A.V: The effect of completing the questionnaire on soft skills in control schools

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Demanding High-School		Grade Retention		Std Test Score grade 8	
School Questionnaire	-0.016 (0.039)	-0.004 (0.057)	0.008 (0.021)	0.007 (0.033)	-0.061 (0.100)	-0.085 (0.125)
Female X School Questionnaire		-0.029 (0.069)		0.004 (0.036)		0.035 (0.140)
Female		0.167 (0.050)		-0.058 (0.025)		0.168 (0.111)
Constant	0.759 (0.027)	0.676 (0.042)	0.052 (0.014)	0.081 (0.023)	0.004 (0.069)	-0.081 (0.087)
Observations	620	620	620	620	552	552
R-squared	0.000	0.031	0.000	0.015	0.001	0.015

*Notes:* This table tests whether control students in schools selected for the questionnaire differ in their high-school choice, grade retention and performance in the standardized test score from students in the control schools not selected for the questionnaire. The dependent variable is a dummy equal to 1 for students choosing the high-track in columns 1 and 2, a dummy equal to 1 if students are retained in grade 7 or 8 in columns 3 and 4 and the standardized value for the test score in columns 5 and 6. Standard errors clustered by school are reported in parentheses.

Table A.VI: Principal component analysis, factor loadings

First principal component: Aspirations		
	Loadings	std. err.
Goal University	1	.
Self efficacy University	1.649	0.050
Self efficacy White collar	0.753	0.030
Self efficacy Manager	0.631	0.030
Second principal component: Perception of barriers		
	Loadings	std. err.
Barriers economic	1	.
Barriers family ideas	1.339	0.087
Barriers prejudice	0.837	0.063
Barriers family plans and marriage	1.001	0.074
Barriers self esteem	0.947	0.072

*Notes:* This table shows estimated factor loadings for the two principal components extracted from psychological measures; Satorra-Bentler robust standard errors are also presented. Measurements are categorical variables in a scale from 1 to 4. “Goal University” is the answer to the following question: *Thinking about your future, do you want to achieve an university degree?*. The measurements related to “Self efficacy” are the answers to the following questions: *Independently from your educational aim but thinking about your abilities, do you think you could get a university degree/ white collar job/ managerial job?*. The measurements related to “Barriers” are the answers to the following questions: *Do you think the following barriers could be an obstacle in the achievement of your educational aims? Economic resources/ The needs and ideas of your family/ Racial prejudice/ Family plans (children, marriage)/ Not feeling good enough.*

Table A.VII: Initial vs. working sample

	Treated	Controls	Difference	P-value	Std. Difference
Initial sample	700	751			
Fraction missing match MIUR-INVALSI	0.043	0.053	- 0.010	[0.72]	-0.049
Number of students with available MIUR-INVALSI	670	711			
Fraction dropped between INVALSI6 and start of EOP	0.109	0.128	- 0.019	[0.51]	-0.059
Final sample	597	620			

*Notes:* This table shows the sample size of treated and control students in our sample. P-value for difference in means are reported in parenthesis. The last column also reports the standardized difference between group averages.

Table A.VIII: Balance table of the nationality of treated and control students

Variable	(1) Treated	(2) Control	(3) Diff.	(4) P-value	Std. Difference
Albania	0.090	0.104	0.014	[0.554]	-0.047
Romania	0.192	0.245	0.053	[0.164]	-0.129
Morocco	0.085	0.082	-0.003	[0.890]	0.011
Philippines	0.106	0.069	-0.038	[0.155]	0.131
Peru	0.058	0.057	-0.001	[0.952]	0.004
Ecuador	0.052	0.062	0.010	[0.591]	-0.043
China	0.056	0.054	-0.003	[0.908]	0.009
Observations	620	597			

*Notes:* This table shows the most common nationality of treated and control students in our sample. P-values for difference in means are reported in square brackets. The last column also reports the standardized difference between group averages.

Table A.IX: EOP effect on track choice using an ordered probit

Dependent variable: Track choice						
	(1)	(2)	(3)	(4)	(5)	(6)
	All immigrants		Male immigrants		Female immigrants	
Top Academic	0.029 (0.026)	0.025 (0.021)	0.062 (0.030)	0.053 (0.026)	-0.011 (0.036)	-0.005 (0.030)
Other academic/Technical	-0.002 (0.002)	-0.001 (0.001)	0.009 (0.007)	0.009 (0.006)	0.003 (0.009)	0.001 (0.007)
Vocational	-0.028 (0.024)	-0.024 (0.020)	-0.071 (0.035)	-0.062 (0.031)	0.008 (0.027)	0.004 (0.023)
Mean Control Top Academic	0.245	0.245	0.195	0.195	0.294	0.294
Mean Control Other	0.507	0.507	0.482	0.482	0.530	0.530
Mean Control Vocational	0.248	0.248	0.322	0.322	0.176	0.176
Observations	1,217	1,217	601	601	616	616
Controls	No	Yes	No	Yes	No	Yes

*Notes:* This table shows marginal effects of an ordered probit, considering tracks split in three groups: top academic (classic and scientific), Other academic/Technical, Vocational. Controls include the second degree polynomial of test score in grade 6 (INVALSI6), a dummy for first generation immigrants, and province fixed effects.

Table A.X: The effect of EOP on educational choices, heterogeneity

Dependent variable: Choosing the high-track			
	(1)	(2)	(3)
<i>Panel A: Heterogeneity by gender and mother education</i>			
Subgroup	EOP Coeff.	P-value	MHT P-value
Boys High Edu Mother	0.017	0.725	0.725
Boys Low Edu Mother	0.160	0.046	0.209
Boys Missing Edu Mother	0.119	0.058	0.211
Girls High Edu Mother	0.050	0.169	0.423
Girls Low Edu Mother	0.185	0.009	0.051
Girls Missing Edu Mother	0.040	0.502	0.750
<i>Panel B: Heterogeneity by gender and terciles of initial test score</i>			
Subgroup	EOP Coeff.	P-value	MHT P-value
Boys Top INVALSI6	0.106	0.052	0.272
Boys Middle INVALSI6	0.067	0.267	0.468
Boys Bottom INVALSI6	0.088	0.189	0.564
Girls Top INVALSI6	0.016	0.648	0.648
Girls Middle INVALSI6	0.085	0.063	0.272
Girls Bottom INVALSI6	0.075	0.250	0.582
<i>Panel C: Heterogeneity by gender and place of birth</i>			
Subgroup	EOP Coeff.	P-value	MHT P-value
Boys EU	0.103	0.058	0.580
Boys Not EU	0.085	0.049	0.179
Girls EU	0.058	0.103	0.456
Girls Not EU	0.004	0.895	0.895

*Notes:* This table shows the heterogeneity of the effect of EOP on immigrant students' educational choices the end of middle school. The dependent variable is a dummy equal to 1 for students choosing the high-track (academic or technical schools) and equal to zero otherwise. EOP is a dummy equal to 1 for students in middle schools assigned to the treatment group and equal to zero for schools assigned to the control group. Highly educated mother is a dummy equal to 1 for students' whose mother has at least a high-school diploma. EU is a dummy equal to 1 for immigrants from EU-member countries.

Table A.XI: RDD effects

	(1)	(2)	(3)	(4)
<i>Panel A:</i>				
Dep. var.	No. CALP Meetings	No. Counseling Meetings	Choice High Track	Grade Retention
RDD Estimate	-5.418 (2.575)	1.006 (0.934)	-0.024 (0.0758)	-0.023 (0.0250)
Observations	565	565	565	565
<i>Panel B:</i>				
Dep. var.	Aspirations	Perception of Barriers	INVALSI8	Teachers' recom.
RDD Estimate	-0.143 (0.127)	-0.020 (0.0877)	-1.900 (2.775)	-0.002 (0.111)
Observations	404	404	512	565

*Notes:* This table shows the RDD effect of CALP on several outcomes, as reported also in Figure 7. Squared Order Local Polynomial.

Table A.XII: Treatment effect on soft skills (by survey question)

ITT on	Coefficient	p-value	p-value FWER
Group 1: aspirations			
Goal University	0.098	0.012	0.072
Self efficacy University	0.174	0.008	0.067
Self efficacy Whitecollar	0.120	0.021	0.078
Self efficacy Manager	0.056	0.342	0.035
Group 2: perception of environmental barriers			
Barriers economic	-0.101	0.041	0.019
Barriers family ideas	-0.084	0.068	0.224
Barriers prejudice	0.033	0.513	0.536
Barriers family formation and marriage	-0.087	0.076	0.254
Barriers self esteem	-0.233	0.000	0.001

*Notes:* Robust standard errors clustered at school level. All regressions include generation of immigration, province and squared test score. P-values are adjusted for multiple hypothesis testing using the free step-down resampling method (Westfall and Young, 1993) to control the family-wise error rate (FWER). Measurements are categorical variables in a scale from 1 to 4. “Goal University” is the answer to the following question: *Thinking about your future, do you want to achieve an university degree?*. The measurements related to “Self efficacy” are the answers to the following questions: *Independently from your educational aim but thinking about your abilities, do you think you could get a university degree/ white collar job/ managerial job?*. The measurements related to “Barriers” are the answers to the following questions: *Do you think the following barriers could be an obstacle in the achievement of your educational aims? Economic resources/ The needs and ideas of your family/ Racial prejudice/ Family plans (children, marriage)/ Not feeling good enough.*

Table A.XIII: Specification test, Males

Outcome: Choosing the high-track	(1)		(2)		(3)	
Mediating factors: $H_0 : \alpha_1^j = \alpha_0^j$	Test Statistic	p-value	Test Statistic	p-value	Test Statistic	p-value
Aspirations	2.24	[0.14]	1.81	[0.18]	2.84	[0.10]
Barriers	1.30	[0.26]	1.08	[0.30]	0.24	[0.63]
INVALSI8			0.02	[0.89]	0.36	[0.55]
Teachers' recommendation					0.36	[0.55]
Controls: $H_0 : \beta_1^i = \beta_0^i$	Test Statistic	p-value	Test Statistic	p-value	Test Statistic	p-value
INVALSI6	0.52	[0.47]	0.31	[0.58]	0.78	[0.38]
INVALSI6 sq.	2.49	[0.12]	2.22	[0.14]	1.93	[0.17]
First generation immigrant	0.91	[0.34]	0.86	[0.39]	0.54	[0.46]
Prov BS	2.35	[0.13]	1.76	[0.19]	0.21	[0.65]
Prov GE	1.20	[0.28]	1.01	[0.32]	1.74	[0.19]
Prov MI	0.08	[0.77]	0.00	[0.96]	0.27	[0.61]
Prov PD	0.36	[0.55]	0.11	[0.75]	0.03	[0.86]
Prov TO	0.95	[0.33]	0.96	[0.33]	0.10	[0.76]
F-test	1.37	[0.21]	1.28	[0.26]	1.35	[0.23]

*Notes:* The first panel tests whether the treatment group regression coefficients in equation 3 are the same as the control group coefficients:  $H_0 : \alpha_1^j = \alpha_0^j$ , for each potential channel  $\theta$ . The second panel tests whether the treatment group regression coefficients are the same as the control group coefficients:  $H_0 : \beta_1^i = \beta_0^i$ , for each potential control variable  $\mathbf{X}$ . In column (1), we consider only two mediating factors, i.e. aspirations and barriers, while in column (2) we include also the achievement test scores and in column (3) teachers' track recommendation. All specifications control in addition for a squared polynomial in INVALSI6, a dummy equal to 1 for first generation immigrants, and province fixed effects.

Table A.XIV: Decomposition of the effect of EOP on high-school choice, male students (Gelbach, 2016)

	(1)		(2)		(3)	
	Explained	p-value	Explained	p-value	Explained	p-value
Aspiration	0.0354	[0.034]	0.0291	[0.042]	0.0264	[0.046]
Barriers	0.0046	[0.602]	0.0067	[0.446]	0.0042	[0.620]
Cognitive skills			0.0178	[0.085]	0.0105	[0.190]
Teachers' recommendation					0.0385	[0.011]
Total explained	0.0400	[0.032]	0.0537	[0.010]	0.0796	[0.001]
EOP effect on choosing high-track	0.091		0.091		0.091	

*Notes:* This table decomposes the effect of EOP between changes in personality skills (aspirations and perception of barriers), increased schooling achievement (as measured by INVALSI8) and teachers' recommendations. The decomposition follows the method devised by Gelbach (2016). All specifications control in addition for a squared polynomial in INVALSI6, a dummy equal to 1 for first generation immigrants, and province fixed effects.



Table A.XV: Cost Benefit Analysis

	Scenario 1	Scenario 2	Scenario 3
<b>Parameters</b>			
Discount rate	3%	3%	3%
Tax rate	28%	28%	28%
Higher salary per month (euros)	500	500	650
Lower unemployment probability	4%	4%	6%
Unemployment insurance benefit per month (euros)	1000	1000	1300
Number of beneficiaries	60	125	125
<b>Costs and Benefits</b>			
Total costs (thousand euros)	2,177	2,177	2,177
Higher taxes on wage (thousand euros)	3,344	7,006	9,108
Lower unemployment insurance (thousand euros)	955	2,002	3,904
<b>Internal Rate of Return</b>			
	2.8%	6.6%	8.8%

*Notes:* Although EOP program has potentially strong effects on health and on reduction of crime rates, we present conservative estimates focusing our cost-benefit analysis only on social benefits coming from higher income taxes and public savings on unemployment insurance. In the first scenario, we consider potential benefits only on 10% of students directly treated by EOP. In the second scenario, keeping all other assumptions constant, we consider also the additional spillovers on 5% of classmates of treated students (close to the share who did not fail the school year or decided to attend a more demanding track compared to classmates of control students). In the last scenario, we slightly reduce the unemployment probability, we slightly increase the expected average higher salary per month and the expected unemployment insurance benefit. We use the discount rate of 3% as in the simulation in Heckman et al. (2010) and a tax rate close to the current one for the income bracket 15,000-28,000 euros.

## B Average Treatment-on-the-treated effect

While in the main text we focus on intention-to-treat (ITT) effects, in this section we estimate the average treatment-on-the-treated (ATT). Since there is one-sided non-compliance with treatment assignment, the ATT effects on the subset of compliers are larger than the ITT. Appendix Figure B.1 shows that the pattern of meetings attendance is quite heterogeneous, with more than 40 percent of immigrant boys and girls attending at least 87.5 percent of the meetings, another 20 percent attending between 75 and 87.5 percent of the meetings, and the remaining fractions attending less. Interestingly, about 15 percent of the students who were assigned to treatment ended up attending less than 12.5 percent of the meetings. Given this heterogeneity, there is no unambiguous way of defining treatment status. For this reason, in Appendix Table B.I we experiment with three alternative definitions.

In Panel A, we classify as treated all students attending at least one meeting (85 percent of the sample). In Panel B we restrict the definition to students attending at least 75 percent of the meetings, in accordance with the program guidelines discussed in Section 3. When adopting these definitions, the ATT effects on males range between a 9.4 to 12.5 percentage point increase in enrollment in the high track, and a 4.3 to 5.7 percentage point decrease in grade retention.

In Appendix Figure A.6 we characterize compliers with treatment assignment, defining the treatment as attending at least 75 percent of the meetings, by the ratio of the first stage effect within specific sub-samples to the overall first stage (Angrist et al., 2016). Compliers are slightly more likely to be female, equally likely to be first and second generation immigrants, and more likely to be in the right grade ('Not late') given their age. The bottom panels of Figure A.6 show that while female compliers are more likely to be from the top part of the initial ability distribution, male compliers are more likely to be from the bottom part, thus more in need of support.

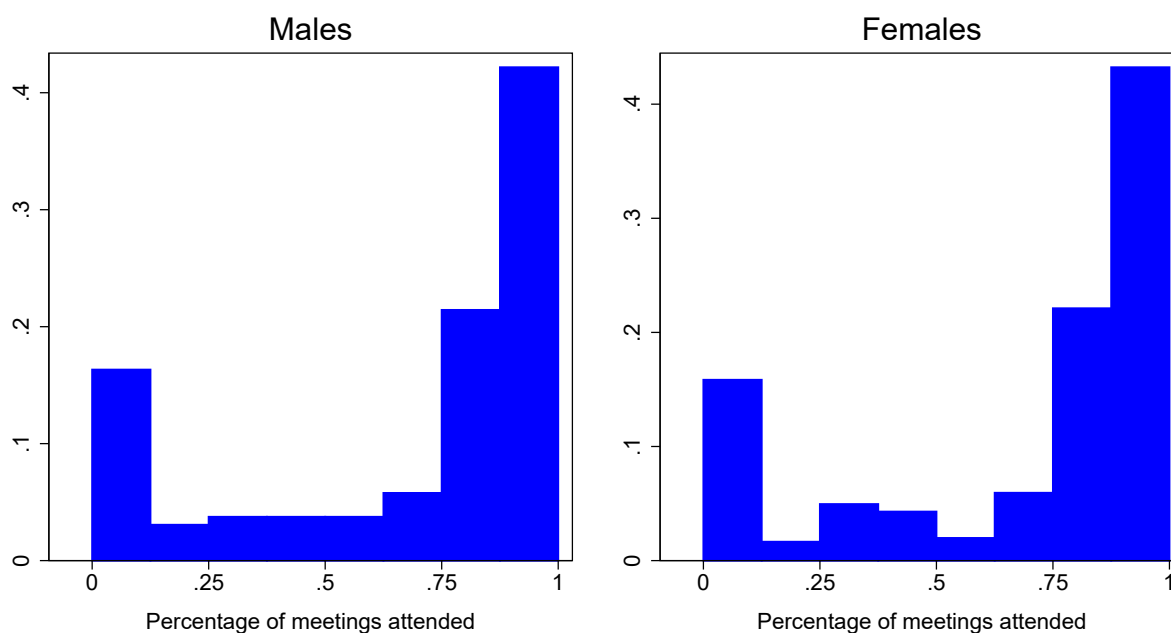
In Panel C of Table B.I we measure treatment 'intensity' by the fraction of meetings attended. The corresponding ATT estimate suggests that one standard deviation increase in the number of meetings attended increases enrollment into the high track by 4.2 percentage points and reduces grade retention by 1.9 percentage points for males. All three approaches in Table B.I recover the ATT effect under strong (and untestable) assumptions about the relationship between number of meetings attended and treatment intensity. For this reason, in the main body of the paper we focus on the ITT.

Table B.I: Effects of EOP, average treatment-on-the-treated (ATT)

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.:	Choosing the high-track			Grade retention		
	All	Males	Females	All	Males	Females
<i>Panel A: Treatment = 1 if attended at least one meeting</i>						
ATT	0.051 (0.028)	0.094 (0.041)	0.010 (0.036)	-0.015 (0.018)	-0.043 (0.024)	0.011 (0.022)
Constant	0.682 (0.038)	0.648 (0.049)	0.720 (0.049)	0.073 (0.028)	0.098 (0.028)	0.047 (0.043)
<i>Panel B: Treatment = 1 if attended at least 75% of meetings</i>						
ATT	0.067 (0.037)	0.125 (0.054)	0.013 (0.047)	-0.020 (0.024)	-0.057 (0.032)	0.014 (0.029)
Constant	0.679 (0.038)	0.641 (0.050)	0.720 (0.049)	0.074 (0.027)	0.101 (0.028)	0.047 (0.043)
<i>Panel C: Treatment = fraction of meetings attended</i>						
ATT	0.064 (0.036)	0.119 (0.052)	0.013 (0.045)	-0.019 (0.023)	-0.055 (0.031)	0.014 (0.028)
Constant	0.681 (0.038)	0.645 (0.049)	0.720 (0.049)	0.074 (0.027)	0.100 (0.028)	0.048 (0.043)
Observations	1.217	601	616	1.217	601	616

*Notes:* This table shows the average-treatment-on-the-treated effect of EOP on immigrant students' educational choices (columns 1-3) and grade retention during middle school (columns 4-6). The ATT is computed as the ratio of the reduced form effect of EOP on such outcomes and the first stage effect on three alternative measures of compliance with treatment assignment: attending at least 1 meeting (Panel A), attending at least 75% of meetings (Panel B), and fraction of meetings attended (Panel C). All specifications control for a squared polynomial in INVALSI6, a dummy equal to 1 for first generation immigrants, and province fixed effects. Standard errors clustered by school are reported in parentheses.

Figure B.1: Meetings attendance of immigrant students assigned to EOP



## C Heterogeneous treatment effects using Causal Forest

The purpose of causal forest approach is not to test whether a particular characteristic (e.g., gender) is associated with significantly different treatment effects, but to predict the heterogeneity in the causal treatment effects. The method adapts the classification and regression tree (CART) methodology from the recent literature on supervised machine learning to the problem of predicting treatment effects rather than outcomes. More precisely, the causal forest allows to estimate the Conditional Average Treatment Effect (CATE), defined as  $E(Y_{1i} - Y_{0i} | X_i = x)$  where  $Y$  is the outcome of interest and  $X$  is a vector of observable baseline characteristics.

We build upon the recent literature of machine learning techniques in the context of randomized control trials and heterogeneous treatment effects (Davis and Heller, 2017; Bertrand et al., 2017). We use causal forest algorithm (Athey and Imbens, 2016; Wager and Athey, 2018; Athey et al., 2019) and the following procedure. First, we train a causal forest by building 100,000 trees and setting the minimum number of treatment and control observations allowed in a leaf to the default value (5).<sup>32</sup> We use the “honest” approach (Athey and Imbens, 2016; Wager and Athey, 2018): we

<sup>32</sup> We use the `causal_forest` command in R of the package `grf` (generalized random forest). As suggested by Athey et al. (2019), the only substantial difference from the method propose in Wager and Athey (2018) is that the split

split the training sample in two parts, we use half of the observations for growing the tree and half of the observations for estimating the treatment effect within each leaf of the tree. We include in the causal forest the following baseline characteristics: gender (dummy variable), squared polynomial of baseline test score (continuous variable), generation of immigration (dummy variable), school province (categorical variable with five values: Milan, Turin, Brescia, Padua, Genoa), parents' education (categorical variable with four values: university, high-school diploma, less than high-school diploma, missing) and occupation (categorical variable with five values: high-level occupation, working class, unemployed, housemaker, missing), area of citizenship (categorical variable with four values: Latin America, East Europe, Africa, Asia), and, crucially given the design of our experiment, we include the clusters at school level. Second, we calculate the out-of-bag predicted CATE and its variance estimates.<sup>33</sup>

We use the predictions on the expected treatment effect on the high-school track choice for each individual, given the covariates, to investigate the treatment heterogeneity in our data. We divide the sample in two groups, considering students in the top 50% and bottom 50% of the predictions. Table C.I reports the balancing test for the Conditional Average Treatment Effect (CATE) and provides the p-value adjusted for multiple hypothesis testing. Overall the results on gender differences are consistent with the main analysis: girls are overrepresented among students with lower CATE. However, additional interesting results emerge.

Students whose parents have white collar jobs or high levels of education are more likely to have a low predicted CATE. Furthermore, interestingly, more than 50% of the students with high predicted CATE have missing values for their parental education and occupation, while this share is lower than 20% for students with high predicted CATE. Parental background is asked to parents from schools and then submitted together with the test scores INVALSI. Targeting the children of parents less responsive to school requests and therefore less involved in school activities may increase the effectiveness of EOP. The average test score in grade 6 (INVALSI6) is slightly lower for students with high predicted CATE on track choice (0.263 vs. 0.352 standard deviations), although the difference is not statistically significant when the p-value is adjusted for multiple hypothesis testing. Figure 6 shows the non-linearities in the CATE considering the deciles of the test score in grade 6 (INVALSI6), gender, and mother education. As suggested by Table C.I, on average boys,

---

id done using a gradient-based loss criterion instead of the exact loss criterion.

<sup>33</sup> Davis and Heller (2017) show that "honest" approach may lead to overfitting if the CATE is assigned to all observations in the sample, including those used to construct the tree and they suggest to obtain out of sample predictions by further splitting the sample and running the causal forest of 20% of the observations. Given our small sample and the purpose of our exercise, we provide predictions for an observation in the original dataset (at  $X_i$ ) using only trees that did not use the  $i$ -th training example. These predictions are not prone to overfitting, as each prediction is only made by learners that did not use the observation for training. In our predictions the 'excess.error' is negligible, with a mean value of 4.21e-07.

students with lower or missing levels of mother education, and students in the central part of the ability distribution<sup>34</sup> benefit the most from participation in EOP. Targeting only boys would have missed a substantially positive impact on these girls.

Finally, Table C.I reports also the descriptive statistics for the school province and the citizenship. There are no substantial differences across provinces in the CATE, suggesting that different psychologists were equally effective in implementing the treatment. However, students from East Europe are more likely to benefit from the intervention (they are 49% of the High Predicted CATE and 34% of the Low Predicted CATE group) and those from Latin America are less likely to benefit from it (they are 20% of the High Predicted CATE and 32% of the Low Predicted CATE group).

---

<sup>34</sup>The deciles are defined among the 1217 immigrant students in the treatment and control group.

Table C.I: Conditional Average Treatment Effect (CATE): Choice of High-Track

Variable	(1) High Predicted CATE	(2) Low Predicted CATE	(3) Diff.	(4) MHT p-value
ITT High Track	0.053	0.031	-0.023	
Female	0.447	0.566	0.119	0.001
First Generation	0.586	0.523	-0.063	0.210
INVALSI 6	0.263	0.352	0.089	0.177
Sq. INVALSI 6	0.364	0.742	0.378	0.001
Older age	0.302	0.212	-0.090	0.001
<b>Mother occupation:</b>				
White Collar	0.049	0.199	0.150	0.001
Working class	0.151	0.322	0.171	0.001
Homemaker	0.213	0.304	0.091	0.005
Unemployed	0.062	0.048	-0.015	0.588
Missing	0.524	0.127	-0.397	0.001
<b>Father occupation:</b>				
Whitecollar	0.122	0.275	0.153	0.001
Working class	0.232	0.482	0.250	0.001
Homemaker	0.013	0.012	-0.002	0.805
Unemployed	0.054	0.071	0.017	0.648
Missing	0.580	0.161	-0.418	0.001
<b>Mother education:</b>				
University	0.053	0.186	0.133	0.001
High-school	0.169	0.426	0.257	0.001
Less than High-school	0.205	0.243	0.038	0.530
Missing	0.573	0.145	-0.428	0.001
<b>Father education:</b>				
University	0.043	0.156	0.114	0.001
High-school	0.149	0.442	0.293	0.001
Less than High-school	0.171	0.224	0.053	0.168
Missing	0.637	0.178	-0.459	0.001
<b>Province:</b>				
PD	0.079	0.031	-0.048	0.001
BS	0.164	0.194	0.030	0.674
MI	0.461	0.531	0.070	0.153
TO	0.227	0.179	-0.047	0.257
GE	0.069	0.064	-0.005	0.937
<b>Citizenship:</b>				
Latin America	0.195	0.317	0.122	0.001
Africa	0.166	0.224	0.058	0.104
Asia	0.148	0.123	-0.024	0.684
East Europe	0.489	0.336	-0.154	0.001
Observations	609	608		

Notes: The Table reports the descriptive statistics of students in the top 50% (column 1) and bottom 50% (column 2) of the predicted Conditional Average Treatment Effect (CATE) on the choice of a demanding high-school. The CATE is computed following the procedure explained in Appendix C. Column 3 reports the difference between column 2 and 1. Column 4 shows the p-value of the t-test adjusted for multiple hypothesis testing.

## D Methodology for mediation analysis

Following Heckman et al. (2013), we decompose the treatment effect on educational choices into experimentally induced changes in the mediating factors listed in Table IV and changes in other (unmeasured) factors. Assume the following linear model for the potential outcome when randomized into the treated ( $d = 1$ ) and into the control group ( $d = 0$ ):

$$Y_d = \tau_d + \sum_{j \in J} \alpha_d^j \theta_d^j + \beta_d \mathbf{X} + \varepsilon_d, \quad d \in \{0, 1\}, \quad (2)$$

where  $Y$  is a dummy for choosing the high track,  $\tau$  is the intercept,  $\Theta = (\theta^j : j \in J)$  is the set of observed mediating factors (cognitive skills, personality traits, and teachers' recommendation),  $\mathbf{X}$  is a vector of pre-program variables unaffected by the treatment (initial test score INVALSI6, generation of immigration, and province fixed effects), and  $\varepsilon_d$  is an error term. With the exception of  $\mathbf{X}$ , all variables and coefficients in equation (2) are allowed to depend on treatment assignment. In particular,  $\tau_d$  captures the effect of experimentally induced changes in other (unobserved) determinants of  $Y$ , in addition to the observed mediating factors in  $\Theta$ .

Separately identifying the components of the treatment effect attributable to  $\tau_d$  and  $\Theta$ , respectively, requires further assumptions, as experimental variation allows us to consistently estimate the effects of EOP on measured factors and final educational decisions, but not the relationship between the former and the latter. Heckman et al. (2013) assume independence of observed and unobserved factors in the no-treatment state, conditional on the vector  $\mathbf{X}$  of pre-treatment characteristics. Maintaining this assumption and imposing the additional testable restriction that coefficients do not vary with treatment assignment (respectively,  $\alpha_d^j = \alpha$  for all  $j$  and  $\beta_d = \beta$ ) allows us to decompose the effect of EOP as:

$$E(Y_1 - Y_0) = \sum_{j \in J} \alpha^j E(\theta_1^j - \theta_0^j) + (\tau_1 - \tau_0), \quad (3)$$

where  $E(Y_1 - Y_0)$  is the average treatment effect;  $E(\theta_1^j - \theta_0^j)$  is the average change induced in the  $j$ -th observed factor, and  $\alpha^j$  is the associated effect on educational choices; finally,  $(\tau_1 - \tau_0)$  is the effect due to other unmeasured factors. In Appendix Table A.XIII we test and do not reject the structural invariance assumptions on  $\alpha^j$  for all  $j$  and  $\beta$ .