

The Labor Hub

Job Market Paper Series No. 3

Advisor: Marco Bertoni

PEER BACKGROUNDS MATTER: CLASSMATES' PARENTS AND HIGH SCHOOL TRACK CHOICE

Filippo Da Re

Peer Backgrounds Matter: Classmates' Parents and High School Track Choice

Filippo Da Re, University of Padova

Abstract.

This paper investigates whether parental occupations of middle school classmates causally influence high school track choice—a critical margin for intergenerational mobility. Using administrative data from Italy, I compare students whose parents hold the same occupation and exploit within-school, within-cohort variation in classroom composition arising from quasi-random class assignment. A one-standard-deviation increase in exposure to classmates with high prestige parental occupations raises academic track enrollment by approximately 2 percentage points, while equivalent exposure to classmates with low prestige parental occupations reduces it by a smaller magnitude. I then disentangle this effect from the role of classmates' academic achievement by instrumenting the latter using partially overlapping groups of classmates. The estimated effect of classmates' parental occupations persists, and I find evidence that classmates' achievement does not mediate it, consistent with non-academic channels such as information transmission, aspiration formation, and social-status considerations. These findings suggest that classroom composition policies can shape educational sorting beyond their effects on student achievement.

Keywords: school tracking, school choice, peers, occupational prestige.

JEL codes: I24; J24.

Acknowledgments: I am grateful to Marco Bertoni, Giorgio Brunello, Sarah Cattani, Colin Green, Rigissa Megalokonomou, Laura Pagani, Enrico Rettore, Kjell Salvanes, Marco Tonello, and the audiences at seminars in Padova, Venice, Trondheim, Bergen, Verona and at the 10th LEER, 40th AIEL and 23rd Brucchi Luchino, 39th ESPE, 38th EALE conferences for comments and suggestions. Please address correspondence to Filippo Da Re, Department of Economics and Management “Marco Fanno”, Via del Santo 33, 35123. Padova, Italy. E-mail address: filippo.dare.3@phd.unipd.it.

1. *Introduction*

High school track choice is a consequential educational decision that shapes students' subsequent access to higher education and, ultimately, their labor market trajectories. Because students from different family backgrounds make systematically different educational choices, track selection is also a key margin through which socioeconomic inequality is transmitted across generations. Crucially, educational choices are shaped by factors beyond academic ability. Prior work identifies several influences that operate alongside measured ability, including family income (Andrews et al., 2020), parental background (Biewen & Tapalaga, 2017), peers (Bertoni et al., 2020), and teacher recommendations (Carlana, 2019). Yet, among these influences, the role of classmates' parental occupations remains largely unexplored. Understanding whether and how the social background of classmates shapes educational choices is therefore central to assessing the role of schools in intergenerational mobility, particularly in Italy, which ranks among the countries with the lowest mobility across generations (Mocetti, 2007; Acciari, Polo and Violante, 2022).

In this paper, I examine whether exposure during middle school to classmates' parental occupations causally shapes high school track choice in Italy. Parental occupations represent a visible and socially salient dimension of family background that may influence students' expectations about future educational and career paths. Through daily interactions with classmates, students are indirectly exposed to the occupational environments of their families. These exposures may affect educational choices by providing information about career trajectories associated with academic education, shaping aspirations, or signaling the social prestige attached to certain occupations.

Focusing on parental occupations allows me to isolate a dimension of students background that is particularly relevant for educational decisions and closely linked to the occupational returns of academic tracks. While occupations are correlated with other aspects of socioeconomic background, such as parental education and income, they also capture the types of careers that students observe in their social environment and may therefore influence educational choices through channels beyond academic ability.

Italy provides an ideal setting to investigate these questions. At the end of lower-secondary school (around age 14), students choose among distinct high school tracks: vocational and technical programs, which emphasize work-based skills, and the academic track, which offers a more general, theoretical education. Choice is subject to minimal formal restrictions and only limited ability-based screening (see Section 2). Students from more affluent backgrounds are disproportionately enrolled in academic streams, raising long-standing concerns about equality of opportunity (Dustmann, 2004). Descriptive evidence from Italy shows that academic track graduates have better cognitive skills, earn higher wages, and work fewer hours than graduates of other tracks (Agarwal et al., 2021). Students in academic high schools are 50 percent more likely to pursue higher education, with 90 percent of them reporting plans to attend college (Almadiploma, 2024). In contrast, students in vocational and technical programs have far lower aspirations to attend college (Figure 1). This may translate into large wage gap, as college graduates earn an average annual earnings premium of €10,576, or approximately 37.5 log points, compared to non-graduates¹. These pathways are also characterized by significant socioeconomic segregation. As Figure 2 shows, students from low prestige parental occupational backgrounds are significantly underrepresented in academic tracks. Nevertheless, isolating the causal effect of track choice is challenging because quasi-experimental variation is scarce².

I use administrative data from INVALSI covering two cohorts of grade 8 students, linked to their grade 10 enrollment. I define classmates as the relevant peer group throughout the paper. I compare students with the same parental occupation profile who are assigned to different classes within the same school-year cohort. This design exploits quasi-random variation in classroom composition generated by school rules that form classes similar along observable dimensions. Classes are then kept together across all three grades, so students remain with the same classmates throughout middle school. I measure classmates' social background using the class shares of students with low, middle, and high prestige parental occupations, assigning each student the more prestigious occupation of their two parents. I

¹ These estimates use data from the Survey on Household Income and Wealth (SHIW) from the Bank of Italy and they are displayed in Appendix A, Table A1.

² Few notable exceptions are Anelli (2020) and Anelli and Peri (2019).

document substantial within-school-cohort variation in these shares and provide balance and randomization tests consistent with quasi-random assignment across classes.

I find sizable effects of classmates' parental occupational prestige on academic track enrollment. A one standard deviation increase in the share of classmates with high prestige parental occupations raises the probability of enrolling in the academic track by about 2 percentage points, while a corresponding increase in low prestige exposure reduces enrollment by a smaller amount. These effects are economically meaningful relative to baseline enrollment rates and align with magnitudes reported in related peer effects studies.

A key empirical challenge is that classmates' parental occupations are correlated with their academic achievement, so the estimated effect may partly capture achievement spillovers. To separate these channels, I conduct a mediation analysis that treats classmates' achievement as an endogenous mediator and instruments it using partially overlapping peer groups (De Giorgi et al., 2010). Specifically, I use the grade 5 classmates of a student's grade 8 classmates who were never that student's own classmates. Building on Cattan et al. (2025), this approach recovers exogenous variation in classmates' achievement and allows me to estimate the effect of classmates' parental occupations conditional on it. Conditional on classmates' achievement, the effect of classmates with high prestige parental occupations remains positive and statistically significant. The indirect effect through classmates' achievement is negative, but it is driven by a coefficient that is imprecisely estimated and statistically insignificant in the instrumental-variable specification. This supports the interpretation that classmates' parental occupations affect track choice mainly through channels other than achievement, such as information transmission, aspiration formation, and social status considerations.

This paper contributes to several related literatures. First, it advances research on peer effects in educational track choice by showing that classmates' parental occupations shape decisions at an earlier and consequential margin—high school track selection—than previously documented. Existing work focuses primarily on university outcomes: Bertoni et al. (2020) show that Danish students exposed to peers with higher parental education are more likely to pursue STEM degrees, while Cattan et al. (2025) find that elite peers reinforce attainment gaps partly through GPA effects. I provide new evidence that peer background

influences educational sorting before tertiary education, at a stage where social influences may have lasting, compounding consequences.

Second, the paper distinguishes classmates' background from classmates' ability using an identification strategy that combines within-school, within-cohort comparisons with instrumentation of peer ability based on partially overlapping peer groups. This complements prior work on peer socioeconomic background and attainment, including Bifulco et al. (2011), Black et al. (2013), Fruehwirth and Gagate-Miranda (2019) and Chung and Zou (2023), by testing directly whether the effect of peer background is attributable to peer academic performance.

Third, it contributes to the emerging literature on peers' parental occupations as a distinct driver of educational and labor market outcomes. Porreca et al. (2026) find that exposure to peers with unemployed parents reduces school completion in Italy; Mertz et al. (2025) show that peers with entrepreneurial parents increase entrepreneurship in Denmark; and Montonen and Solomon (2024) document that exposure to same-gender peers with white-collar parents raises the probability of entering a white-collar occupation in Finland. I complement these findings by examining how peer occupational exposure shapes high school track choice and by refining the definition of peers as classmates – the group students interact with most intensively – rather than school grade-mates.

Beyond the average effect, the results vary in ways consistent with two channels. First, status-linked aspirations and information about returns to demanding curricula predict effects concentrated in the most selective academic streams, scientific and classical, and absent in the remaining academic curricula. Second, if peer composition partly reflects broader local economic conditions, its influence should be amplified where those conditions most shape opportunity: exposure to low prestige peers has larger negative effects in low-mobility provinces. Exposure to high prestige peers, by contrast, benefits students across the socioeconomic spectrum, and most strongly those from the least advantaged families, pointing to peer influence as a channel for reducing school stratification. I find no statistical evidence that immigrant students respond differently to peer background than their native classmates, although they enroll in the academic track at substantially lower rates irrespective of the peer environment they face.

The findings have direct public policy relevance. The most immediate policy implication leverage peers' assignment. In my setting, this would entail reallocating high prestige peers from high prestige students to low prestige students. Although high prestige students also benefit from exposure to high prestige peers, the estimated effect is substantially smaller than that for low prestige students. Such a reallocation therefore need not be Pareto improving, and its desirability ultimately depends on the social welfare function adopted by the policymaker. Beyond the allocation of peers, my findings point to several other policy avenues. Because peer composition shapes access to academic tracks, policies that alter social exposure or mitigate negative peer environments—through targeted mentoring, information provision, or track-choice counselling—could improve equality of opportunity at relatively low cost, especially when directed towards students and places most at risk.

The paper unfolds as follows. In Section 2, I describe the Italian educational system. Section 3 describes the data. Section 4 details the empirical strategy. Section 5 presents the baseline results followed by robustness checks. Section 6 presents the mediation analysis. Section 7 concludes.

2. Institutional setting

Several features make Italy a compelling setting for studying peer effects in high school track choice. Education is tuition-free and publicly provided, and the track decision is taken while school attendance is still compulsory, which minimizes selection through early dropout. The curriculum is established centrally, and students are not formally sorted by ability³. The decision is also taken relatively late, at age 14, giving students greater autonomy over their educational paths.

Compulsory schooling begins at age 6 and continues until age 16. The first cycle covers primary and lower secondary education (grades 1–8) under a national curriculum common to all students. Primary school lasts five years (grades 1–5, ISCED 1), followed by three years of middle school (grades 6–8, ISCED 2). Upon completing middle school students sit a national exam, but its score does not determine track placement. Teachers issue a non-binding recommendation for the most suitable track (*consiglio orientativo*), while families

³ Norway (Bütikofer et al., 2023), Finland (Böckerman et al., 2021), or Romania (Malamud and Pop-Eleches, 2011).

retain full autonomy over the final choice. Data on these recommendations are not available; Section 4.3 discusses how the empirical strategy mitigates the resulting concerns.

The second cycle—upper secondary education (grades 9–13, ISCED 3)—begins at age 14 and lasts up to five years. Students choose from three main tracks. The academic track (*liceo* hereafter) prepares students for university⁴ and comprises several streams: scientific, classical, and streams focusing on humanities, arts, foreign languages, and social sciences. The technical track (*istituto tecnico*) combines general education with a specialization, while the vocational track (*istituto professionale*) focuses on job-specific training. Admission is not selective: any student from any middle school may enroll in any program, and all tracks confer eligibility for university admission. Nationally, about 53 percent of students enroll in academic programs, 37 percent in technical programs, and 10 percent in vocational ones. Parental background and academic achievement are relatively homogeneous within tracks but differ substantially across them, being highest on average in the academic track.

Education is moreover widely regarded in Italy as a means of achieving social recognition, particularly among students from disadvantaged families: as Figure 3 shows, students in vocational and technical tracks are more likely than those in academic tracks to report that their motivation to pursue college stems from a desire to fulfil their parents' aspirations or to enjoy the prestige of a university degree. The occupations of classmates' parents are therefore likely to bear on students' educational aspirations.

The primary criterion for assigning students to a middle school is residence within the school's catchment area (Pagani and Pica, 2025). Once admitted, students are allocated to a particular class. Class formation is managed by the school administration—typically the principal in collaboration with a team of teachers—and does not consider parental preferences for specific teachers or classmates. Instead, students are assigned to classes with the aim of creating heterogeneous groups in terms of gender, immigrant status, previous academic achievement, and birth semester, as mandated by law. Students remain in the same class with the same group of peers and teachers throughout middle school, as no reshuffling occurs across subjects or grades. During middle school, the Italian educational curriculum is centrally determined and is uniform for all students nationwide. Students do

⁴ University education is the third education cycle.

not have the option to choose elective subjects or enroll in advanced or honors courses. Classmates attend the same set of subjects and spend the entire school day together in the same classroom. As a result, most student interactions occur within classes rather than across classes within the same cohort.

3. Data and descriptive statistics

Since the 2009–10 school year, INVALSI (the Italian National Institute for the Evaluation of the Education System) has administered annual standardized tests to the entire student population in grades 2, 5, 8, and 10. The empirical analysis draws on administrative data from INVALSI for two cohorts of Italian students enrolled in the final year of middle school (grade 8) during the 2015–16 and 2016–17 academic years. Starting in 2012–13, INVALSI introduced an anonymized student identifier that allows students to be linked across grades. This feature enables me to follow students from grade 8 to grade 10 (high school) in the 2017–18 and 2018–19 academic years. The data additionally provide sociodemographic information on students, including gender, age, citizenship, and parental education and occupation, drawn from school administrative records. To conduct the mediation analysis, I also link students to their primary school grade 5 classmates in the 2012–13 and 2013–14 academic years. Unfortunately, the data include no information on teachers.

3.1 Sample selection

The sample includes 400,537 students in 27,283 classes across 3781 schools in the Center-North of Italy, for whom there is no missing information on the key variables of interest (test scores, demographics, and parental occupations) and who I can link to their high school records in grade 10. Southern regions are excluded for several reasons, including teacher strikes (Meschi and Pavese, 2023), low reliability of test scores (Lucifora and Tonello, 2015), and issues with longitudinal linkage (Bertoni and Parkham, 2025). Additional sample selection criteria reflect the institutional rules governing class formation in Italy and the requirements of the empirical strategy (see Section 4.3). Appendix A discusses sample selection in more detail and reports additional descriptive statistics.

3.2 Variables of interest: defining occupation groups

School records report each parent's occupation from a set of nine predefined categories: unemployed, homemaker, manager, entrepreneur, professional, retailer, teacher, workman, and retiree (De Benedetto and De Paola, 2023). While these nine categories are coarse, each aggregates a fairly homogeneous set of underlying occupations. Appendix A discusses the occupation classification in more details and reports in Table A.3 the full mapping between the nine categories and the individual job titles available in the data. These categories roughly correspond to the ISCO one-digit classification but also capture labor market participation—a relevant indicator of socioeconomic background in the Italian context. Throughout the paper, I use parental occupation as a measure of parental background. To construct an interpretable measure of occupational prestige, I group occupations according to skill level, income profile, and labor force attachment:

- i) Low prestige: unemployed, homemaker, retiree
- ii) Middle prestige: retailer, teacher, workman
- iii) High prestige: manager, entrepreneur, professional

Figure A1 confirms a monotonic earnings⁵ gradient across these three groups.

3.3 Demographics

Table 1 reports descriptive statistics for the estimation sample. The sample is evenly split between male and female students. The variable *regular age* indicates that most students are at the expected age for grade 8. A small share of students—about 3 percent—are first-generation immigrants, while 6 percent are second-generation, reflecting modest levels of immigration during the study period. Grade 8 repetition is relatively rare. Italian law allows parents to enroll their children in primary school before the standard age; these early enrollees are signaled by the variable *early starters*.

4. Empirical framework

4.1 Measuring occupation group class shares

The main explanatory variables capture the share of classmates' parental occupations in each middle school class. For each student i in class c , school s , and cohort t , I define:

⁵ Data on earnings come from the Survey of Income and Wealth (SHIW).

$$Share_{cst}^o = \frac{\sum_{j \neq i} 1\{occgroup_j = o\}}{N_c - 1} \quad (1)$$

where $o \in \{\text{low, middle, high}\}$ denotes the occupational prestige group of classmates' parents and N_c is class size. Shares are computed leaving out the individual i . Where classmates constitute the relevant peer group, I then use the terms *classmates* and *peers* interchangeably throughout the paper. To capture social prestige, I computed the share of classmates whose most prestigious household occupation falls into group o . This approach mitigates the potential misclassification of temporarily unemployed individuals, an issue discussed further in Appendix A. As robustness check, I construct three additional measures of class composition using respectively the share of fathers, mothers and pooled parents⁶, hereafter – collectively referred to as *occupation group shares*. Figure 4 displays the distribution of class level shares, revealing substantial variation across classes within school⁷ for each occupation group and across parental roles. Notably, the shares of classmates with low prestige parental occupations differ substantially between fathers and mothers.

4.2 Identification assumption

The main objective of the research design is to examine how exposure to peers from different parental occupation backgrounds in middle school affects high school choice. In an ideal experiment, students would be randomly assigned to classes within the same school. In this case, two identical students would be exposed to different occupation group shares only because they were by chance assigned to different classes. While true random assignment is rare, Italian institutional rules can generate quasi-random variation in class composition within each school-cohort. This allows me to compare two students with parents belonging to the same occupation group within a school cohort, leveraging the variation coming from assignment to different classes.

⁶ In the online Appendix D, I provide the estimates using these additional shares.

⁷ To isolate within-school variation across classes, I demean each class level average by subtracting the corresponding school level average.

The first threat to identification concerns the potential sorting of students into middle school classes, which could generate systematic differences in the distribution of socioeconomic background across classes. In the Italian institutional context, student sorting does occur but primarily at the school level. Admission to a specific school mainly depends on residential proximity, as each school serves a designated catchment area within which students have the right to enroll. Families may, however, apply to schools outside their catchment area if vacancies exist. Consequently, both observed and unobserved school characteristics – such as location, quality, or prestige – may drive selection patterns that bias estimates of the occupation group shares. To address this concern, the model includes school-by-year fixed effects, which absorb school-specific, time-varying confounders.

Once families have selected a middle school, the Italian law mandates that students be assigned to classes independently of their preferences for specific teachers or classmates. School staff are generally instructed to form classes that are as heterogeneous as possible in terms of prior academic achievement, gender, citizenship status, and age (first- versus second-semester births), in accordance with legal guidelines (Carlana, 2019). Thus, in principle, students should be randomly assigned to classes within each middle school. To comply with these prescriptions, some schools even rely on a random public draft of classes, but this is not a general rule⁸.

Finally, to further control for potential confounders and isolate the effect of class occupational composition, I augment the school-by-year fixed effects with both parents' occupation group fixed effects. Conditioning on own parental occupation further restricts comparisons to students with identical family occupational background who differ only in the composition of their classmates. Importantly, the exclusion bias highlighted by Guryan et al. (2009) cannot be formally tested in this setting, as conditioning on own type is an identifying assumption.

In summary, to identify the causal effect of early exposure to peers' parental occupations, I exploit quasi-random variation in grade 8 class composition. Concretely, I compare two

⁸ For example, the procedure of a school random assignment to classes can be found here: <https://www.icsanfrancescoanguillara.edu.it/oldsite/2020/07/22/sorteggio-classi-i-scuola-secondaria-a-s-2020-2021/>

students who have identical combination of parental occupation group but attended different classes in the same middle school in the same year.

4.3 The empirical approach

I estimate the effect of peers' parental occupation group in student i 's year t , in class c , in school s :

$$Y_{icst} = \alpha + \beta_1 \text{low}_{-icst} + \beta_2 \text{high}_{-icst} + \lambda_{ist} + \varepsilon_{icst} \quad (2)$$

where Y is the binary outcome variable for being enrolled in *liceo* in grade 10. The coefficients of interest in equation (2) are β_1 and β_2 which are associated with the class shares of the low and high occupation group respectively⁹. I isolate school-by-year variation in peers exposure by including a set of school-by-year-by-parental occupation group fixed effects¹⁰ ($\lambda_{i,s,t}$), that allow me to attribute the variation in peers' share (*low* and *high*) solely to variation in class composition within school and year. Pupils who are exposed to different peers within the same school and year are indeed assigned to different classes. For instance, I compare student i , from school s , year t , with father in the *low* occupation group and mother in the *high* occupation group, with student j , from school s , year t , with father in the *low* occupation group and mother in the *high* occupation group. The fixed effect λ controls for school-by-year effects, including local business cycle at a granular level. Finally, λ controls for job characteristics, and particularly for the combination of both parents' occupation group to absorb potential family spillover. Unfortunately, I have no data on teacher recommendations for high school track placement. While teachers' recommendations potentially correlate with class composition, part of this variation is likely related to peer parental background, reflecting inherited ability or stereotypes (Van Maarseveen, 2023¹¹), and is therefore absorbed by the school-by-year-by-parental occupation group fixed effects used in the empirical strategy.

⁹ To verify the sign of the coefficient of interest, parental occupation groups are reclassified in two ways: first, *low* and *middle* are combined to estimate the effect of *high*; second, *middle* and *high* are combined to estimate the effect of *low*. This approach enables estimation of a simplified one-regressor model that isolates each regressor contribution.

¹⁰ Where parental occupations information enters in i .

¹¹ He shows that, conditional on ability, rural students receive systematically less ambitious track recommendations in the Netherlands.

I cluster standard errors at the school level, rather than –less conservatively– at the class level, as the latter assumes independence across classes within the same school. Students from different classes but in the same school may interact, be taught by the same teachers, and are exposed to common school level influences, including the local labor market context. Given that I compare effects across multiple occupation groups I conservatively adjust inference for multiple testing and report False Discovery Rate-adjusted q-values along with clustered standard errors (Anderson, 2008).

4.4 Evidence in support of the identification strategy

The credibility of the identification strategy relies on two main assumptions: sufficient within-cohort variation in peer composition across classes, and the quasi-random assignment of students to classes within schools.

Figure 5 illustrates this identifying variation graphically using binned scatter plots. Panel A (low share) and Panel B (high share) each show four stages: the raw relationship between the peer share variable and the outcome, followed by the same relationship after partialling out school fixed effects, school-by-year fixed effects, and finally my baseline fixed effects (school-by-year-by parental occupation group). The raw correlations are relatively strong, as expected given between-school sorting; however, a clear and stable relationship persists even after the full set of fixed effects is absorbed, confirming that estimates are identified from meaningful within cell variation (details in Appendix B.1). After partialling out the fixed effects from the baseline specification, the standard deviation of the residuals for the low share is 0.036 and for high share is 0.091. These values indicate that, respectively, 71 percent and 53 percent of the original variation remains, leaving ample residual variation on which the analysis relies.

To assess the validity of the quasi-random class assignment assumption, I regress student demographic characteristics on the occupation group shares in equation (2). Reassuringly, Table 2 shows that the peer composition variables are largely uncorrelated with predetermined student traits. While the coefficient on foreign citizen status is statistically significant, its magnitude is minimal: a one standard deviation increase in the share of high prestige peers is associated with just a 0.5 percentage point reduction in the probability of

being a foreign citizen. This suggests little concern for endogenous sorting. Nevertheless, I further investigate this issue in Section 5.3.

Appendix B.2 reports a range of further balance and sensitivity checks, including class level balancing tests based on repeated random reassignment, comparisons of observed and simulated distributions of occupation group shares under random reallocation, and balance tests following Hoxby (2000), Porreca et al. (2026) and Bietenbeck (2025). None indicates systematic sorting of students across classes by parental background.

5. The effect of peer backgrounds on high school track choice

5.1 Results

Table 3 presents the baseline results. To facilitate interpretation, all peer composition variables are standardized in the full sample. Column (1) reports each peer-share coefficient from a separate regression that includes the baseline fixed effects and no additional controls, isolating the low-share and high-share relationships before the two are combined. Column (2) reports the benchmark specification in equation (2) with both peer shares included jointly. The results show that a one standard deviation increase in the share of peers with parents in high prestige occupations raises the probability of enrolling in the academic track by approximately 2.3 percentage points, while a one standard deviation increase in the share of peers from low prestige backgrounds lowers this probability by approximately 0.4 percentage points. These magnitudes are economically meaningful: relative to a baseline academic track enrollment rate of roughly 56 percent, the high share estimate corresponds to a 4 percent change relative to the sample mean. The estimated effect sizes are consistent with those reported in related work on peer background and educational choices (e.g., Cattán et al., 2025; Montonen and Solomon, 2024; Bertoni et al., 2020)¹².

Columns (3)-(5) progressively add controls for gender, foreign citizenship, and parental education to the baseline specification: at the individual level in Column (3), at the class average level (excluding observation i) in Column (4), and at both levels jointly in Column

¹² Cattán et al., 2025 reported that one standard deviation increase in the proportion of elite educated parents in a school-cohort leads to a 2.6 percentage point increase in the likelihood that students in the same school-cohort enroll in an elite degree. Montonen and Solomon, 2024 showed a one standard deviation increase in the white-collar share has a 0.4 percentage point effect on the likelihood that the child will be in a white-collar occupation at age 30. Bertoni et al., 2020 found that an increase of one year in the average years of peers' parental education increases the probability of completing tertiary education by 0.4 percentage points.

(5). The high share coefficient declines more under individual level controls (from 0.0234 to 0.0145) than under class level controls alone (to 0.0179). This pattern suggests that individual characteristics — plausibly correlated with a student's own academic trajectory and track preferences — account for a larger share of the attenuation than peer characteristics do. At the same time, the peer shares regressors likely still absorb some of the correlated variation in peer educational background, since parental occupation correlates with other peers' traits; the remaining decline under class level controls is consistent with this partial overlap. Only under the most demanding specification (Column 5) does the low share coefficient lose statistical significance.

Since the identification strategy relies on class assignment, I adopt Column (3) — which includes individual level controls for gender, citizenship, and parental education alongside school-by-year-by-parental-occupation-group fixed effects — as the preferred specification, as these characteristics are directly relevant to class assignment and therefore to potential confounding.

5.2 Robustness tests

Before proceeding with the mediation analysis on peers' ability, I present several robustness checks on the main results.

First, as previously discussed, a limitation of the data is that parental occupation is observed only for a subset of classmates, which introduces measurement error in the peer variables. To the extent that missing values in the variable of interest are randomly distributed, they bias the estimated effects of peers' parental occupation toward zero. This issue is common in the literature and arises here as well, as roughly one-fifth of pupils lack parental occupation information. I follow the approach of Sojourner (2013) and Fruehwirth and Gagate-Miranda (2019) to correct for this measurement error. Specifically, Sojourner's (2013) correction implies to multiply the occupation group shares by the share of classmates for whom occupation information is observed and includes this share as a control. This correction assumes that missingness in parental occupation is random with respect to observed student characteristics. In Appendix A, I show that this assumption may not hold, although the correction remains preferable to ignoring the issue. Column (1) of Table 4

confirms that the main results survive this correction; I therefore use these measurement error corrected shares in the estimates that follow.

Second, attrition between grades 8 and 10 is non negligible: about 18 percent of students cannot be linked across the two grades. However, this does not pose a threat to the identification strategy, since the treatment variables – the shares of classmates by parental occupation – are defined at the class level in grade 8, before attrition occurs. As such, the peer measures are fixed and identical for all students within a class, regardless of whether they are later observed in grade 10. Consequently, attrition can only affect the sample size used to estimate the outcome, but it cannot influence treatment intensity or generate spurious variation in peer exposure. A characterization of students who cannot be tracked based on demographic traits can be found in Table A6. Nevertheless, I conduct a conservative robustness check by coding students who cannot be followed to grade 10 as not enrolled in a *liceo* (i.e., setting the outcome to zero). This approach is akin to assuming that untracked students did not continue to *liceo*. Column (2) of Table 4 shows that the estimates remain very similar to the baseline, suggesting that attrition is unlikely to substantially affect the results.

Finally, the empirical specification compares students with the same combination of parental occupations, thereby accounting for potential family spillovers. However, this raises the question of whether mothers and fathers contribute differently to peer effects. To examine this, Column (3) of Table 4 reports estimates based on the separate occupation group shares of mothers and fathers. All four coefficients are statistically significant, indicating that both parents' occupational backgrounds are relevant. Nonetheless, the coefficient associated with fathers' occupations is substantially larger, indicating that paternal occupation is the stronger driver of the peer effect documented here. This finding is consistent with the fact that, in 80 percent of the sample, the father holds the most prestigious occupation within the household, and echoes Chise et al. (2021), who similarly find that paternal influence outweighs maternal influence in the intergenerational transmission of college choices.

5.3 Foreign pupils

Given the small but statistically significant correlation between foreign citizen status and class composition previously documented, I assess the robustness of the results by estimating the model separately for Italian and non-Italian students. Table 5 reports these estimates. I cannot reject the null of equal peer share coefficients across the two groups, and therefore find no statistical evidence that immigrant students respond differently to classmates' parental occupational background than their Italian peers.

A related concern is that the share of low prestige occupations among peers may partly proxy for exposure to immigrant classmates, since immigrant families in Italy are on average more likely to hold lower skilled jobs. Recomputing the occupation shares using Italian students¹³ only yields similar distributions and similar estimates (Table 5, Columns 3 and 4). Immigrant background thus correlates with exposure to low prestige occupations without driving differential peer effects. Baseline *liceo* enrollment nonetheless differs sharply by citizenship, at 0.579 for native and 0.382 for foreign students, consistent with Carlana (2022): a gap that predates the peer composition effects estimated here.

5.4 Decomposition of parental occupations and parental education

Next, I investigate whether parental occupation constitutes a distinct dimension of peer background, or instead proxies for parental education. I follow the same procedure used for occupations. I define three groups of parental education — low, middle, and high — corresponding to less than a high school diploma, a high school diploma, and more than a high school diploma, respectively, and assign each household the highest educational level attained by either parent. I then compute the class share of parents with low and high education. Correspondingly, I define fixed effects for school-by-year, mother's occupation, father's occupation, mother's education, and father's education, mirroring the baseline specification and resulting in a highly saturated model.

Column (1) of Table 6 reproduces the occupation baseline: a one standard deviation increase in the share of high prestige occupation peers raises the probability of *liceo* enrollment by 1.62 percentage points, while an equivalent increase in the share of low prestige-occupation peers lowers it by 0.31 percentage points. Column (2) reports the analogous specification using peer parental education: a one standard deviation increase in the share of highly

¹³ Distributions can be seen in figure A2.

educated peers raises *liceo* enrollment by 1.83 percentage points, while a one standard deviation increase in the share of peers with low parental education lowers it by 0.58 percentage points.

Column (3) includes both sets of peer shares jointly. All four coefficients remain statistically significant, indicating that occupation and education each capture variation in peer background not fully absorbed by the other. The attenuation is similar for the two low share coefficients, but more pronounced for the high share coefficients. The high prestige occupation coefficient is therefore more sensitive to the inclusion of the education measure than the reverse, suggesting that part of the estimated effect of exposure to high prestige occupation peers overlaps with peer parental education, while the education coefficient captures a comparatively more stable component of peer background. Considering, that both sets of coefficients remain statistically significant when included jointly suggests that parental occupation captures variation in peer background that education does not fully absorb, plausibly related to channels such as aspirations, information, or social norms tied to occupational identity. I do not have direct evidence to establish which of these channels drives this residual independent role of occupation, and I leave a full decomposition of mechanisms to future work.

6. The role of peers' ability and own ability on high school track choice

6.1 Own ability and peers' ability

Students' academic ability is partly associated with parental background, as parental occupations is a trait of parental socioeconomic status, which is in turn correlated with children's own ability (Sirin, 2005). Because academically stronger students tend to come from more advantaged backgrounds, exposure to classmates from high prestige families may also entail exposure to more academically able peers. This raises the possibility that any effect of peer background on high school choice operates, at least in part, through an ability channel.

Table 7 investigates this scenario using the same specification as the baseline model, with peer occupation group shares as regressors and the same set of fixed effects and demographic controls; the two columns differ only in the outcome variable, student i 's own grade 8 ability in Column (1) and the average ability of student i 's classmates in Column (2).

Academic performance is measured using students' grade 8 standardized test scores in mathematics and reading. To ensure comparability across cohorts and subjects, I compute each student's percentile position in the national distribution for each subject and average the two percentiles¹⁴.

The coefficient on peer occupation shares is statistically insignificant in Column (1), indicating that peer background composition does not directly affect students' own ability¹⁵. In Column (2), peer occupation shares are significantly associated with classmates' average ability, confirming that peer occupation composition is correlated with the actual ability level of the peer group. Taken together, these results motivate the choice of classmates' average ability as a candidate mediator in the mediation analysis presented in the next section, which formally tests whether peer ability accounts for part of the estimated effect of peer background on high school choice.

6.2 Mediation analysis

The paper so far has shown that exposure to high(low)-prestige peers increases(reduces) the probability that students enroll in the academic high school track and impacts average class ability.

In this section, I quantify the contribution of students' ability channel to the overall effect of peers' parental occupation on the probability to enroll in an academic high school. Ideally one would want to control for both parental occupations and peers' ability, but the latter is determined by their parents' ability which is reflected by their occupations. Fruehwirth and Gagete-Miranda (2019) faced a similar challenge when testing spillovers of peer parental education on children's achievement at the beginning of kindergarten, after controlling for peer achievement. Nevertheless, parental occupations could be a proxy for parental ability, that affects their pupils' scores. This makes students' test score a bad control to estimate the impact of parental background. Consider the following model:

$$Y_{icst} = \alpha + \gamma_1 low_{-icst} + \gamma_2 high_{-icst} + \gamma_3 TS_{-icst} + \gamma_4 X_{icst} + \lambda_{ist} + \varepsilon_{icst} \quad (3)$$

¹⁴ The standardized test scores available in INVALSI are constructed to have a mean of 200 and a standard deviation of 40 and are corrected for cheating behavior (see Quintano et al., 2009).

¹⁵ This result also provides evidence against teachers sorting students into classes according to academic ability.

where TS is the average class ability measured through percentile position in the nationally standardized test scores. The parameters γ_1 and γ_2 are the direct effect of peer parental background conditioning on peers' ability. Estimating equation (3) by OLS would not identify γ_1 and γ_2 because, as shown in Table 7 Column (2), parental occupations are correlated with average class ability and it is likely to be correlated with unobserved individual determinants of academic high school enrollment (and hence endogenous). In order to identify the effect of parental background, I instrument TS in equation (3). This strategy is akin to the mediation analysis with endogenous mediators used in Cattani et al. (2025). The directed acyclic graph in Figure 5 shows the mediating role of peers' ability.

To instrument ability, I take advantage of the longitudinal dimension of the data to recover the partially overlapping peer groups (see De Giorgi et al., 2010 and Nicoletti et al., 2023). In this context, this approach translates into identifying students who were in the same class in grade 5 of i 's grade 8 peers, but who were not in the same class as i in grade 5. In other words, partially overlapping peer groups of pupil i are previous classmates of i 's peers who were not classmates of i . This strategy generates individual-specific groups, which helps in controlling for potential endogeneity caused by correlated influences within groups¹⁶. The first stage of equation (3) is:

$$TS_{-icst} = \alpha + \delta_1 low_{-icst} + \delta_2 high_{-icst} + \delta_3 Z_{-icst} + \delta_4 X_{icst} + \lambda_{ist} + u_{icst} \quad (4)$$

where the instrument (Z) is the average ESCS¹⁷, a comprehensive socioeconomic background indicator, among i 's grade 5 class. I do not rely on the ability of partially overlapping peer groups to avoid simultaneity: I only measure ability after five years of primary school interactions, thus test scores may be reflected among peers (see Manski,

¹⁶ The seminal paper by De Giorgi et. al, 2010 stated that "having peer groups that vary at the individual level guarantees the presence of excluded classmates. The exogenous characteristics of such excluded peers are a natural set of instruments to overcome potential endogeneity generated by common (correlated) group effects".

¹⁷ This indicator is built in accordance with the one proposed in the OECD-PISA framework and considers parental occupations, parental education and other items (computer availability, number of books at home) collected in the students' questionnaire, see De Benedetto and De Paola, 2023. In equations (3) and (4), occupation group shares are included directly, thus the ESCS generates variation in peer ability that is residual to these occupational channels and plausibly related to variables such as household wealth and income, that shape children's academic performance.

1993). Following De Giorgi et al. (2010), I instead use a measure of parental background that can be considered as time-invariant, predetermined and exogenous. Since partially overlapping peers and i can only influence each other through i 's grade 8 peers (with whom the partially overlapping peers were in the same grade 5 class), confirming the exclusion restriction. The partially overlapping peers categorization mechanically excludes the grade 8 i 's peers who were peers of i also in grade 5. For the sake of consistency, the occupation group shares (*top – low share* and *top -high share*) and average class ability (TS) in equations (3) and (4) are computed on new peers (peers of i in grade 8 who were not peers of i in grade 5, see Meschi and Pavese (2023) and Gibbons and Telhaj (2016))¹⁸. I standardized these measures in the sample for ease of interpretation. Due to data linkage limitations, only 385,668 pupils can be followed from primary school (grade 5) through middle school (grade 8) to high school (grade 10), allowing the reconstruction of their complete classmates' history.

The exclusion restriction would fail if the socioeconomic composition of the partially overlapping peer group were correlated with unobserved determinants of i 's track choice through channels other than peer ability, for instance if primary schools in wealthier neighborhoods systematically fed into the same middle schools. Three features of the design limit this concern. First, school-by-year fixed effects absorb time-varying school level confounders, including the socioeconomic composition of the catchment area. Second, the instrument varies at the individual level rather than at the school or neighborhood level, and third, it is based on grade 5 class composition, itself subject to the same heterogeneous class formation rules that govern middle school assignment. Appendix B.3 develops these arguments and reports a test of conditional random assignment of the instrument, which finds no association large enough to generate meaningful bias.

In this mediation analysis, I assume that the reduced-form effect of peers, which captures the total effect (i.e. β_2 in Equation (2)) can be decomposed into a direct effect (the effect of peers' parental occupation on academic high school track, γ_2) and an indirect effect

¹⁸ Table A4 exemplifies the different type of peers discussed in the paper.

operating through peers' average ability (γ_3)¹⁹. This decomposition considers these two channels but does not preclude the existence of others.

Column (3) of Table 8 reports the first stage estimates of δ_1 and δ_2 from equation (4), which show the effects of new peers' parental occupational prestige on new peers' average class ability. A higher average ESCS in the grade 5 classes of i 's new grade 8 peers is associated with higher measured ability among such peers at the end of middle school (δ_3). This finding suggests that i 's new peers score higher in grade 8 when the primary school classmates of those peers - whom i never met - came from more advantaged backgrounds. These estimates support the relevance of the instrument, with the Kleibergen-Paap rk Wald F-statistic equal to 10.07, which lies at the margin of the conventional threshold.

Column (4) reports the reduced-form effect of the average ESCS of the partially overlapping peer group on i 's probability of enrolling in the academic track, that turns out to be a precisely estimated zero though not statistically significant²⁰.

Finally, Column (5) reports the IV estimates. The instrumented coefficient on new peers' average ability (parameter γ_3 in equation 3) is negative but imprecisely estimated, and the null of no effect cannot be rejected. The occupation group shares retain the pattern of Column (1): the coefficient on the *New peers top - high share* in the IV specification is 0.0254 and statistically significant, which means that an increase in exposure to high prestige new peers by 1 standard deviation fosters enrollment in the academic high school track by 2.5 percentage points.²¹

Panel B of Table 8 decomposes the total high prestige new peer effect on student enrollment of an academic high school track - β_2 in equation (2) - into the direct effect γ_2 from equation (3) and the indirect effect ($\delta_2 \times \gamma_3$) working through new peers' average class ability. The

¹⁹ Similarly, for low prestige is: $\beta_1 = \gamma_1 + \gamma_3$.

²⁰ This negative effect of peers ability on academic high school track can be explained through rank effect see Bertoni and Nisticò (2023).

²¹ As the first-stage F-statistic reported in Column (3) is only marginally above 10, I follow Angrist and Pischke (2009), who advise reporting both just-identify 2SLS and over-identify LIML estimates and treating agreement between the two as reassuring evidence against weak-instrument or many-instrument bias. I use the partially overlapping peers' class average female share as an additional instrument, making the model overidentified. LIML is approximately median-unbiased under weak identification, whereas 2SLS is biased toward OLS in finite samples when instruments are weak, although LIML is less efficient. The estimates are reported in Table B.6. The coefficients on new peers' average class ability in Table 8, column (5), and Table B.6 are very similar: -0.0863 (s.e. = 0.0596) and -0.0812 (s.e. = 0.0598), respectively, providing reassurance against weak-instrument bias.

total effect matches the total effect OLS (i.e., unconditional on new peers' class average ability) in Column (1). Interestingly, the direct effect of high prestige new peers is large enough to cancel out the negative effect coming from new peers' class average ability (though it is not statistically significant). The direct effect of high prestige new peers could reflect the effect of such peers on students' information set (information, aspirations, and perceived social prestige directly linked to the occupational returns of academic high school track) and/or through their preferences for such academic high school tracks. The current strategy and data do not allow me to be more precise on the nature of these underlying mechanisms. To isolate such mechanisms, one would need additional data about within-class friendship network and parent-to-parent interactions.

6.3 Heterogeneity analysis

Although the baseline specification relies on a linear peer effects model, meaningful heterogeneity may still exist, indicating that some students benefit more than others from exposure to specific peer backgrounds. Understanding these distributional effects is essential for designing targeted interventions, particularly in contexts where peer composition is difficult to manipulate. First, I examine whether peer effects differ across high school tracks, and then I explore additional heterogeneity by geographic socioeconomic context and individual student backgrounds.

I begin by disaggregating the results by academic track. Table 9 shows that peer effects are concentrated in the most academically demanding programs – scientific and classic – while they are negligible or absent in other academic curricula such as music, social science, art, and language. Coherently, for technical and vocational tracks, the pattern reverses: greater exposure to peers from low prestige occupational backgrounds increases the likelihood of enrolling in these programs, whereas exposure to high prestige peers reduces it. These findings complement the interpretation that parental occupation influences students' educational choices not only through informational channels but also by signaling status-linked expectations and aspirations.

Next, I investigate heterogeneity across provinces using absolute upward mobility ranking from Acciari et al. (2022), where provinces with lower rank exhibit higher levels of social mobility. Table 10 shows that exposure to low prestige peers has a stronger negative effect

in low-mobility provinces (Column 2), whereas it is absent in high-mobility areas (Column 1). This pattern underscores the interaction between peer effects and local opportunity structures. This result is confirmed when the sample is split by the median ESCS measured in grade 5, an alternative measure of socioeconomic status that varies within both schools and provinces (Table 10, Columns 3 and 4). One possible explanation is that richer provinces or more advantaged families can mitigate the negative influence of disadvantaged peers through better schools or greater parental support. Importantly, these findings highlight the importance of the estimated peer effects to improve intergenerational mobility, even in depressed economic locations. I leave a deeper examination of these contextual mechanisms for future research. Considering that the analysis excludes Southern regions – the poorest areas of Italy – these findings are particularly striking.

Finally, Table 11 examines whether the effect of peer exposure varies by students' own parental occupational group. The coefficient on the high prestige peer share is positive and significant for all three groups but declines monotonically in own family prestige. Exposure to high prestige peers therefore has the largest effect on students whose own family background would otherwise predict lower academic track enrollment, a pattern that is naturally consistent with an information or aspiration channel than with a strict social comparison account, and that carries a direct implication for intergenerational mobility, as it indicates that this margin of peer exposure works to equalize rather than amplify existing advantage.

The coefficient on the low prestige peer share is instead significant only for students from middle occupation families. Mean academic track enrollment is 0.501 for them confirming they are closest to the enrollment margin. This finding supports an interpretation in which these students are more sensitive to a reduction in high prestige peer exposure because they are closer to the margin of track choice, although social comparison or other mechanisms tied to parental occupation cannot be ruled out.

7. Conclusion

This paper shows that exposure to classmates' parental occupations during middle school causally influences students' high school track decisions in Italy. Using administrative data that follow students over time and an identification strategy based on within-school, within-

cohort variation in class composition, I distinguish the social influence of peers' parental backgrounds from that of peers' academic ability. Italy provides a well-suited setting for this research question: students select their high school track at age 14 and face virtually no formal restrictions across programs, which makes track choice a credible example of a partially informed yet discretionary decision.

A higher share of high prestige peers substantially increases the probability of enrolling in an academic track (*liceo*), while a higher share of low prestige peers reduces it. These effects are economically meaningful and robust across specifications, alternative measures, and samples. Exposure to high prestige peers benefits students from all backgrounds, and most strongly those from low occupation families, whereas the effects of low prestige exposure are concentrated among students in the middle of the occupational distribution. Effects are also stronger in provinces with lower intergenerational mobility, which indicates that peer composition interacts with local opportunity structures.

Combining quasi-random classroom assignment with a mediation design based on partially overlapping peer groups, I find evidence that peers' academic ability does not mediate the effect of peers' parental occupations. The effect of high prestige exposure survives instrumentation of peer ability, suggesting that information, aspirations, and perceived social prestige may be distinct channels through which peers' parental occupations shape educational decisions.

Three main conclusions emerge. First, peer networks can help mitigate school stratification in a system where social stratification remains persistent despite substantial freedom of educational choice. Second, students from disadvantaged families benefit most from exposure to peers from more advantaged families, while immigrant students appear not to benefit from comparable exposure, which raises the question of what prevents them from translating favorable peer environments into track choices. Third, classmates' parental occupations are especially salient in explaining enrollment in the more demanding academic tracks, those most closely linked to prestigious careers. The magnitude of the estimated effects aligns with prior research on peer influences in education (Cattan et al.,

2025; Montonen and Solomon, 2024; Bertoni et al., 2020): peer exposure accounts for roughly one third of the effect of a student's own parental occupation on high school track choice²².

The effect of peer exposure is asymmetric. A one standard deviation increase in the share of high prestige peers raises the probability of enrolling in the academic track by 2 percentage points, whereas an equivalent increase in the share of low prestige peers reduces it by only 0.5 percentage points. This asymmetry is qualitatively consistent with a reference dependent account of aspiration formation, in which students form aspirations partly in relation to a reference point anchored on the resources of their own family (Fongoni et al., 2022). Under this account, family background sets a floor that limits downward revisions of aspirations after exposure to lower prestige peers, while leaving more scope for upward revisions after exposure to higher prestige peers. I also show that the gap between the high prestige and low prestige coefficients declines monotonically in own family prestige. An alternative explanation rests on proximity to the enrolment margin. Since academic track enrolment is the default outcome for students from high prestige families, exposure to low prestige peers has little scope to push them towards the vocational track, whereas students from low and middle prestige families are closer to the margin and therefore more responsive to high prestige peers, who make the academic track more salient and attainable. The two accounts make different predictions for students at the same distance from the margin: the asymmetry should vanish under the second but persist, and vary with family prestige, under the first. Testing this would require either direct measures of aspirations, elicited before and after peer exposure, or a setting in which students from high prestige families are not concentrated at the default outcome. I cannot distinguish between these two explanations with the present data and treat them as plausible and complementary accounts of the observed asymmetry.

While this study provides valuable insights, it is important to acknowledge limitations that could be addressed in future research. While the mediation analysis provides evidence that ability does not mediate the network influence, the residual effect remains only partly resolved, at two levels. At a coarser level, factors correlated with parental background but

²² The estimates are obtained by regressing *liceo* on two binary variables taking value 1 for having a parent in the low occupation group and high occupation group respectively, including demographic controls and school-by-year fixed effects.

conceptually distinct from it may still contribute to the estimated effect, including peers' non-cognitive skills (Shan and Zölitz, 2025), teacher behavior (Carlana, 2019 and 2026), and the fact that parental background shapes friendship formation among classmates (Agostinelli et al., 2026). Because the identification strategy compares students who share the same combination of own parental occupations, any teacher bias or friendship formation channel that correlates with the student's own socioeconomic status is differenced out by construction. This does not, however, address bias or friendship sorting driven by the composition of peers' backgrounds, which is the treatment of interest, so these channels cannot be ruled out. Data on teacher behavior and non-cognitive skills are, however, not available.

At a finer level, parental background itself aggregates several potentially distinct channels that this design cannot separate: role models, in the sense of peers' parents as examples of feasible trajectories; transmission through close social ties; and preference for social status, interpreted as aspirational alignment with the socioeconomic level of the group. Isolating them would require additional data on the within-class friendship network, in order to test whether the effect operates along close ties or uniformly across classmates regardless of tie strength; on students' awareness of their classmates' parental occupations; and on parent-to-parent contact, in order to isolate channels operating through parents rather than through children. Chung and Zou (2023), for instance, suggest that parental investment behavior is one such channel, with parents adjusting their engagement and responsiveness when exposed to higher-status peers. My results are consistent with such an account, although the present design cannot test it directly.

Overall, the findings show that peer social environments influence educational decisions, often to the disadvantage of students from lower socioeconomic backgrounds, and identify peer exposure as a possible lever for reducing school stratification. The most immediate policy implication concerns the allocation of peers. In our setting, this would entail reallocating high prestige peers from high-prestige students to low prestige students. Although high prestige students also benefit from exposure to high prestige peers, the estimated effect is substantially smaller than that for low prestige students. Such a reallocation therefore need not be Pareto improving, and its desirability ultimately depends on the social welfare function adopted by the policymaker.

Beyond the allocation of peers, policy could also focus on mitigating negative peer environments through targeted interventions for students most exposed to them. Identifying these students by parental occupation, socioeconomic status, or geographic location would allow schools to implement relatively low-cost support programs, such as mentoring schemes (Resnjanskij et al., 2024), information treatments (Facchetti et al., 2025), or summer learning programs (Azzolini et al., 2023). The precise policy implications, however, depend on the underlying mechanism, which the present analysis cannot fully identify, as the candidate channels have different implications for policy design. When information is the dominant mechanism, targeted provision about the returns to different tracks may be a cost-effective intervention. Where social prestige or identity considerations dominate instead, changing classroom composition represents a more structural intervention, and lighter-touch policies are unlikely to suffice on their own. If aspiration formation is the main driver, the timing of interventions matters accordingly. Thus, while peer reallocation represents the most direct structural response to the estimated effects, targeted interventions provide a complementary approach when reshuffling classroom composition is undesirable or infeasible, with their effectiveness likely depending on the mechanism at work.

References

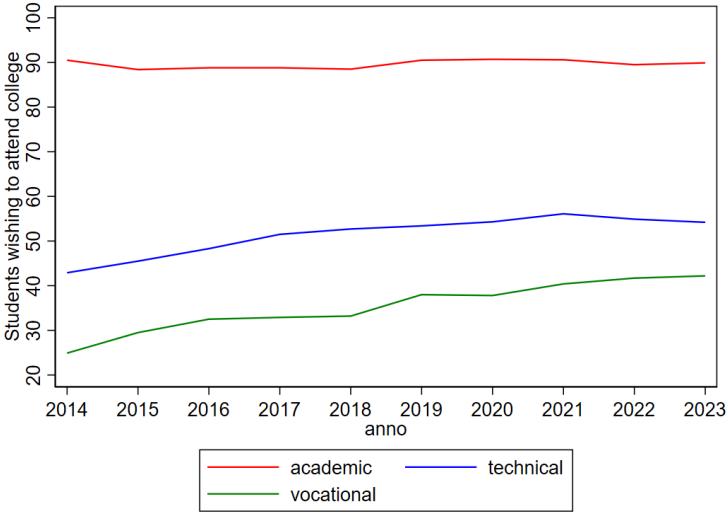
- Acciari, P., Polo, A., & Violante, G. L. (2022). And yet it moves: Intergenerational mobility in Italy. *American Economic Journal: Applied Economics*, 14(3), 118-163.
- Adamecz-Völgyi, A., Henderson, M., & Shure, N. (2023). The labor market returns to “first-in-family” university graduates. *Journal of Population Economics*, 36(3), 1395-1429.
- Agarwal, L., Brunello, G., & Rocco, L. (2021). The pathways to college. *Journal of Human Capital*, 15(4), 554-595.
- Agostinelli, F., Doepke, M., Sorrenti, G., & Zilibotti, F. It Takes a Village: The Economics of Parenting with Neighborhood and Peer Effects. *Journal of Political Economy*.
- Almadiploma, 2024. <https://www.almadiploma.it/indagini/profilo/profilo.aspx>. Consulted on 30 July 2024.
- Anderson, M. L. (2008): “Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects,” *Journal of the American statistical Association*, 103, 1481-1495.
- Andrews, R. J., Imberman, S. A., & Lovenheim, M. F. (2020). Recruiting and supporting low-income, high-achieving students at flagship universities. *Economics of Education Review*, 74, 101923.
- Anelli, M. (2020). The returns to elite university education: A quasi-experimental analysis. *Journal of the European Economic Association*, 18(6), 2824-2868.
- Anelli, M., and Peri, G. (2019). The effects of high school peers’ gender on college major, college performance and income. *The Economic Journal*, 129(618), 553-602.
- Azzolini, D., Bazzoli, M., Burlacu, S., Rettore, E. (2023). “Starting the School Year on the Right Foot. Effects of a Summer Learning Program Targeting Vulnerable Students in Italy”. IZA discussion paper IZA DP No. 16498.
- Bertoni, M., Brunello, G., & Cappellari, L. (2020). Who benefits from privileged peers? Evidence from siblings in schools. *Journal of Applied Econometrics*, 35(7), 893-916.
- Bertoni, M., & Parkham, S. (2025). When do ordinal ability rank effects emerge? Evidence from the timing of school closures. *Economics of Education Review*, 105, 102633.
- Bertoni, M., & Nisticò, R. (2023). Ordinal rank and the structure of ability peer effects. *Journal of Public Economics*, 217, 104797.
- Bietenbeck, J. (2025). Do motivated classmates matter for educational success?. *The Economic Journal*, 135(665), 36-58.
- Biewen, M., & Tapalaga, M. (2017). Life-cycle educational choices in a system with early tracking and ‘second chance’ options. *Economics of Education Review*, 56, 80-94.
- Bifulco, Robert, Jason M. Fletcher, and Stephen L. Ross. 2011. "The Effect of Classmate Characteristics on Post-secondary Outcomes: Evidence from the Add Health." *American Economic Journal: Economic Policy*, 3 (1): 25-53.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. "Under pressure? The effect of peers on outcomes of young adults." *Journal of Labor Economics* 31.1 (2013): 119-153.
- Böckerman, P., Haapanen, M., Jepsen, C., & Roulet, A. (2021). School tracking and mental health. *Journal of Human Capital*, 15(2), 291-345.
- Bütikofer, A., Ginja, R., Løken, K. V., & Landaud, F. (2023). Higher-Achievement Schools, Peers and Mental Health. *The Economic Journal*, 133(655), 2580-2613.
- Carlana, M. (2019). Implicit stereotypes: Evidence from teachers’ gender bias. *The Quarterly Journal of Economics*, 134(3), 1163-1224.

- Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti. 2022. "Implicit Stereotypes in Teachers' Track Recommendations." *AEA Papers and Proceedings*, 112: 409–14.
- Carlana, Michela and Miserocchi, Francesca and Patacchini, Eleonora, Tracking Inequality: Teachers and the Allocation of Educational Opportunities (September 2026). NBER Working Paper No. w35701, Available at SSRN: <https://ssrn.com/abstract=7426428>
- Cattan, S., Tominey, T., & Salvanes, K.G., (2025). First generation elite: the role of school social networks. *American Economic Review*, *Forthcoming*.
- Chise, D., Fort, M., & Monfardini, C. (2021). On the intergenerational transmission of STEM education among graduate students. *The BE Journal of Economic Analysis & Policy*, 21(1), 115-145.
- Chen, Liwen, Bobby W. Chung, and Guanghua Wang. 2025. Composition of peer mothers and gender norms: Class randomization and short-run effects." *Journal of Economic Behavior & Organization* 234:107008.
- Chung, B.W. & Zou, J. (2023) Understanding spillover of peer parental education: randomization evidence and mechanisms. *Economic Inquiry*, 61(3), 496–522.
- De Benedetto, M. A., & De Paola, M. (2023). Immigration and teacher bias towards students with an immigrant background. *Economic Policy*, 38(113), 107-154.
- De Giorgi, G., Pellizzari, M., & Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, 2(2), 241-275.
- Dustmann, C. (2004). Parental background, secondary school track choice, and wages. *Oxford Economic Papers*, 56(2), 209-230.
- Facchetti, N., Ghisolfi, S., Marcolongo, G., Le Moglie, M. (2025) "HERO Teachers: Harboursing Empathy to Raise Opportunities. A Study of Teachers' Burnout risks and Students' Outcomes"
- Fongoni, M., Norris, J., Romiti, A., & Shi, Z. (2022). Reference dependent aspirations and peer effects in education.
- Fruehwirth, J. C., & Gagete-Miranda, J. (2019). Your peers' parents: Spillovers from parental education. *Economics of Education review*, 73, 101910.
- Gibbons, S., & Telhaj, S. (2016). Peer effects: Evidence from secondary school transition in England. *Oxford Bulletin of Economics and Statistics*, 78(4), 548-575.
- Guryan, J., Kroft, K., & Notowidigdo, M. J. (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics*, 1(4), 34-68.
- Goulas, S., Gunawardena, B. N., Megalokonomou, R., and Zenou, Y. (2024). Gender role models in education. *Monash University working paper*, ISSN number 1141-5429.
- Hoxby, Caroline M., "Peer Effects in the Classroom:," NBER Working Paper, 2000.
- Imberman, S. A., Kugler, A. D., & Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review*, 102(5), 2048-2082.
- Lucifora, C., & Tonello, M. (2015). Cheating and social interactions. Evidence from a randomized experiment in a national evaluation program. *Journal of Economic Behavior & Organization*, 115, 45-66.
- Malamud, O., & Pop-Eleches, C. (2011). School tracking and access to higher education among disadvantaged groups. *Journal of Public Economics*, 95(11-12), 1538-1549.
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3), 531-542.

- McConnell, B., Munisamy, P., Zapryanova, M. Community Schools and Student Behavioral Outcomes: Evidence From New York City
- Mertz, Mikkel and Ronchi, Maddalena and Salvestrini, Viola, Female Representation and Talent Allocation in Entrepreneurship: The Role of Early Exposure to Entrepreneurs, 2025. IZA DP No. 17801.
- Meschi, E., & Pavese, C. (2023). Ability composition in the class and the school performance of immigrant students. *Labour Economics*, 85, 102450.
- Mocetti, S. (2007). Intergenerational earnings mobility in Italy. *The BE Journal of Economic Analysis & Policy*, 7(2).
- Montonen, Jerry and Solomon, Samuel, The Impact of Peers on Occupational Choice and Long-Term Outcomes (October 09, 2024). Available at SSRN: <https://ssrn.com/abstract=4981877> or <http://dx.doi.org/10.2139/ssrn.4981877>. Working paper.
- Nicoletti, C., Salvanes, K. G., & Tominey, E. (2023). Mothers working during preschool years and child skills: does income compensate?. *Journal of Labor Economics*, 41(2), 389-429.
- Olivetti, Claudia, Eleonora Patacchini, and Yves Zenou. 2020. Mothers, peers, and gender-role identity." *Journal of the European Economic Association* 18 (1):266-301.
- Pagani, L., Pica, G., A peer like me? Early exposure to high achievers in math and later educational outcomes, *The European Economic Review*, Volume 181, January 2026, 105177.
- Porreca, E., Rizzica, L., Tonello, M., (2026). A Great Depression? Students' aspirations during the Great Recession. *Bank of Italy Working paper number 1528*.
- Resnjanskij, S., Ruhose, J., Wiederhold, S., Woessmann, L., & Wedel, K. (2024). Can mentoring alleviate family disadvantage in adolescence? a field experiment to improve labor market prospects. *Journal of Political Economy*, 132(3), 1013-1062.
- Shan, X., & Zölitz, U. (2025). Peers affect personality development. *Review of Economics and Statistics*, 1-45.
- Sojourner, A. (2013). Identification of peer effects with missing peer data: Evidence from Project STAR. *The Economic Journal*, 123(569), 574-605.
- Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417-453.
- Van Maarseveen, Raoul, Location and educational signals (June 22, 2023). Available at SSRN: <https://ssrn.com/abstract=4488043> or <http://dx.doi.org/10.2139/ssrn.4488043>.

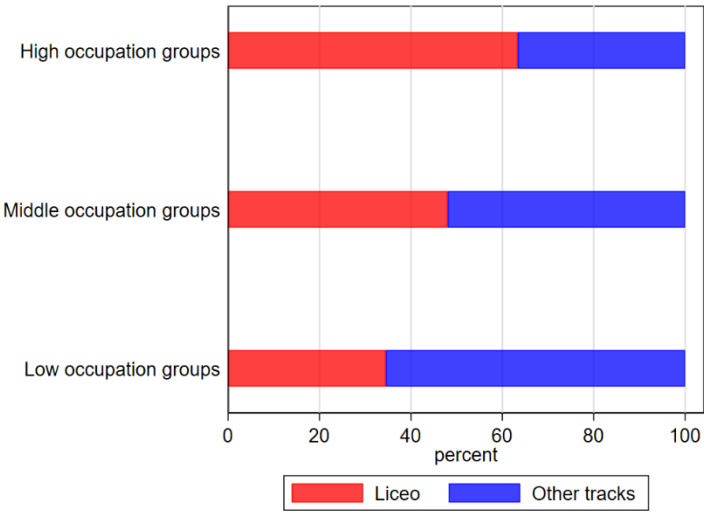
Figures

Figure 1. Percentage of last year high school students who reported that they wished to attend college the following year, by different high school tracks. Source: Almadiploma, 2024.



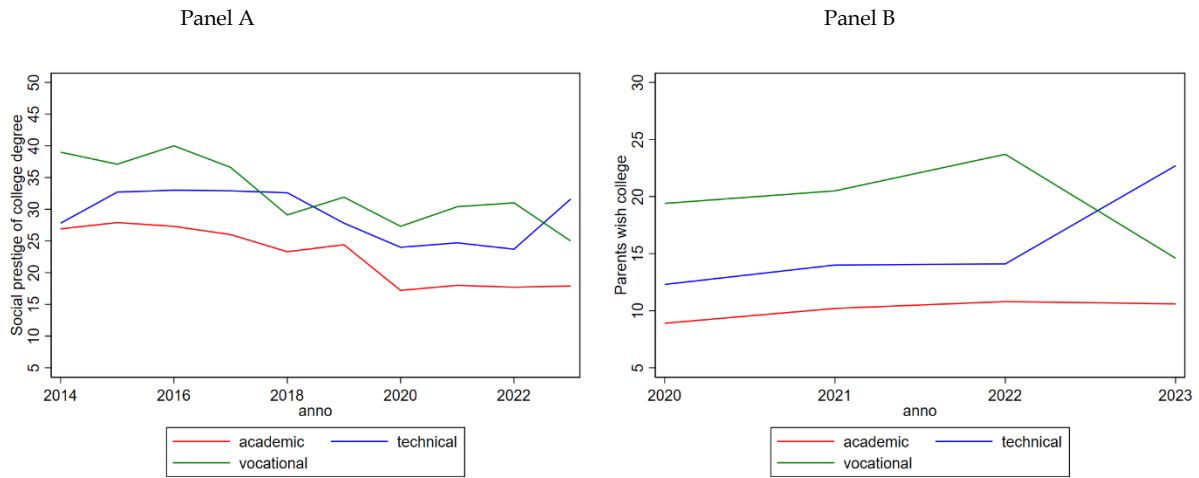
Notes: The figure reports college enrolment intention over years by high school tracks. Students can choose among i) intend to continue education. ii) Undecided about to intend to continue education ii) do not intend to continue education.

Figure 2. Percentage of students enrolled in academic high school track by parental occupation groups. Source: INVALSI.



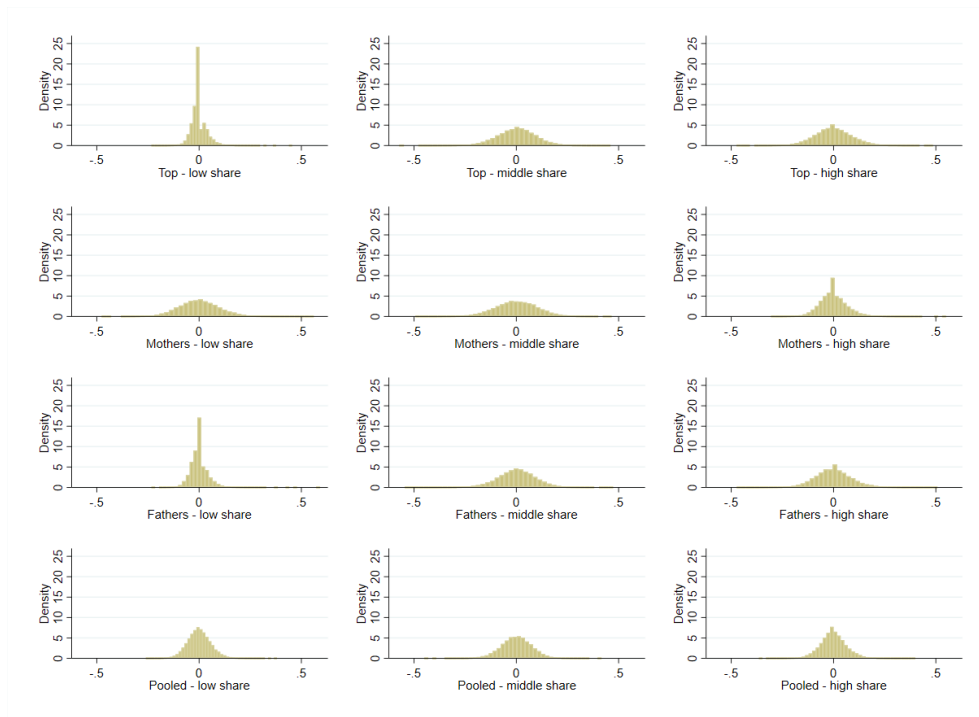
Notes: Occupation group: low (unemployed, homemaker, retiree), middle (retailer, teacher, workman), high (manager, entrepreneur, professional). *Liceo* refers to academic high school track in Italy. *Others* refer to technical and vocational high school track in Italy.

Figure 3. Percentage of students who wish to attend college because of social prestige (Panel A) or to fulfil parents' wishes (Panel B). Source: Almadiploma, 2024.



Notes: Almadiploma data are collected during high school last year. Students are asked which are the most important motivations for them if they intend to enroll in college. They can choose among i) want to deepen their cultural interests. ii) university education is needed to do the job they are interested in. iii) want to have a well-paid job. It is difficult for a high school graduate to find a job. iv) want the social prestige of a degree. v) are interested in social contacts and leisure opportunities that come with being a student. vi) parents would like them to continue their studies at university.

Figure 4. Density of occupation group shares.

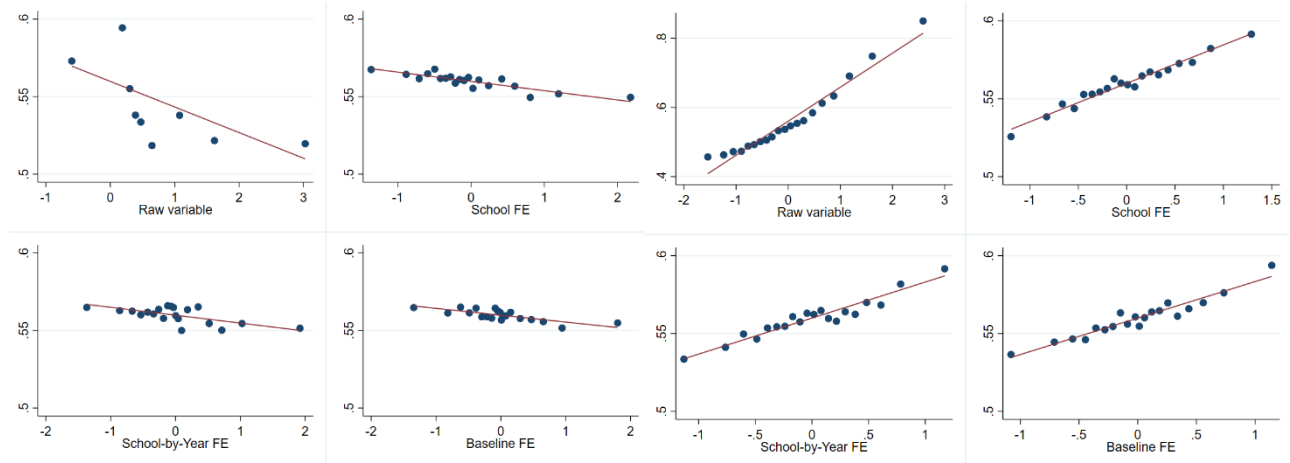


Notes: Occupation groups: low (unemployed, homemaker, retiree), middle (retailer, teacher, workman), high (manager, entrepreneur, professional). Density of the occupations group shares by classes within the same school. The top share is based on the most prestigious occupation group within each household while the mothers', fathers', and pooled shares are computed as the proportion of mothers, fathers, or all observed parents, respectively, in each occupation group relative to their corresponding totals. Number of classes: 27,283.

Figure 5. Binned scatter plots of the relationship between peer occupation shares and *liceo* enrollment, before and after partialling out fixed effects.

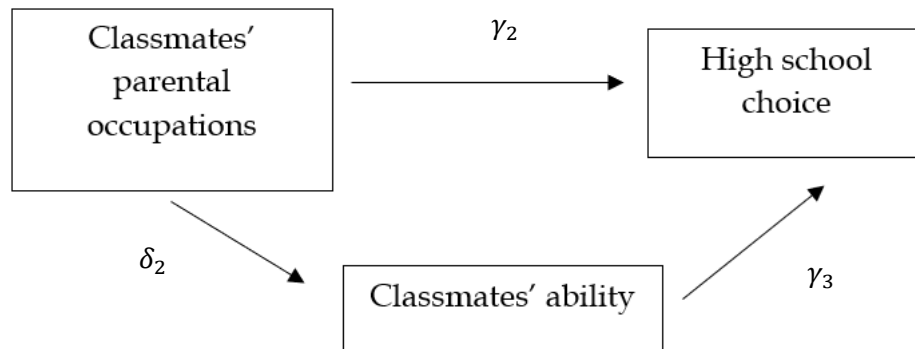
Panel A. Low share.

Panel B. High Share.



Notes: For each panel, left top corner reports the raw correlation with no fixed effects. Right top corner (2) adds school fixed effects. Left bottom corner adds school-by-year fixed effects. Right bottom corner reports baseline specification, which absorbs own-parental-occupation-background $\times$ school $\times$ year fixed effects.

Figure 6. Mediating role of classmates' academic ability.



Notes: Decomposition of total effect (i.e. β_2 in Equation (2)) into a direct effect (the effect of peers' parental occupation on academic high school track, γ_2) and an indirect effect operating through peers' average ability (γ_3).

Tables

Table 1. Summary statistics of the sample.

	N	Mean	Standard deviation
<i>Occupations group class shares</i>			
Top - low	400,537	0.032	0.053
Top - middle	400,537	0.691	0.176
Top - high	400,537	0.277	0.178
Mothers - low	400,537	0.304	0.157
Mothers - middle	400,537	0.571	0.155
Mothers - high	400,537	0.125	0.117
Fathers - low	400,537	0.039	0.059
Fathers - middle	400,537	0.723	0.169
Fathers - high	400,537	0.237	0.169
Pooled - low	400,537	0.173	0.090
Pooled - middle	400,537	0.646	0.128
Pooled - high	400,537	0.181	0.130
<i>Covariates</i>			
Female	400,537	0.509	0.500
Repeating grade 8	400,537	0.001	0.025
Regular age	400,537	0.912	0.283
Pre-school	357,010	0.953	0.212
Foreign student	400,537	0.096	0.295
Foreign student – first generation	400,537	0.033	0.178
Foreign student – second generation	400,537	0.063	0.243
Early starters	400,537	0.004	0.065
First quarter of birth	400,537	0.236	0.425
Second quarter of birth	400,537	0.252	0.434
Third quarter of birth	400,537	0.262	0.440
Fourth quarter of birth	400,537	0.250	0.433
<i>Students academic ability</i>			
<i>Liceo</i>	400,537	0.560	0.496
Average test scores (percentiles)	400,537	58.783	24.525
Class average test scores (percentiles)	400,537	54.343	9.219

Notes: Class share refers to the class share of occupation groups. Occupation group: low (unemployed, homemaker, retiree), middle (retailer, teacher, workman), high (manager, entrepreneur, professional). *Liceo* is the academic high school track in Italy. The lower number of observations for Pre-school program is due to missing data.

Table 2. Balance test.

	(1) Female	(2) Repeating grade 8	(3) Regular age	(4) Month of birth	(5) Early starter	(6) Pre-school	(7) Foreign citizen
Top – low share	0.0008 (0.0011) [0.947]	-0.0000 (0.0001) [1.000]	-0.0006 (0.0007) [0.283]	-0.0050 (0.0081) [1.000]	-0.0000 (0.0001) [1.000]	-0.0003 (0.0005) [1.000]	0.0040*** (0.0008) [0.001]
Top – high share	0.0014 (0.0015) [0.947]	-0.0000 (0.0001) [1.000]	0.0015 (0.0009) [0.283]	0.0053 (0.0109) [1.000]	0.0000 (0.0002) [1.000]	0.0003 (0.0007) [1.000]	-0.0052*** (0.0010) [0.001]
P-value joint significance	0.541	0.711	0.162	0.695	0.938	0.762	0
Observations	400,537	400,537	400,537	400,537	400,537	356,160	400,537
Mean of outcome	0.509	0.001	0.912	6.570	0.00430	0.953	0.0960
Clusters	3781	3781	3781	3781	3781	3600	3781

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects. Outcome variables are reported in column headings. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table 3. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*).

	(1) <i>Liceo</i>	(2) <i>Liceo</i>	(3) <i>Liceo</i>	(4) <i>Liceo</i>	(5) <i>Liceo</i>
Top – low share	-0.0075*** (0.0012)	-0.0044*** (0.0012) [0.001]	-0.0028** (0.0011) [0.007]	-0.0031** (0.0013) [0.008]	-0.0015 (0.0012) [0.100]
Top – high share	0.0254*** (0.0017)	0.0234*** (0.0017) [0.001]	0.0145*** (0.0018) [0.001]	0.0179*** (0.0018) [0.001]	0.0091*** (0.0016) [0.001]
P-value joint significance		0	0	0	0
Observations	400,537	400,537	400,537	398,004	398,004
Mean of outcome	0.560	0.560	0.560	0.560	0.560
Clusters	3781	3781	3781	3763	3763

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects. Clustered standard errors at school level are displayed in parentheses. In column (1) each coefficient comes from a separate regression of the outcome on each explanatory variable. Column (2) and (4) include citizen status, gender and parental education control. Column (3) and (4) include citizen status, gender and parental education control at average class level. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table 4. Sensitivity analysis on the effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*).

	(1) <i>Liceo</i>	(2) <i>Liceo</i>	(3) <i>Liceo</i>
Top - low share	-0.0029** (0.0012) [0.001]	-0.0028*** (0.0010) [0.004]	
Top - high share	0.0151*** (0.0015) [0.001]	0.0149*** (0.0013) [0.001]	
Mothers - low share			-0.0053*** (0.0013) [0.001]
Mothers - high share			0.0030** (0.0014) [0.017]
Fathers - low share			-0.0025** (0.0011) [0.017]
Fathers - high share			0.0114*** (0.0016) [0.001]
P-value joint significance	0	0	0
Observations	400,537	495,390	400,537
Mean of outcome	0.560	0.449	0.560
Clusters	3781	4424	3781

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects and includes citizen status, gender and parental education control. Clustered standard errors at school level are displayed in parentheses. Column (1) uses occupation group shares robust to measurement error. In Column (2) students who cannot be tracked in high school are coded not to be in *liceo*. Column (3) uses *mothers* and *fathers* shares (see Section 4.1). "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table 5. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by citizenship status and using only native shares.

Outcome variable: <i>Liceo</i>	(1) Italian citizen	(2) Foreign citizen	(3) All	(4) Italian citizen
Top - low share	-0.0023* (0.0012) [0.034]	-0.0067* (0.0037) [0.074]	-0.0018* (0.0011) [0.051]	-0.0013 (0.0012) [0.152]
Top - high share	0.0152*** (0.0016) [0.001]	0.0110* (0.0059) [0.074]	0.0128*** (0.0014) [0.001]	0.0130*** (0.0015) [0.001]
P-value joint significance	0	0.02	0	0
Observations	359,631	29,582	400,515	359,613
Mean of outcome	0.579	0.382	0.560	0.579
Clusters	3779	2863	3781	3779

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects and includes citizen status, gender and parental education control. Clustered standard errors at school level are displayed in parentheses. Outcome variable is reported in the upper left column heading. The sub-sample is reported in column headings. Columns (3) and (4) uses *top* shares computed only on Italian citizen students (see Section 5.3). "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. P-values from testing the null hypothesis of equal effects for immigrant and non-immigrant students: 0.2562 for the low-share coefficient, 0.4945 for the high-share coefficient. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table 6. Decomposition of parental occupations and parental education.

	(1) <i>liceo</i>	(2) <i>liceo</i>	(3) <i>liceo</i>
Top - low share	-0.0031** (0.0013) [0.011]		-0.0023* (0.0014) [0.027]
Top - high share	0.0162*** (0.0018) [0.001]		0.0087*** (0.0019) [0.001]
Top - share low educated		-0.0058*** (0.0019) [0.001]	-0.0044** (0.0019) [0.014]
Top - share high educated		0.0183*** (0.0019) [0.001]	0.0150*** (0.0021) [0.001]
P-value joint significance	0	0	0
Observations	309,345	309,345	309,345
Mean of outcome	0.559	0.559	0.559
Clusters	3766	3766	3766

Notes: Each regression includes school-year-mother's occupation-father's occupation group- mother's education-father's education group fixed effects and includes citizen status. Top - share consider the highest education level within the household. Low: less than high school. Middle educated: high school. High educated: more than high school. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table 7. Effect of classmates' parental occupations on own ability and class ability.

	(1) Own average test score grade 8	(2) Peers' average test score grade 8
Top – low share	-0.0154 (0.0705) [1.000]	-0.7993*** (0.0672) [0.001]
Top – high share	-0.0542 (0.1072) [1.000]	1.8642*** (0.1055) [0.001]
P-value joint significance	0.867	0
Observations	400,537	400,537
Clusters	3781	3781

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects and includes citizen status, gender and parental education control, and measurement error robust occupation group shares. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table 8. Instrumental variable estimates and decomposition of the total effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*).

	(1) OLS Total effect <i>Liceo</i>	(2) OLS Naïve <i>Liceo</i>	(3) FS New peers' class average test score	(4) RF <i>Liceo</i>	(5) IV <i>Liceo</i>
<i>Panel A. Estimates of equation (3) and (4)</i>					
Average ESCS (partially overlapping peers in grade 5)			0.0352*** (0.011)	-0.0029 (0.0018)	
New peers' top – low share	-0.0009 (0.0012)	-0.0021* (0.0011)	-0.0885*** (0.0058)	-0.0009 (0.0011)	-0.0081 (0.0054)
New peers' top – high share	0.0104*** (0.0014)	0.0131*** (0.0014)	0.1830*** (0.0091)	0.0105*** (0.0014)	0.0254** (0.0110)
New peers' class average test score (<i>TS</i>)		-0.0143*** (0.0012)			-0.0812 (0.0598)
Kleibergen-Paap F statistic			10.07		
P-value for joint significance of "Top – low share" and "Top – high share"		0			0.001
<i>Panel B. Decomposition of "Top - high share"</i>					
Direct effect (γ_2)					0.0254
Indirect effect ($\delta_2 \times \gamma_3$)					-0.0148
Total effect					0.0104
Observations	385,668	385,668	385,668	385,668	385,668
Clusters	3736	3736	3736	3736	3736

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects and includes citizen status, gender and parental education control, and measurement error robust occupation group shares. Outcome variables are reported in column headings. The occupation group shares are computed on new peers. The instrument *Average ESCS* is computed considering the partially overlapping peers. *New peers' average test score grade 8* refers to the average percentile positions in standardized national test scores (reading and maths) averaged over class excluding student *i*. Clustered standard errors at school level are displayed in parentheses. The reported p-value corresponds to a joint test of the coefficients on top-low share and top-high share. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table 9. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in different high school tracks.

	(1) <i>Liceo</i>	(2) Scientific or classic	(3) Other types of <i>liceo</i>	(4) Technical or vocational	(5) Technical	(6) Vocational
Top – low share	-0.0029** (0.0012) [0.006]	-0.0009 (0.0011) [0.241]	-0.0020* (0.0011) [0.147]	0.0029** (0.0012) [0.006]	0.0011 (0.0011) [0.192]	0.0018** (0.0008) [0.018]
Top – high share	0.0151*** (0.0015) [0.001]	0.0167*** (0.0017) [0.001]	-0.0016 (0.0016) [0.189]	-0.0151*** (0.0015) [0.001]	-0.0116*** (0.0015) [0.001]	-0.0035*** (0.0010) [0.004]
P-value joint significance	0	0	0.130	0	0	0
Observations	400,537	400,537	400,537	400,537	400,537	400,537
Mean outcome	0.560	0.312	0.248	0.440	0.315	0.125
Clusters	3781	3781	3781	3781	3781	3781

Notes: Each regression includes school-year-mother's occupation-father's occupation fixed effects and includes citizen status, gender and parental education control, and measurement error robust occupation group shares. Outcome variables are reported in column headings. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table 10. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by provinces social mobility and by ESCS (socioeconomic background).

Outcome variable: <i>Liceo</i>	(1) Higher social mobility	(2) Lower social mobility	(3) Below median ESCS	(4) Above median ESCS
Top – low share	0.0000 (0.0018) [0.416]	-0.0051*** (0.0015) [0.004]	-0.0051*** (0.0019) [0.001]	-0.0010 (0.0019) [0.978]
Top – high share	0.0130*** (0.0023) [0.001]	0.0172*** (0.0021) [0.001]	0.0177*** (0.0026) [0.001]	0.0105*** (0.0022) [0.001]
P-value joint significance	0	0	0	0
Observations	198,726	201,811	155,332	165,802
Clusters	1939	1842	3737	3717

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects and includes citizen status, gender and parental education control, and measurement error robust occupation group shares. In Column (1) higher social mobility refers to be below the AUM rank. Column (4) refers to be above or at the median of ESCS. Outcome variable is reported in the upper left column heading. The sub-sample is reported in column headings. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. The p-values from testing the null hypothesis of equal effects across the two subsamples in each pair of columns are respectively: 0.031 and 0.1733; 0.119 and 0.026. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table 11. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by own parental occupations.

Outcome variable: <i>Liceo</i>	(1) Low occupation group	(2) Middle occupation group	(3) High occupation group
Top – low share	-0.0007 (0.0065) [0.836]	-0.0035*** (0.0013) [0.004]	-0.0014 (0.0022) [0.362]
Top – high share	0.0250** (0.0118) [0.073]	0.0162*** (0.0019) [0.001]	0.0124*** (0.0025) [0.001]
P-value joint significance	0.093	0	0
Observations	8,430	274,730	117,377
Mean outcome	0.398	0.501	0.710
Clusters	1835	3778	3665

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects and includes citizen status, gender and parental education control, and measurement error robust occupation group shares. Outcome variable is reported in the upper left column heading. The sub-sample is reported in column headings. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. P-values from testing the null hypothesis of equal effects across columns (1),(2) and (3) are: 0.6518 for the low-share coefficient, 0.2898 for the high-share coefficient. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p <0.01, ** p <0.05, and * p <0.10.

Online Appendix for

Peer Backgrounds Matter: Classmates' Parents and High School Track Choice

Filippo Da Re, University of Padova

A. Additional descriptive statistics

A1. Sample selection

The starting sample includes 661,217 students attending 4981 schools in the Center-North of Italy, grouped into 33,049 classes. Southern regions are excluded for several reasons, including teacher strikes (Meschi and Pavese, 2023), low reliability of test scores (Lucifora and Tonello, 2015), and longitudinal linkage issues (Bertoni and Parkham, 2025). I exclude 15,902 students with missing data on key variables and 2272 with inconsistent gender information across grades, following Pagani and Pica (2025). For the 88 students who repeat grade 8 and thus appear twice, only the most recent observation is retained (though both appearances contribute to peers' composition). I also exclude 11,440 students in classes with more than 30 or fewer than 10 students — the maximum and minimum class sizes generally allowed in secondary schools, following the rules on class formation in Italy (see Ministerial Decree no. 331/1998 and Presidential Decree no. 81/2009). Due to the lack of variation across classes, single-class schools do not contribute to the analysis (29,725 students). To avoid undue influence from outlier schools, I retain at most ten classes per school (2907 observations). These criteria result in a 9.5 percent reduction from the original sample, comparable with previous studies (Bertoni and Parkham, 2025). The empirical strategy further omits 18,048 observations belonging to singleton school-year-parental occupation cells (see Section 4.3) and 80,139 observations with missing information on both parents' occupations. Separately, 100,159 grade 8 students cannot be followed up to grade 10; following Merlino et al. (2019), I retain them when computing peer exposure, since the analysis targets contemporaneous peer effects, and because their inclusion reduces measurement error in class composition by making the data closely resemble a school level

census. The final sample includes 400,537 students in 27,283 classes across 3781 schools. Section 5.2 presents robustness checks addressing potential sample selection concerns and shows that, despite data limitations, the imposed selection criteria are unlikely to affect the results. Table A5 compares the starting sample and the estimation sample among students with non-missing demographic information. While the differences are modest, students in the estimation sample are likely to be female and native-born, and more likely to have attended preschool and to display higher ability. Finally, Table A6 investigates the characteristics of students that cannot be followed between middle school to high school. Comparing Tables A5 and A6, and consistent with Meschi and Pavese (2023), attriters are more likely to be male, foreign-born, and of lower ability. Table A7 empirically confirms this pattern by regressing individual traits on the probability of exiting the sample. These compositional differences should be borne in mind when interpreting the external validity of the findings.

As an additional check, Table A7 examines whether missingness in parental occupation correlates with observed student traits by regressing an indicator for both parental occupations being missing on a set of covariates. Students are more likely to have missing data if they have an irregular enrollment history, did not attend preschool, are from the Center regions, are in larger classes, or have lower ability. While these correlations indicate that the missing-at-random assumption underlying the correction in Section 5.2 does not hold exactly, the correction still improves on disregarding the issue entirely.

A.2 Occupation classification

Unemployment can be transitory and need not reflect a parent's usual occupational prestige: an unemployed parent may, for instance, be between prestigious jobs. The low category should therefore be interpreted as capturing whether a parent is currently not working, rather than as a stable indicator of low occupational prestige. This concern is mitigated by the use of the highest prestigious occupation across the two parents in the household: as long as at least one parent holds a non-transitory position, transitory unemployment of the other parent does not affect the household's assigned prestige category. Reassuringly, the share of unemployed parents in the INVALSI 2016 data (5.07%) closely tracks the 4.87% recorded in the Survey on Household Income and Wealth (SHIW) for the same year,

suggesting that this measure does not systematically over- or understate overall unemployment.

Residual misclassification — cases where the higher prestige parent is themselves temporarily unemployed — can still bias the estimates, and the direction of this bias depends on which of two scenarios holds. If such spells are largely transitory and uncorrelated with the parent's usual occupational standing, this constitutes classical measurement error. It dilutes the true prestige gradient by mixing genuinely low prestige households with households that, in substance, transmit high or middle prestige signals — attenuating the estimated low prestige peer effects toward zero and making my estimates conservative. Alternatively, unemployment spells — even short ones — may be correlated with a household-level shock (income loss, reduced social network, geographic mobility) that independently affects a family's capacity to transmit information or aspirations to the child's peers. In this case, the low category would partly capture the effect of the shock itself, and the direction of the resulting bias is a priori ambiguous. While the data at hand cannot fully adjudicate between these two explanations, the first is the standard benchmark case for a discrete proxy of an underlying persistent trait, implying that my estimates are, if anything, conservative.

A similar concern applies to retirees, who may have held prestigious occupations before retiring; since I have no information on parents' previous occupations, this could analogously bias the low category. To address this, I re-estimate equation (2) excluding retirees from the sample; the results are unchanged — unsurprising given the relatively young age of the sampled students, as retired parents represent only 2% of observations.

Tables

Table A1. Descriptive statistics for Italian labor market. (SHIW 2016).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Employed	Ever not employed	Log earnings	Work fulltime	Bonus	College	Log earnings
College	0.1280*** (0.0164)	-0.0811*** (0.0132)	0.3725*** (0.0391)	0.0708*** (0.0197)	0.1130*** (0.0251)		
<i>Liceo</i>						0.5135*** (0.0316)	0.1822*** (0.0418)
Observations	10,152	10,152	5,195	4,390	4,390	4,435	3,018

Notes: Each regression includes controls for gender, experience and geographic region. Standard errors are clustered at NACE level. Data comes from SHIW 2016 (The Survey on Households Income and Wealth) collected by the Bank of Italy. Estimates are comparable to the ones documented in Anelli (2020).

Table A2. Timing of the cohorts by academic year in the sample.

	Academic year	
	Cohort 1	Cohort 2
Grade		
5 th (elementary school)	2012-2013	2013-2014
8 th (middle school)	2015-2016	2016-2017
10 th (high school)	2017-2018	2018-2019

Table A3. List of occupations for each category available in INVALSI questionnaire and ISCO 1-digit skill.

Category name	Complete list of occupations (INVALSI)	ISCO 1-digit skill
Unemployed	Unemployed	
Homemaker	Homemaker	
Manager	Manager, University professor, Military officer	White-collar high-skilled
Entrepreneur	Entrepreneur, Farm holder	White-collar high-skilled
Retailer	Professional employee, Doctor, Lawyer, Psychologist, Researcher	White-collar high-skilled
Teacher	Teacher, Office worker	White-collar low-skilled
Workman	Workman, Service worker, Cooperative member	Blue-collar low-skilled
Retiree	Retiree	

Notes: The first column refers to the category occupations in INVALSI (see De Benedetto and De Paola, 2023). The second column provides the complete list of occupations for each category in INVALSI. The third column maps the occupation category onto the ISCO 1-digit level skill following the categorization of Dumont (2006).

Table A4. Definitions of peers.

	Peers	New Peers	Partially overlapping peers
Definitions	Classmates in grade 8	Classmates in grade 8 but in different classes in grade 5	Classmates in grade 5 of my peers (of grade 8) but not my classmates in grade 5.
Grade 5 (primary school)		<div>Class H</div> <div>a b c</div> <div>Class J</div> <div>d e f</div>	<div>Class H</div> <div>a b c</div> <div>Class J</div> <div>d e f</div>
Grade 8 (middle school)	<div>Class K</div> <div>a b c</div>	<div>Class K</div> <div>a b d</div>	<div>Class K</div> <div>a b d</div>
	a, b and c are <i>peers</i>	a and d are <i>new peers</i> b and d are <i>new peers</i>	e and f are <i>partially overlapping peers</i> of a

Table A5. Summary statistics of starting and final sample.

	Estimation sample			Starting sample		
	N	Mean	Standard deviation	N	Mean	Standard deviation
Female	400,537	0.509	0.500	645,315	0.490	0.500
Repeating grade 8	400,537	0.001	0.025	645,315	0.002	0.049
Regular age	400,537	0.912	0.283	645,315	0.866	0.340
Pre-school	357,010	0.953	0.212	535,261	0.928	0.259
Foreign student	400,537	0.096	0.295	645,315	0.130	0.336
Foreign student – first generation	400,537	0.033	0.178	645,315	0.051	0.221
Foreign student – second generation	400,537	0.063	0.243	645,315	0.078	0.269
Early starters	400,537	0.004	0.065	645,315	0.004	0.066
First quarter of birth	400,537	0.236	0.425	645,315	0.235	0.424
Second quarter of birth	400,537	0.252	0.434	645,315	0.251	0.433
Third quarter of birth	400,537	0.262	0.440	645,315	0.263	0.440
Fourth quarter of birth	400,537	0.250	0.433	645,315	0.251	0.434
Average test scores (percentiles)	400,537	58.782	24.525	645,315	53.591	25.878

Notes: The starting sample refers to 645,315 observations with information on basic demographic variables. The lower number of observations for Pre-school program is due to missing data.

Table A6. Summary statistics of attriters.

	N	Attriters	
		Mean	Standard deviation
Female	100,159	0.415	0.493
Repeating grade 8	100,159	0.008	0.089
Regular age	100,159	0.717	0.450
Pre-school	88,280	0.921	0.269
Foreign student	100,159	0.246	0.430
Foreign student – first generation	100,159	0.110	0.313
Foreign student – second generation	100,159	0.136	0.343
Early starters	100,159	0.003	0.051
First quarter of birth	100,159	0.228	0.420
Second quarter of birth	100,159	0.244	0.429
Third quarter of birth	100,159	0.270	0.444
Fourth quarter of birth	100,159	0.258	0.437
Average test scores (percentiles)	100,159	35.114	22.160

Notes: Attriters refers to the 100,159 observations who cannot be followed from grade 8 to grade 10. The lower number of observations for pre-school program is due to missing data.

Table A7. Missingness of occupations information and individual characteristics.

	Missing information on parental occupations	Probability of not being followed in grade 10
Female	-0.0004 (0.0007)	-0.0429*** (0.0012)
Repeating grade 8	0.0242** (0.0096)	0.2700*** (0.0175)
Regular age	-0.0002 (0.0019)	-0.1882*** (0.0026)
Pre-school	-0.2221*** (0.0152)	-0.0160*** (0.0036)
Foreign student	0.0049 (0.0032)	0.0871*** (0.0024)
Early starters	-0.0001 (0.0061)	-0.1907*** (0.0082)
Numbers of classes per school	-0.0003 (0.0013)	-0.0023*** (0.0004)
Number of students per school	0.0010** (0.0005)	-0.0032*** (0.0003)
Northeast	0.0030 (0.0050)	-0.0344*** (0.0022)
Center	0.0446*** (0.0076)	-0.0912*** (0.0023)
Average test scores (percentiles)	-0.0001*** (0.0000)	-0.0049*** (0.0000)
Observations	535,261	447,929
Clusters	4755	4330

Notes: The regressions include binary variables for students' gender, citizenship status, age, attendance of pre-school programs and anticipatory careers. It includes binary variable for school geographic location (North West as reference group) and continuous variable controlling for students' ability and planned number of classes per school and students per school. The regression includes cohort fixed effects and standard errors are clustered at school level. Outcome variable Column (1) having missing information on both parental occupations. Outcome variable Column (2) students that cannot be followed from grade 8 to grade 10.

B. Additional robustness tests

B.1 Specification choice

Table B1 reports how the coefficients on the peer-share variables evolve as fixed effects are progressively added to the model, complementing the binned scatter plots in Figure 5. Column (1) reports the raw, unconditional relationship with no fixed effects; column (2) adds school fixed effects; column (3) adds school-by-year fixed effects; column (4) reports the baseline specification with the full school-by-year-by-occupation group fixed effects; and column (5) adds individual level controls. The coefficients change substantially between columns (1) and (2), reflecting the confounding influence of between-school sorting on family background: the low-share coefficient is small and insignificant in the raw data but becomes negative and significant once school fixed effects are included, while the high-share coefficient falls by roughly three-quarters. From column (2) onward, the estimates are comparatively stable, indicating that most of the identifying variation is absorbed once between-school differences are accounted for, with the progressively finer fixed effects and individual controls refining rather than overturning the baseline result.

B.2 Threat to identification strategy and other validation tests

As the main identifying assumption concerns the quasi-random allocation of students across classes within a school cohort, I present additional evidence to corroborate this hypothesis. Following Goulas et al. (2024), I regress mean class level predetermined characteristics on classroom numbers (i.e., class 1, class 2) on school-by-year fixed effects and on dummies for the within school-cohort class progressive identifier. This test detects whether systematic differences exist across classes within a school in the same year. The results in Table B2 demonstrate that classroom number is not systematically associated with class characteristics. Furthermore, as in by Pagani and Pica (2025), I run 10,000 replications of such class level regressions randomly reassigning them within school-cohort progressive class identifier. Figure B1 shows, for each observable characteristic, the distribution of the p-values of the F-test for joint significance of the class dummies in the 10,000 regression replications. In the case of random assignment, the null hypothesis that the coefficients on the class dummies are jointly zero at the 5 percent significance level should be rejected in

approximately 5 percent of cases. Figure B1 confirms that this is always true, except for *repeating grade 8* variable, which is rejected in 10 percent cases.

A potential concern is that students may be non-randomly clustered by parental background within school-cohort - for instance, if principals deliberately group students by socioeconomic status. In this case, one's parental occupation group would be highly correlated to their peers' one. To investigate this, I randomly reallocate students to classes within school-cohort, keeping the existing class sizes, and I replicate the exercise 100 times: then, I compare the distribution of the occupation group shares in the actual sample and in the randomly reallocated samples (see Pagani and Pica, 2025). If students are not grouped by parental background, the observed and simulated distributions should be similar. Figure B2 displays the densities of occupation group shares from the actual data and from these 100 random reallocations. The two distributions largely overlap, providing evidence of no sorting by parental background.

Most studies on peer effects exploit idiosyncratic variation across cohorts within the same school (Hoxby, 2000). I do not adopt this strategy for two reasons. First, this approach is most suitable for panel data covering many cohorts per school, where it is reasonable to assume that demographic shifts generate quasi-random variation in peer composition. My data include only two cohorts reducing variability over time. Second, focusing on within-school, within-cohort variation allows me to compare students exposed to the same institutional environment and local business cycle—an essential consideration when parental occupations serve both as a measure and as a mechanism—while exploiting class level variation, which represents the relevant peer network. For completeness, Table B2 (Panel A) presents balance tests using a within-school, across-cohort specification. The results are comparable to the baseline estimates in Table 2. As suggested in Bietenbeck (2025), I perform a balance test by estimating separate regressions for each peer variable included in the main model, which is particularly appropriate when multiple peer measures are of interest. Table B3 (Panel B) reports these results. Most coefficients are close to zero and statistically insignificant, confirming the absence of systematic relationships between predetermined student characteristics and occupation group shares.

As an additional robustness check, Table B3 reports the results of a test proposed by Porreca et al. (2026), which involves regressing peer composition variables as dependent variable on students' observable characteristics to detect any residual correlation.

As a complement to the balance tests presented above, I examine the magnitude of differences in occupation group shares between classes within the same school and cohort. Systematic sorting of students by parental occupation across classes would produce large differences in these shares. To examine this, I collapse the data to the school by class by year level and, for each school-cohort, compute the pairwise difference in each occupation group share between every pair of classes (see Merlino et al. 2019). Since classes within a school-cohort have no natural ordering, the sign of each pairwise difference is arbitrary, and I therefore focus on the magnitude of the differences rather than on their distribution around zero. Figure B3 plots these distributions, and the mean absolute change in the low (high) share is 0.040 (0.124), with signed mean close to zero (-0.0011 and 0.0063 , respectively). These magnitudes are consistent with the balance tests reported above.

B.3 Validity of the instrument

The instrument used in Section 6.2 is the average ESCS in the grade 5 classes of i 's partially overlapping peers. The exclusion restriction requires that this variable affects i 's high school track choice only through the ability of i 's grade 8 peers. It would fail if the socioeconomic composition of the partially overlapping peer group were correlated with unobserved determinants of i 's track choice through other channels. This could occur, for example, if primary schools in wealthier neighborhoods systematically fed into the same middle schools, so that the ESCS of the grade 5 classmates of i 's peers partly reflected local socioeconomic conditions that also shaped i 's educational aspirations directly.

As reported in Section 6.2, three features of the empirical design mitigate this concern. First, the school-by-year fixed effects absorb all time-varying school level confounders, including the socioeconomic composition of the catchment area. Second, the instrument varies at the individual level, since each student's partially overlapping peer group is unique, rather than at the school or neighborhood level, which limits the scope for area-level confounding. Third, the instrument is based on grade 5 class composition, which is itself subject to the

same heterogeneous class formation rules that govern middle school assignment²³. While these features do not eliminate all possible threats, they substantially narrow the set of confounders that could violate the exclusion restriction.

I further assess the conditional independence assumption underlying the instrument by testing for conditional random assignment. In the spirit of McConnell et al. (2025), I regress the instrument on the full set of grade 8 demographic controls; results are reported in Table B5. Of the six covariates, only regular age enters significantly, and its coefficient is small in absolute terms. The remaining coefficients are close to zero and imprecisely estimated. The demographic controls are jointly significant at the 10 percent level but not at conventional levels ($p = 0.080$), thus it establishes that any residual association between the instrument and observable student characteristics is too small to generate meaningful bias, which is the relevant condition here.

²³ In addition, school principals do not observe prior ability before primary school (see Pagani and Pica, 2025).

Tables

Table B1. Coefficients on peer parental occupation shares across progressively richer fixed-effect specifications.

	(1) <i>Liceo</i>	(2) <i>Liceo</i>	(3) <i>Liceo</i>	(4) <i>Liceo</i>	(5) <i>Liceo</i>
Top – low share	0.0019 (0.0016)	-0.0060*** (0.0011)	-0.0051*** (0.0012)	-0.0044*** (0.0012)	-0.0028** (0.0011)
Top – high share	0.0988*** (0.0019)	0.0247*** (0.0016)	0.0231*** (0.0018)	0.0234*** (0.0017)	0.0145*** (0.0018)
P-value joint significance	0	0	0	0	0
Type of fixed effects	None	School	School × year	Baseline	Baseline + controls
Observations	400,537	400,537	400,537	400,537	400,537
Mean of outcome	0.560	0.560	0.560	0.560	0.560
Clusters	3781	3781	3781	3781	3781

Notes: Column (1) reports the raw correlation with no fixed effects. Column (2) adds school fixed effects. Column (3) adds school-by-year fixed effects. Column (4) reports baseline specification, which absorbs own-parental-occupation-background × school × year fixed effects. Column (5) adds individual level controls (citizen status, gender, and parental education) to the baseline specification. Standard errors are clustered at the school level and reported in parentheses. "P-value joint significance" reports the p-value for the joint significance of the two peer-share coefficients in each column. *** p<0.01, ** p<0.05, * p<0.10.

Table B2. Balance test at class level.

	(1) % Female	(2) % Repeating grade 8	(3) % Regular age	(4) % Early starters	(5) % Pre- school	(6) % Foreign pupil
Class 2	0.0019 (0.0018)	-0.0001 (0.0002)	-0.0025* (0.0014)	0.0005* (0.0003)	-0.0001 (0.0011)	-0.0005 (0.0015)
Class 3	0.0031 (0.0022)	-0.0000 (0.0002)	0.0016 (0.0015)	0.0002 (0.0003)	0.0011 (0.0013)	-0.0008 (0.0017)
Class 4	-0.0014 (0.0024)	-0.0001 (0.0003)	0.0025 (0.0018)	-0.0000 (0.0003)	0.0002 (0.0015)	-0.0014 (0.0020)
Class 5	-0.0030 (0.0030)	0.0001 (0.0003)	-0.0014 (0.0021)	0.0000 (0.0004)	0.0014 (0.0017)	0.0002 (0.0024)
Class 6	0.0045 (0.0034)	-0.0002 (0.0004)	0.0025 (0.0026)	-0.0002 (0.0005)	0.0018 (0.0021)	0.0015 (0.0030)
Class 7	-0.0035 (0.0047)	-0.0004 (0.0004)	-0.0012 (0.0037)	-0.0003 (0.0007)	-0.0000 (0.0031)	-0.0050 (0.0037)
Class 8	-0.0070 (0.0059)	-0.0004 (0.0005)	0.0037 (0.0043)	0.0003 (0.0011)	0.0001 (0.0051)	0.0039 (0.0056)
Class 9	0.0096 (0.0089)	-0.0017*** (0.0006)	-0.0008 (0.0063)	-0.0000 (0.0014)	-0.0027 (0.0076)	0.0047 (0.0064)
Class 10	-0.0061 (0.0099)	0.0004 (0.0009)	0.0032 (0.0082)	-0.0014 (0.0015)	0.0001 (0.0089)	0.0060 (0.0104)
Observations	27,283	27,283	27,283	27,283	25,310	27,283
Model p-value	0.165	0.261	0.122	0.723	0.989	0.813
Cluster	3781	3781	3781	3781	3620	3781

Notes: The table presents the estimated coefficients of binary indicators for different class numbers on a variety of outcomes. For instance, Class 1 is a binary indicator that takes the value of 1 if the class average of the relevant variable comes from class 1 and 0 otherwise. Class number 10 is omitted from the regression as the reference category. The unit of observation is the class. Each regression includes school-by-year fixed effects. Outcome variables are reported in column headings and have been averaged at class level. F-statistics for the joint significance of the regressors and the related P-value are also reported. Clustered standard errors at school level are displayed in parentheses. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table B3. Balance test following Hoxby method, panel A. Balance test for each single regressor (see Bietenbeck, 2025), panel B.

	(1) Female	(2) Repeating grade 8	(3) Regular age	(4) Month of birth	(5) Early starters	(6) Pre-school	(7) Foreign pupil
<hr/> Panel A <hr/>							
Top – low share	0.0004 (0.0010) [0.264]	-0.0000 (0.0001) [1.000]	-0.0008 (0.0007) [0.528]	-0.0042 (0.0078) [0.440]	-0.0001 (0.0001) [0.793]	-0.0005 (0.0005) [0.177]	0.0044*** (0.0008) [0.001]
Top – high share	0.0019 (0.0014) [0.264]	-0.0000 (0.0001) [1.000]	0.0015 (0.0009) [0.528]	0.0018 (0.0104) [0.440]	0.0002 (0.0002) [0.793]	0.0002 (0.0007) [0.465]	-0.0057*** (0.0010) [0.001]
P-value joint significance	0.406	0.863	0.116	0.834	0.591	0.609	0
Observations	400,537	400,537	400,537	400,537	400,537	356,981	400,537
Clusters	3781	3781	3781	3781	3781	3616	3781
<hr/> Panel B <hr/>							
Top – low share	0.0006 (0.0011)	-0.0000 (0.0001)	-0.0008 (0.0007)	-0.0055 (0.0080)	-0.0000 (0.0001)	-0.0004 (0.0005)	0.0046*** (0.0008)
Top – high share	0.0013 (0.0015)	-0.0000 (0.0001)	0.0016 (0.0009)	0.0063 (0.0108)	0.0000 (0.0002)	0.0003 (0.0007)	-0.0059*** (0.0010)
Observations	400,537	400,537	400,537	400,537	400,537	356,981	400,537
Clusters	3781	3781	3781	3781	3781	3616	3781

Notes: Panel A: Each regression includes school-mother's occupation-father's occupation group fixed effects and year fixed effect. Outcome variables are reported in column headings. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. Panel B: each estimate comes from a separate regression of the regressor in the first column on the outcome variables reported in column headings. Clustered standard errors at school level are displayed in parentheses. *** p < 0.01, ** p < 0.05, and * p < 0.10. Dependent variables are displayed in the first row.

Table B4. Balance test. Covariates as independent variable (see Porreca et al., 2026).

	(1) Top - low share	(2) Top - high share
Female	0.0012 (0.0024)	0.0021 (0.0019)
Repeating grade 8	-0.0366 (0.0563)	-0.0244 (0.0379)
Regular age	0.0038 (0.0053)	0.0030 (0.0040)
Pre-school	-0.0002 (0.0004)	0.0001 (0.0003)
Month of birth	-0.0178 (0.0199)	0.0098 (0.0179)
Early starters	-0.0020 (0.0108)	-0.0010 (0.0086)
Foreign pupil	0.0289*** (0.0059)	-0.0213*** (0.0042)
Observations	356,981	356,981
Clusters	3599	3599

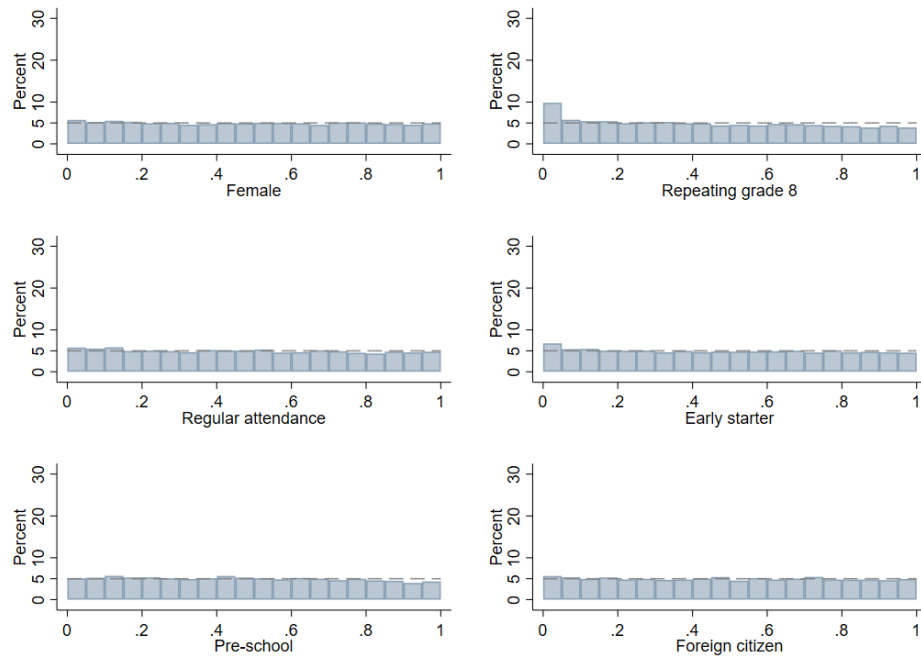
Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects. Outcome variables are reported in column headings. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table B5. Testing for conditional random assignment of the instrument

	(1) Average ESCS partially overlapping peer groups
Female	0.0002 (0.0005)
Repeating grade 8	-0.0012 (0.0101)
Regular age	0.0033*** (0.0010)
Early starters	0.0018 (0.0049)
Foreign pupil - first generation	0.0014 (0.0019)
Foreign pupil - second generation	0.0017 (0.0013)
P-value for joint test	0.0803
Observations	385,668
Clusters	3736

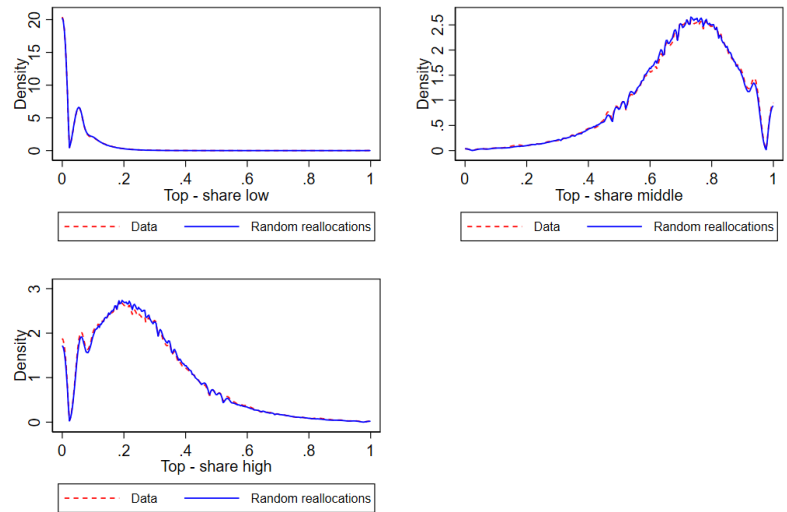
Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects. Outcome variables are reported in column headings and have been averaged at class level. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Figure B1. P-values of F-tests for joint significance of class dummies.



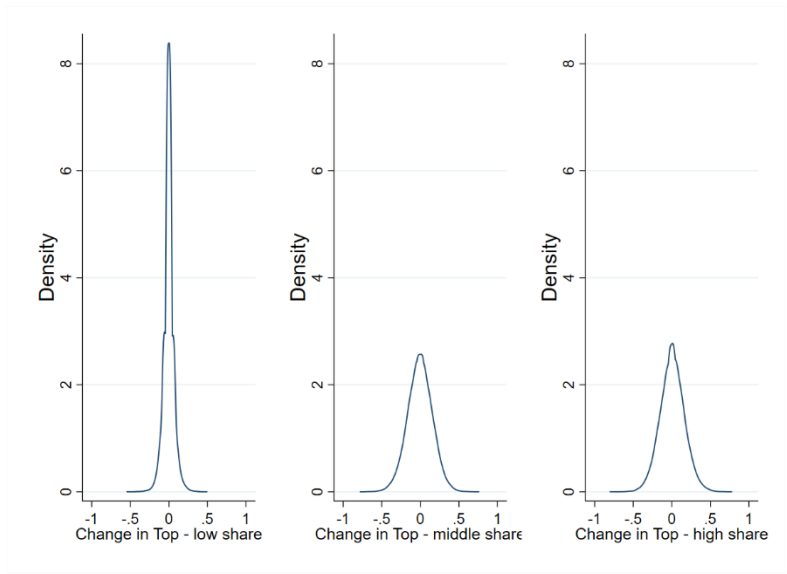
Notes: Each panel shows the distribution of the p-values of the F-tests for joint significance of the class dummies from 10,000 replications of class level regressions of class characteristics on class dummies and school-by-year fixed effects. Class dummies are obtained by randomly ranking classes within schools, and then assigning each class a dummy according to the within-school ranking. In each of the 10,000 replications, classes are randomly re-ranked within schools.

Figure B2. Densities of occupation group class shares from actual data and random reallocation of students within school-cohort.



Notes: Each panel shows the density of the occupation groups shares both from actual data and from 100 random reallocation of students within schools, keeping the existing class sizes constant. The bottom panel does the same for the share of high achieving girls.

Figure B3. Kernel density of change in occupation group class share.



Notes: Each panel shows the kernel density of change in occupation group class share.

C. Heterogeneous effects

Tables

Table C1. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by gender.

Outcome variable: <i>Liceo</i>	(1)	(2)	(3)	(4)	(5)	(6)
	All Females	All Males	Italian pupil Females	Italian pupil Males	Foreign pupil Females	Foreign pupil Males
Top – low share	-0.0044*** (0.0017) [0.005]	-0.0012 (0.0017) [0.313]	-0.0038** (0.0018) [0.014]	-0.0002 (0.0018) [0.900]	-0.0080 (0.0062) [0.255]	-0.0143*** (0.0056) [0.015]
Top – high share	0.0138*** (0.0022) [0.001]	0.0169*** (0.0026) [0.001]	0.0134*** (0.0023) [0.001]	0.0172*** (0.0023) [0.001]	0.0156 (0.0102) [0.255]	0.0097 (0.0095) [0.192]
P-value joint significance	0	0	0	0	0.085	0.015
Observations	194,743	187,570	173,759	168,836	12,858	10,978
Mean outcome	0.676	0.439	0.697	0.458	0.493	0.257
Clusters	3781	3779	3776	3772	2083	2036

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects and includes citizen status, gender and parental education control. Clustered standard errors at school level are displayed in parentheses. Outcome variable is reported in the upper left column heading. The sub-sample is reported in column headings. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. The p-values from testing the null hypothesis of equal effects across the two subsamples in each pair of columns are respectively: 0.184 and 0.306; 0.161 and 0.238; 0.467 and 0.678. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table C2. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by gender, both parents shares.

Outcome variable: liceo	(1) Females	(2) Males
Mothers – low share	-0.0052*** (0.0018) [0.009]	-0.0049*** (0.0019) [0.014]
Mothers – high share	0.0044** (0.0020) [0.020]	0.0027 (0.0021) [0.108]
Fathers – low share	-0.0033** (0.0016) [0.023]	-0.0026 (0.0016) [0.077]
Mothers – high share	0.0083*** (0.0022) [0.001]	0.0144*** (0.0023) [0.001]
P-value joint significance	0	0
Observations	194,743	187,570
Clusters	3781	3779

Notes: Each regression includes school-year-mother's occupation-father's occupation group fixed effects and includes citizen status, gender and parental education control and uses *mothers* and *fathers* shares (see Section 4.1). Clustered standard errors at school level are displayed in parentheses. Outcome variable is reported in the upper left column heading. The sub-sample is reported in column headings. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. *** p < 0.01, ** p < 0.05, and * p < 0.10.

D. Alternative shares definitions

For the sake of consistency, in this section I present the estimates for each table in the main text and in appendix B and C, using other three occupation group definitions: namely, the share of *mothers*, *fathers* and *parents* in each occupation group, which are computed with the following procedures:

- i) *Mothers' share*: based only on mothers' occupations.
- ii) *Fathers' share*: based only on fathers' occupations.
- iii) *Pooled share*: total number of parents in group o divided by the number of parents observed (students may contribute one or two parents depending on data availability).

Table

Table D1. Balance test. Other class shares.

	(1) Female	(2) Repeating grade 8	(3) Regular age	(4) Month of birth	(5) Early starter	(6) Pre-school	(7) Foreign citizen
Panel A							
Mothers – low share	-0.0017 (0.0013) [0.327]	0.0000 (0.0001) [1.000]	-0.0003 (0.0008) [1.000]	0.0033 (0.0090) [1.000]	0.0000 (0.0002) [0.650]	-0.0001 (0.0006) [0.865]	0.0015* (0.0009) [0.097]
Mothers – high share	0.0016 (0.0014) [0.327]	-0.0000 (0.0001) [1.000]	0.0002 (0.0008) [1.000]	0.0015 (0.0100) [1.000]	0.0003 (0.0002) [0.295]	0.0008 (0.0006) [0.643]	-0.0015* (0.0008) [0.097]
P-value joint significance	0.110	0.809	0.808	0.932	0.286	0.366	0.021
Panel B							
Fathers – low share	-0.0009 (0.0011) [0.701]	-0.0000 (0.0001) [1.000]	-0.0005 (0.0007) [1.000]	-0.0089 (0.0079) [1.000]	-0.0001 (0.0001) [1.000]	-0.0009* (0.0005) [0.229]	0.0040*** (0.0008) [0.001]
Fathers – high share	0.0016 (0.0014) [0.701]	0.0000 (0.0001) [1.000]	0.0008 (0.0009) [1.000]	-0.0029 (0.0106) [1.000]	0.0001 (0.0003) [1.000]	-0.0001 (0.0007) [0.724]	-0.0046*** (0.0009) [0.001]
P-value joint significance	0.357	0.959	0.505	0.522	0.793	0.243	0
Panel C							
Pooled – low share	-0.0020 (0.0013) [0.252]	0.0000 (0.0001) [1.000]	-0.0005 (0.0008) [0.997]	-0.0023 (0.0091) [1.000]	0.0000 (0.0002) [0.291]	-0.0005 (0.0006) [1.000]	0.0030*** (0.0009) [0.001]
Pooled – high share	0.0020 (0.0016) [0.252]	-0.0000 (0.0001) [1.000]	0.0007 (0.0010) [0.997]	-0.0008 (0.0117) [1.000]	0.0003 (0.0003) [0.291]	0.0003 (0.0008) [1.000]	-0.0042*** (0.0010) [0.001]
P-value joint significance	0.079	0.953	0.572	0.968	0.511	0.511	0
Observations	400,537	400,537	400,537	400,537	400,537	356,160	400,537
Clusters	3781	3781	3781	3781	3781	3600	3781

Notes: Each regression includes school-year-mother's occupation-father's occupation fixed effects. Outcome variables are reported in column headings. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. The corresponding table in the text is Table 2. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table D2. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) and sensitivity tests. Other class shares.

	(1) <i>Liceo</i>	(2) <i>Liceo</i>	(3) <i>Liceo</i>	(4) <i>Liceo</i>	(5) <i>Liceo</i>	(6) <i>Liceo</i>
Panel A						
Mothers – low share	-0.0099*** (0.0014) [0.001]	-0.0058*** (0.0013) [0.001]	-0.0060*** (0.0015) [0.001]	-0.0020 (0.0013) [0.067]	-0.0059*** (0.0013) [0.001]	-0.0064*** (0.0011) [0.001]
Mothers – high share	0.0112*** (0.0015) [0.001]	0.0067*** (0.0013) [0.001]	0.0088*** (0.0015) [0.001]	0.0043*** (0.0013) [0.003]	0.0066*** (0.0013) [0.001]	0.0067*** (0.0012) [0.001]
P-value joint significance	0	0	0	0	0	0
Panel B						
Fathers – low share	-0.0051*** (0.0012) [0.001]	-0.0029*** (0.0011) [0.005]	-0.0038*** (0.0012) [0.001]	-0.0016 (0.0011) [0.079]	-0.0030*** (0.0011) [0.004]	-0.0034*** (0.0010) [0.001]
Fathers – high share	0.0212*** (0.0017) [0.001]	0.0132*** (0.0015) [0.001]	0.0158*** (0.0017) [0.001]	0.0081*** (0.0016) [0.001]	0.0131*** (0.0015) [0.001]	0.0136*** (0.0013) [0.001]
P-value joint significance	0	0	0	0	0	0
Panel C						
Pooled – low share	-0.0105*** (0.0014) [0.001]	-0.0060*** (0.0013) [0.001]	-0.0077*** (0.0015) [0.001]	-0.0028** (0.0015) [0.019]	-0.0062*** (0.0013) [0.001]	-0.0067*** (0.0011) [0.001]
Pooled – high share	0.0218*** (0.0017) [0.001]	0.0135*** (0.0016) [0.001]	0.0177*** (0.0018) [0.001]	0.0090*** (0.0016) [0.001]	0.0133*** (0.0016) [0.001]	0.0137*** (0.0014) [0.001]
P-value joint significance	0	0	0	0	0	0
Observations	400,537	400,537	398,004	398,004	400,537	495,390
Mean of outcome	0.560	0.560	0.560	0.560	0.560	0.449
Clusters	3781	3781	3763	3763	3781	4424

Notes: Each regression includes school-year-mother's occupation-father's occupation fixed effects. Clustered standard errors at school level are displayed in parentheses. Column (2) and (4) include citizen status, gender and parental education control. Column (3) and (4) include citizen status, gender and parental education control at class level. Column (4) uses occupation group shares robust to measurement error. In Column (6) students who cannot be tracked in high school are coded not to be in *liceo*. Column (7). The lower number of observations is due to longitudinal linkage issue to primary school. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. The corresponding tables in the text are Tables 3 and 4. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table D3. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by immigrant status and using only-native shares. Other class shares.

	(1)	(2)	(3)	(4)
Outcome variable: <i>liceo</i>	Italian citizen	Foreign citizen	All	Italian citizen
Panel A				
Mothers – low share	-0.0062*** (0.0014) [0.001]	-0.0014 (0.0046) [1.000]	-0.0051*** (0.0013) [0.001]	-0.0051*** (0.0013) [0.001]
Mothers – high share	0.0068*** (0.0014) [0.001]	0.0021 (0.0054) [1.000]	0.0058*** (0.0013) [0.001]	0.0063*** (0.0013) [0.001]
P-value joint significance	0	0.854	0	0
Panel B				
Fathers – low share	-0.0030** (0.0012) [0.008]	-0.0008 (0.0036) [0.699]	-0.0017* (0.0011) [0.060]	-0.0020* (0.0012) [0.049]
Fathers – high share	0.0130*** (0.0016) [0.001]	0.0092* (0.0054) [0.222]	0.0118*** (0.0014) [0.001]	0.0117*** (0.0015) [0.001]
P-value joint significance	0	0.201	0	0
Panel C				
Pooled – low share	-0.065*** (0.0014) [0.001]	-0.0016 (0.0046) [0.623]	-0.0050*** (0.0013) [0.001]	-0.0053*** (0.0013) [0.001]
Pooled – high share	0.0133*** (0.0016) [0.001]	0.0079 (0.0060) [0.623]	0.0120*** (0.0015) [0.001]	0.0122*** (0.0016) [0.001]
P-value joint significance	0	0.328	0	0
Observations	359,631	29,582	400,515	359,613
Mean of outcome	0.579	0.382	0.560	0.579
Clusters	3779	2863	3781	3779

Notes: Each regression includes school-year-mother's occupation-father's occupation fixed effects and includes citizen status, gender and parental education control. Clustered standard errors at school level are displayed in parentheses. Outcome variable is reported in the upper left column heading. The sub-sample is reported in column headings. Column (3) includes primary school own ability and primary school class average ability. The lower number of observations is due to longitudinal linkage issue to primary school. Columns (3) and (4) uses *top* shares computed only on Italian citizen students (see Section 5.3). "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. The corresponding table in the text is Table 5. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table D4. Effect of classmates' parental occupations on own ability and class ability. Other class shares.

	(1) Own average test score grade 8	(2) Peers' average test score grade 8
<hr/> Panel A <hr/>		
Mothers – low share	-0.1658** (0.0793) [0.079]	-1.5073*** (0.0775) [0.001]
Mothers – high share	-0.0354 (0.0913) [0.537]	0.6843*** (0.0881) [0.001]
P-value joint significance	0.112	0
<hr/> Panel B <hr/>		
Fathers – low share	-0.0107 (0.0705) [1.000]	-0.6801*** (0.0666) [0.001]
Fathers – high share	-0.0431 (0.1021) [1.000]	-1.5953*** (0.0987) [0.001]
P-value joint significance	0.911	0
<hr/> Panel C <hr/>		
Pooled – low share	-0.1522* (0.0815) [0.141]	-1.6100*** (0.0776) [0.001]
Pooled – high share	0.0724 (0.1121) [0.350]	1.4542*** (0.1085) [0.001]
P-value joint significance	0.168	0
Observations	400,537	400,537
Clusters	3781	3781

Notes: Each regression includes school-year-mother's occupation-father's occupation fixed effects and includes citizen status, gender and parental education control, and measurement error robust occupation group shares. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. The corresponding table in the text is Table 6. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table D5. Effect of classmates' parental occupations on the probability of enrolling in different high school tracks. Other class shares.

	(1) <i>Liceo</i>	(2) Scientific or classic	(3) Other types of <i>liceo</i>	(4) Technical or vocational	(5) Technical	(6) Vocational
Panel A						
Mothers – low share	-0.0059*** (0.0013) [0.001]	-0.0042*** (0.0013) [0.001]	-0.0017 (0.0012) [0.391]	0.0059*** (0.0013) [0.001]	0.0029** (0.0013) [0.017]	0.0030*** (0.0009) [0.002]
Mothers – high share	0.0066*** (0.0013) [0.001]	0.0081*** (0.0014) [0.001]	-0.0014 (0.0013) [0.391]	-0.0066*** (0.0013) [0.001]	-0.0045*** (0.0013) [0.001]	-0.0021 (0.0008) [0.006]
P-value joint significance	0	0	0.268	0	0	0
Panel B						
Fathers – low share	-0.0030*** (0.0011) [0.004]	-0.0013 (0.0011) [0.115]	-0.0017 (0.0010) [0.149]	0.0030*** (0.0011) [0.004]	0.0017 (0.0011) [0.071]	0.0013* (0.0008) [0.081]
Fathers – high share	0.0131*** (0.0015) [0.001]	0.0153*** (0.0016) [0.001]	-0.0022 (0.0015) [0.149]	-0.0130*** (0.0015) [0.001]	-0.0110*** (0.0014) [0.001]	-0.0020** (0.0010) [0.081]
P-value joint significance	0	0	0.120	0	0	0.013
Panel C						
Pooled – low share	-0.0062*** (0.0013) [0.001]	-0.0037*** (0.0013) [0.003]	-0.0025** (0.0012) [0.081]	0.0062*** (0.0013) [0.001]	0.0027*** (0.0014) [0.023]	0.0035*** (0.0009) [0.001]
Pooled – high share	0.0133*** (0.0016) [0.001]	0.0160*** (0.0017) [0.001]	-0.0028* (0.0016) [0.081]	-0.0133*** (0.0016) [0.001]	-0.0108*** (0.0015) [0.001]	-0.0025** (0.0010) [0.007]
P-value joint significance	0	0	0.051	0	0	0
Observations	400,537	400,537	400,537	400,537	400,537	356,160
Mean of outcome	0.560	0.312	0.248	0.440	0.315	0.125
Clusters	3781	3781	3781	3781	3781	3600

Notes: Each regression includes school-year-mother's occupation-father's occupation fixed effects and includes citizen status, gender and parental education control, and measurement error robust occupation group shares. Outcome variables are reported in column headings. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. The corresponding table in the text is Table 8. *** p < 0.01, ** p < 0.05, and * p < 0.10.

Table D6. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by provinces social mobility and by ESCS (socioeconomic background), and own parental occupations. Other class shares.

Outcome variable: <i>Liceo</i>	(1) Higher social mobility	(2) Lower social mobility	(3) Below median ESCS	(4) Above median ESCS	(5) Low occupation group	(6) Middle occupation group	(7) High occupation group
Panel A							
Mothers – low share	-0.0033* (0.0019) [0.042]	-0.0085*** (0.0018) [0.001]	-0.0047** (0.0022) [0.001]	-0.0078*** (0.0020) [0.001]	-0.0045 (0.0094) [0.463]	-0.0063*** (0.0015) [0.001]	-0.0045* (0.0023) [0.037]
Mothers – high share	0.0068*** (0.0020) [0.002]	0.0064*** (0.0018) [0.001]	0.0089*** (0.0023) [0.001]	0.0036** (0.0018) [0.022]	0.0160 (0.0106) [0.369]	0.0076*** (0.0016) [0.001]	0.0046** (0.0019) [0.037]
P-value joint significance	0	0	0	0	0.197	0	0.001
Panel B							
Fathers – low share	-0.0029* (0.0017) [0.048]	-0.0030** (0.0015) [0.022]	-0.0039** (0.0018) [0.015]	-0.0033* (0.0018) [0.035]	-0.0009 (0.0068) [0.798]	-0.0040*** (0.0013) [0.002]	-0.0006 (0.0021) [0.655]
Fathers – high share	0.0103*** (0.0022) [0.001]	0.0158*** (0.0020) [0.001]	0.0159*** (0.0024) [0.001]	0.0095*** (0.0021) [0.001]	0.0205* (0.0112) [0.155]	0.0141*** (0.0018) [0.001]	0.0106*** (0.0022) [0.001]
P-value joint significance	0	0	0	0	0.167	0	0
Panel C							
Pooled – low share	-0.0041*** (0.0019) [0.019]	-0.0081*** (0.0018) [0.001]	-0.0056*** (0.0022) [0.005]	-0.0080*** (0.0020) [0.001]	-0.0022 (0.0091) [0.675]	-0.0070*** (0.0015) [0.001]	-0.0041* (0.0024) [0.045]
Pooled – high share	0.0115*** (0.0023) [0.001]	0.0149*** (0.0021) [0.001]	0.0171*** (0.0027) [0.001]	0.0085*** (0.0022) [0.001]	0.0246** (0.0118) [0.081]	0.0146*** (0.0019) [0.001]	0.0102*** (0.0023) [0.001]
P-value joint significance	0	0	0	0	0.079	0	0
Observations	198,726	201,811	155,332	165,332	8,430	356,160	400,537
Mean outcome					0.398	0.501	0.710
Clusters	1939	1842	3737	3737	1835	3778	3665

Notes: Each regression includes school-year-mother's occupation-father's occupation fixed effects and includes citizen status, gender and parental education control, and measurement error robust occupation group shares. In Column (1) higher social mobility refers to be below the AUM rank. Column (4) refers to be above or at the median of ESCS. Outcome variable is reported in the upper left column heading. The sub-sample is reported in column headings. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the two coefficients reported in the same column. Sharpened False Discovery Rate-adjusted q-values are reported in squared brackets. The corresponding tables in the text are Tables 9 and 10. *** p < 0.01, ** p < 0.05, and * p < 0.10.

References

- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- Bertoni, M., & Parkham, S. (2025). When do ordinal ability rank effects emerge? Evidence from the timing of school closures. *Economics of Education Review*, 105, 102633.
- Bietenbeck, J. (2025). Do motivated classmates matter for educational success?. *The Economic Journal*, 135(665), 36-58.
- De Benedetto, M. A., & De Paola, M. (2023). Immigration and teacher bias towards students with an immigrant background. *Economic Policy*, 38(113), 107-154.
- Dumont, M. (2006). The reliability–or lack thereof–of data on skills. *Economics Letters*, 93(3), 348-353.
- Goulas, S., Gunawardena, B. N., Megalokonomou, R., and Zenou, Y. (2024). Gender role models in education. *Monash University working paper*, ISSN number 1141-5429.
- Hoxby, Caroline M., “Peer Effects in the Classroom;” NBER Working Paper, 2000.
- Lucifora, C., & Tonello, M. (2015). Cheating and social interactions. Evidence from a randomized experiment in a national evaluation program. *Journal of Economic Behavior & Organization*, 115, 45-66.
- Merlino, L. P., Steinhardt, M. F., & Wren-Lewis, L. (2019). More than just friends? School peers and adult interracial relationships. *Journal of Labor Economics*, 37(3), 663-713.
- Meschi, E., & Pavese, C. (2023). Ability composition in the class and the school performance of immigrant students. *Labour Economics*, 85, 102450.
- Pagani, L., Pica, G., A peer like me? Early exposure to high achievers in math and later educational outcomes, *The European Economic Review*, Volume 181, January 2026, 105177.
- Porreca, E., Rizzica, L., Tonello, M., (2023). A Great Depression? Students’ aspirations during the Great Recession. *Working paper*.

Data availability statement

This paper uses proprietary data and therefore the author cannot provide direct access to the raw data. However, access can be obtained from INVALSI. Access is free conditional on application. Please see <https://serviziostatistico.invalsi.it/> for information on the application procedure.