FILIPPO DA RE

University of Padova

CHOOSING A HIGH SCHOOL TRACK: THE ROLE OF CLASSMATES PARENTAL OCCUPATIONS

April 2025

Marco Fanno Working Papers – 320

dSEA Dipartimento di scienze economiche e aziendali 'Marco Fanno'

Università degli Studi di Padova

Choosing a High School Track: The Role of Classmates Parental Occupations

Filippo Da Re, University of Padova

Abstract.

This paper examines how the parental occupations of grademates influence students' choice to enrol in an academic high school track. Exploiting variation in the distribution of parental occupations across classes within Italian middle school cohorts, I find that a one standard deviation increase in the share of classmates with prestigious (humble) parental occupations raises (lowers) the likelihood of academic track enrolment by 2 percentage points. Instrumental variable estimates suggest this effect is not driven by individual or peer ability, indicating a direct influence of peer networks. The negative impact of peers from disadvantaged backgrounds is particularly pronounced for low-SES students and in provinces with low social mobility. The effect is concentrated in the most prestigious academic curricula, pointing to the role of social prestige and networks. Notably, immigrant students do not enrol in the academic track regardless of their peers.

Keywords: High school track choice, peer effects, occupations.

JEL codes: I2; J24.

Acknowledgments: I am grateful to Claudia Amadei, Marco Bertoni, Giorgio Brunello, Rigissa Megalokonomou, Martina Miotto, Lorenzo Rocco, Enrico Rettore, Kjell Salvanes, Marco Tonello and the audiences at seminars in Padova, Venice, Trondheim and at FAIR NHH Bergen, LEER Leuven, 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*

School tracking refers to the practice of sorting students into different curricula, programs, or majors (i.e. tracks) on the basis of ability during secondary or post-secondary education.

A large body of evidence shows that school tracking shapes important student outcomes, such as enrolment in further education, employment, occupation, and wages (see Terrin and Triventi, 2023, for a recent review). For instance, in Italy – the country of interest of this study – graduates from academic tracks are more likely to enrol in university, earn more and work less hours than those from technical or vocational tracks (see Agarwal et al., 2021).

According to an established literature, other key factors influencing high school track choice on top of ability are family income (Andrews et al., 2020), parental background (Biewen and Tapalaga, 2020) and teacher recommendations (Carlana, 2019). Relatedly, Carlana et al., (2022) show that, in Italy, immigrant students disproportionately enrol in vocational tracks, and that teachers are more likely to recommend vocational education for immigrant students than for natives with comparable abilities and socio-economic background.

Rather surprisingly, little evidence is available on the role that peers and their background play in track choice. Interactions with class peers can affect several individual outcomes such as test scores, college choice, employment and career decisions in the labour market (Sacerdote, 2014). Another way of measuring peers' quality is by looking at peers' parental occupations. These provide information about the quality of the family background (see Fruehwirth and Gagete-Miranda, 2019). Additionally, they deliver information about the labour market, bring networks and signal social prestige. These features point to occupations having a direct impact on educational choices, rather than simply measuring socio-economic status.

Using administrative data from Italy (INVALSI), I evaluate the effect of classmates' parental occupation during middle school on pupils' high school track choice. Leveraging comparisons across pupils with the same parental occupations assigned to different classes within the same school cohort, I find that, on average, pupils with a higher share of classmates whose parents are employed in prestigious occupations experience higher chances of enrolling in the academic high school track. This effect is also relevant in magnitude, as it improves the probability of enrolling by 2 percentage points. Conversely, I

find that exposure to classmates with parents employed in humble occupations has the opposite effect, and of a similar magnitude. The results are robust across different specifications and definitions of occupation groups.

Peer parental occupation may influence pupils' choice directly – by serving as role models, providing access to information about the returns to the prestigious occupations one can access via academic tracks, and helping to establish connections with those in prestigious jobs – or indirectly. If parents employed in prestigious occupations are more academically able, then their children will themselves be abler. Exposure to peers with prestigious parental occupation would then generate an indirect ability peer effect, improving a pupil's own academic achievement and, in turn, the probability to choose an academic track.

To separate direct and indirect effects, I use variation in peers' ability and own ability generated by peers' previous classmates (partially overlapping peers, see De Giorgi et al., 2010; Nicoletti et al., 2023). Unlike previous studies, I dismiss indirect effects due to both peers' ability and students' own ability, emphasizing instead the direct effect of parental occupations.

Peer effects have important implications for school assignment policies. If peer effects were linear in means, assignment to different peers would be a zero-sum game. Heterogeneous peer effects generate instead relevant policy implications, as they help education decision makers to identifying *swing students*, who would enrol in different tracks if they were exposed to different peers. I find that parental occupation peer effects are salient in the most prestigious curricula of academic tracks, in provinces with low social mobility, and for students from disadvantaged social background. Furthermore, track choices of immigrant students are wholly unaffected by exposure to peers with different parental occupations. These findings have relevant implications for assignment policies and for the targeting of remedial interventions such as mentoring or summer programs.

This paper contributes to three strands of literature. First, it contributes to the literature on peer effects in education tracking choice. Bertoni et al., 2020, using Danish administrative data, found that privileged[1] students who are exposed to better peers are more likely to choose STEM college majors. Their identification strategy relied on comparing siblings

[1] Defined by parents' average years of education.

attending the same school at different points in time. Cattan et al., 2024 documented that students with elite educated parents are beneficial to elite students. This implies that the presence of peers with elite parental education[2] plays a crucial role in explaining the education gap between elite and non-elite students. Moreover, elite peers have a negative effect on non-elite students' GPA, lending support to the idea that the mechanism at work in explaining college enrolment is through school social networks. To the best of my knowledge, my study is the first to assess the role of peers in track choice prior to college education. I show that the effects of peers' parents operate even at younger ages than those analysed by Bertoni et al., 2020 and Cattan et al., 2024.

Second, this paper relates to the recent literature on peers' parental occupations. Porreca et al., 2023, using INVALSI data found that students exposed to school peers with at least one unemployed[3] parent report to be less likely to complete secondary education and to enrol in tertiary education. Mertz et al., 2023, using Danish registry data, found that female student exposed to school peers who have entrepreneur parents are more likely to become entrepreneurs. Montonen and Solomon, 2024, using Finnish data found that one standard deviation increase in exposure to same gender schoolmates' children from a white-collar parental background at age 15 has a significant effect on the likelihood of being in a white-collar occupation at age 30. Finally, they also provided evidence that the relevant network is that of school peers rather than that of neighbours. This paper is among the few papers, that look at parental occupations of peers – a yet overlooked matter in the peer effects literature. Differently from the standard approach taken in this literature, that looks at grademates, I am able to refine the definition of the relevant peer group to classmates, with whom students are likely to interact the most. My findings complement those of Mertz et al., 2023 and Montonen and Solomon, 2024 by proposing high school choice as a potential mechanism for shaping future occupational choices. Additionally, I work with different classifications of occupations, which allow me to investigate the role of three groups of peers i.e. low-, middle-, and high-prestige parental occupations, and how they impact on individual outcomes.

[2] Elite education is defined as combinations of field of study in certain university.
[3] They interpret unemployment as a measure of business cycle.

Lastly, I contribute to the literature focusing on the effects of peers' parental background on individual attainment. Bifulco et al., 2011, showed that an increase in the share of peers with college-graduated mothers increase the probability of enrolling in post-secondary education but has no effect on current GPA. Black et al., 2013, relied on idiosyncratic variation within schools across cohorts and studied the role of mothers' education and fathers' earnings, on future labour market outcomes. They found that boys benefit from being exposed to peers coming from families in the top quintile of earnings distribution. Fruehwirth and Gagete-Miranda, 2019 reported that pupils gain positive spillover from kids with better educated parents in U.S. kindergarteners in reading and math tests. Conversely, gains are null for socio-emotional skills. With respect these papers, thanks to instrumental variable strategy, I disentangle the direct effect of peers' parental occupations on the outcome of interest, from the indirect effect it might have via peer ability.

The remainder of the paper is structured as follows. In section 2 I describe the Italian educational system. The data are described in section 3. The empirical approach is presented in section 4. Section 5 presents the baseline results and some sensitivity checks. The mechanism is investigated in section 6. Section 7 concludes.

2. Institutional framework

In Italy, education begins at the age you turn six and it is compulsory up to 16. The education system is divided into three cycles. The first one comprises primary education and lower secondary education, for a total of eight years, and offers the same curriculum for all students. In primary schools, students spend five years (grades 1-5, ISCED level 1), whereas the remaining three are spent in middle schools (grades 6-8, ISCED level 2). At the end of grade 8 students perform an exit exam. This exam grade is not binding for the choice of high school tracks. In fact, the Italian system does not formally allow for ability tracking. Students also receive a recommendation (*consiglio orientativo*), through which teachers recommend the type of high school track they consider most appropriate for the student. Still, teachers' recommendation is not binding for students' final enrolment choice. Conditioning on students' choice, proximity of residence should be the unique requirement in high school enrolment. The second cycle refers to upper secondary education, where students attend high schools (grades 9-13, ISCED level 3). This cycle lasts up to five years and students can

choose among different fields and several programs within each field. The academic field track (*liceo*) is supposed to prepare students for college education. The vocational track (*istituti professionali*) is intended to prepare students to directly access the labour market. The technical track (*istituti tecnici*) is a mix of the former programs. It prepares students with in-depth knowledge of specific subjects, enabling them to enter the labour market or to continue to be educated. 53% of students are enrolled in the academic track, while 37% pursue a technical track, and the remaining 10% are in vocational programs. Finally, the tertiary (college) education lasts five years, consisting in three years of bachelor's degree and two years of master's degree.

The consequences of school tracking are lasting. College graduates earn a wage premium about 3.5[4] log points (SHIW data, see table A1). Those who studied at *liceo* (academic high school track) are 50 percent more likely to pursue higher education at college. 90 percent of students enrolled in the last year of *liceo* report that they wish to attend college (Almadiploma, 2024[5]). This percentage dramatically drops when we move to the vocational and technical high school tracks (figure 1). High school tracks differ in the socioeconomic background of their students. In figure 2, parental occupations are divided into three categories based on their social prestige. Sons and daughters of parents in the least prestigious occupations are less likely to be enrolled in *liceo*. Given that Italy is characterised by a largely public education system, but where school leaving grades are not binding for further enrolment in either high school or university, the persistence of tracks and their impact on the labour market, as well as the differences in the self-selection of students into tracks by socio-economic and native background, are quite remarkable.

3. Data

I use administrative data on population of Italian students between year 2012 and year 2019, called INVALSI[6]. These data are collected in grade 2, 5 (primary school), 8 (middle school) and 10 (high school). INVALSI administers the same tests in reading and maths for all Italian students, conditioning on grade, to keep track of schools' performance. Both demographic

[4] Similar estimates can be found in Anelli, 2020.
[5] https://www.almadiploma.it/indagini/profilo/profilo.aspx. Consulted on 30 July 2024.
[6] INVALSI stands for "*Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione*" that is: National Institute for the Evaluation of the Education and Training System.

characteristics and students' questionnaires are filled in by the students themselves during test days.

3.1. Sample selection

The longitudinal nature of the INVALSI data allows to link middle class peers in grade 8 to the high school track where pupils are enrolled in grade 10. Most of the analysis concentrates on grade 8, for which I have data on students' demographics and parents' occupation. From grade 10 I retrieve high school choices. To investigate mechanisms, I also link students with their grade 5 class peers. Table A2 details cohorts' grades and year of observation.

I start with 661,217 students attending schools in the Centre-North of Italy. Southern regions are excluded for various reasons, including teacher strikes (Meschi and Pavese, 2023), low reliability of test scores (Lucifora and Tonello, 2015), and longitudinal linkage issues (Bertoni and Parkham, 2024). I exclude from the sample 5742 students with missing information on gender, immigrant background and test scores. I also exclude 18,702 pupils in classes with more than 30 pupils and fewer than 10 pupils - 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), - and 30,363 pupils in single-class schools due to the lack of variation across classes. 712 pupils repeated grade 8, for them I their last appearance in grade 8. Finally, the empirical strategy omits 17,764 observations belonging to singleton cells. I omit 87,898 observations with missing information on both parental occupations because no parental information could be used for them. 98,285 observations that cannot be traced to grade 10. The latter two exclusions are discussed in detail in section 5.3.2. The final sample includes data on 401,627 students in 26,944 classes in 3737 schools.

3.2. Descriptive statistics

3.2.1. Occupations

Students are asked to report both parents' occupations and they can choose among nine options. These are: unemployed, homemaker, manager, entrepreneur, professional, self-employed, teacher, workman, retiree (see De Benedetto and De Paola, 2023). These categories try to reflect the ISCO classification at one digit. I group occupations by social

prestige (see figure 2); hence I defined the following three occupations group, which from hereafter I refer to as: low (unemployed, homemaker, retiree[7]), middle (self-employed, teacher, workman), and high (manager, entrepreneur, professional). Evidence of the rationale behind this classification can be found by looking at earnings profile of this occupations. Figure A2 plots average earnings for different occupations. The figure reveals a clear increasing pattern as we move towards most prestigious occupations, with the low group performing below 9 log points, the medium between 9 and 10, and the high one above 10.

3.2.2. Occupation class share

I define the variable class share in the following way:

$$Share_{-icst} = \frac{\sum_{cst \neq i} group_{cst}}{N_{sct} - 1} \quad (1)$$

Where *c*, *s*, *t*, stand for class, school, and cohort respectively and N identifies the total number of students within the class. The variable *group* refers to the occupations group I defined in the previous subsection. Crucially the peers' shares related to observation *i* are computed leaving them out. I compute four different types of shares in the class:

i) Share of *rank*: for each student I consider solely the parent belonging to the more prestigious occupations group[8], over the number of students in the class. Similar to Montonen and Solomon, 2024.

ii) Share of *mothers*: mothers in each occupations group over the number of mothers in the class.

iii) Share of *fathers*: fathers in each occupations group over the number of fathers in the class.

[7] Retirees may have belonged to prestigious occupations. Since I have no information on previous occupations, I re-estimate equation (2) by excluding retirees. The results are unchanged. This is not surprising as, given the relatively young age of students, retired parents are just 2 percent of the sample.

[8] As an example, if the mother of students *i* is in occupations group high and the father is in occupations group middle, the mother is the parent with the highest rank occupations. Hence, I considered the mother, and the pupil has "high" parental occupation status.

iv) Share of *parents*: parents in each occupations group over the number of parents in the class. In this computation, a pupil can contribute with one or two parents, depending on the available information on occupations.

The main analysis is conducted using share i) to investigate social prestige in a finer way. Shares ii) and iii) are computed to account for gender differences in labour market. Share iv) gives a general overview of the class parental background. I use the additional shares ii), iii), iv) as robustness tests in appendix B1. However, since for 80 percent of the sample the parent with the highest rank occupation is the father, there is a strong correlation between share i) and iii). Figure 3 shows the density of the class shares computed. The figure highlights interesting variability in the occupations group shares across classes. Furthermore, it signals some variability across parental roles. In this sense, the differences in the low occupations group between fathers and mothers are striking (see also table 1).

3.2.3. Demographics

INVALSI provides further demographic information, such as: year and month of birth, gender, immigrant status. Summary statistics are reported in table 1. Females and males are equally represented in the sample. Most of the students are fourteen years old (the age at which you are expected to conclude the grade 8). 3 percent of students are first generation immigrants, whereas 6 percent are second generation ones. Repeating grade 8 is quite uncommon in Italy as it can be seen by the data. Finally, some students in Italy may decide to enter school one year before (i.e. at age they turn five); these are signalled by the variable *early starters*.

3.2.4. Outcomes of parental background

Parental background of peers signals social prestige, social networks and labour market information. Similarly, parental background determines students' ability, as the latter is correlated with parental ability, which is reflected by parental occupations. Better ability peers may affect individual ability and thus make the pupil more likely to enrol in the academic track choice. To investigate such mechanism, I use the average ability of class peers in math and reading, measured through standardized test scores[9]. To ensure

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

comparability across subjects and cohorts I take the average of both test scores, and I keep the percentile position of the students with respect to the national distribution of test scores.

4. The empirical approach

I estimate the effect of peers' parental occupations group in student *i*'s cohort *t*, in class *c*, in school *s*:

$$Y_{icst} = \alpha + \beta_1 ranklow_{-icst} + \beta_2 rankhigh_{-icst} + \lambda_{sto_m o_f} + \varepsilon_{icst} \quad (2)$$

Where Y_{icst} 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 occupations group respectively. λ are school-cohort-mother's occupation-father's occupation fixed effects. So, equation (2) compares students from the same school, same cohort and with the same combination of father and mother occupations group. For example, I compare student *i,* from school *s*, cohort *t,* with father in the *low* occupations group and mother in the *high* occupations group, with students *j,* from school *s*, year *t,* with father in the *low* occupations group and mother in the *high* occupations group. This specification allows to use the variation stemming from different classes within the same school-cohort. Thus, λ controls for many important features of this design. First, it accounts for school and year specific trends. Second, it controls for local business cycle, thanks to the inclusion of time and geography indicators given by cohort and school variables. Finally, it controls for job characteristics, and particularly for the combination of father and mother occupations to absorb potential family spillover. Standard errors are clustered at the school level. Given that I compare effects across multiple occupations group I conservatively adjust inference for multiple testing and report False Discovery Rate-adjusted q-values along with standard clustered standard errors (Anderson, 2008).

5. Results

5.1 Baseline results

Table 2 reports baseline results. For the sake of comparison, I standardize the class shares (see table 1) in equation (2). The sign of the coefficients is as expected. An increase in the share of peers' parents in the low (high) occupations group decreases (increases) the probability of enrolling in *liceo*. In terms of magnitude, a one standard deviation increase in

the occupations class share is associated with a reduction in the probability of enrolling in *liceo* of about 0.5 percentage point. The magnitude is larger for the high occupations group share, namely 2 percentage points. After removing school-cohort-mother's occupation-father's occupation fixed effects the standard deviation of residuals is 0.038 for the share of low occupations group, and 0.091 for the share of high occupations group. Which are respectively 71 and 53 percentage points of residual variation left. The results are robust using the other occupations class shares, as can be seen in table B1. The effects estimated in the baseline results suggest that the quality of class peers matters in determining students' high school track choice.

5.2. Validation of the identification strategy

The empirical strategy assumes so far that the variation in class composition within school cohorts is as good as random once I control for school-cohort-mother's occupation-father's occupation fixed effects (λ). Exploiting within class variation implies that once conditioned by this fixed effect, the variation in parental occupations group is solely attributable to classes, which I assume to be exogenous. To test this assumption, I regress the available covariates as dependent variable in equation (2). Table 3 reports the estimates of such exercise. Exposure to classmates' parental occupations is largely unrelated to predetermined individual characteristics. These results are encouraging for considering the treatment variable as not correlated with observed individual characteristics. Appendix C Suggests additional tests to validate the underlining identification strategy. The regression of immigrant status in table 3 reports statistically significant coefficients. This depicts a potential correlation between immigrant status and peers' parental occupations. The coefficients are, however, seemingly negligible in terms of magnitude. To put it into perspective: the second row implies that following a one standard deviation increase in the share of the high occupations group, observations are 0.5 percent less likely to be immigrants. Nonetheless, in section 5.3.3. I present some tests to control for this potential confounder.

5.3. Robustness tests

Before proceeding with the analysis of the role of peers' ability and own ability, I present some sensitivities to my main results.

5.3.1. Adding controls

The first set of tests examines whether results are sensitive to the addition of controls. I estimate equation (2) augmented with a vector of student *i*'s characteristics (immigrant status, gender, age, repeating grade 8, mother's education and father's education); the results are shown in table 4 column (1). Compared to the baseline results, the two main coefficients remain statistically significant, and the magnitude is roughly unchanged. Notice that although immigrant status is unbalanced, controlling for it does not alter the findings. Section C2 provides tests with a similar approach and confirms this intuition.

5.3.2. Attrition and measurement error

Attrition in INVALSI data is a non-negligible problem. In fact, 18 percent of the students cannot be linked between grade 8 and grade 10. I conduct a placebo test by setting the outcome variable equal to 0 for observations that could not be followed. This is equivalent to assume that if I cannot follow students, they are not enrolled in *liceo*. The results are shown in table 4 column (2) and are comparable to the baseline ones. Non-random attrition seems to play no role in this setting. Following Sojourner, 2013 and Gagete-Miranda et al., 2024 I apply a correction method to address the problem of missing peers' data. Since missing data on parental occupations are considerable in INVALSI, this may introduce random measurement errors that would bias the results toward zero. Consider the following equation:

$$E[Y] = E[Y|X_i]P(X_i) + E[Y|X_{-i}]P(X_{-i}) \quad (3)$$

In equation (3), $E[Y|X_i]$ represents the outcome of interest conditioned on parental occupations (X_i). However, this quantity is computed only for a sub-sample (i.e., observations with non-missing information on parental occupations). The idea proposed by Sojourner, 2013 is simple: multiply $E[Y|X_i]$ by $P(X_i)$, which is the proportion of non-missing values in the class. In other words, equation (3) weighs the classmates' parental occupations shares by the proportion of peers with observable data on these occupations. Estimates from such regression are shown in table 4 column (3), which shows that results are robust to this additional correction.

5.3.3. Native shares

As outlined in section 5.2., even though smaller in magnitude, the immigrant status trait is not well balanced in the specification. Section 5.3.1. partially address this problem by including an indicator for immigrant status among other controls. Here, I conduct two additional tests. Table 5 columns (1) and (2), table D1 and table D2 (appendix D) report the heterogeneous estimates by immigrant status and by immigrant status and gender respectively. From those tables it can be noticed that immigrants do not drive the results. This can be interpreted as the fact that immigrants tend not to enrol in *liceo* regardless to peers' exposure, in line with the findings of Carlana, 2022. Since in Italy immigrants are concentrated on less skilful jobs one issue may raise if in computing occupations' share, I pick up an immigrant effect. If it was the case this would mean that exposure to low occupations group is indeed exposure to more immigrants' classmates, since their parents' occupations are likely to be the less prestigious once. I run a placebo test where I compute occupations group shares only considering natives. The distributions of the shares are unchanged as can be seen by comparing figure 3 and figure A2. Table 5 columns (3) and (4) report the estimates of such regression for natives and immigrants and only for natives. The results are identical to the baseline ones.

6. The role of peers' ability and own ability

6.1. Peers' ability

Peers with prestigious (humble) parental backgrounds may also possess high (low) academic ability. In this section, I explore whether it is the peers' abilities or their parents' occupations that influence their decision to enrol in an academic track. 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, tested whether spillovers from peer parental education remain after controlling for peer achievement at the start of kindergarten. This may not be the correct approach. Parental background could be a proxy for parental ability, that affects their kids' scores. This makes their kids' score a bad control to estimate the impact of parents' education. Consider the following model*:*

$$Y_{icst} = \alpha + \beta_1 ranklow_{-icst} + \beta_2 rankhigh_{-icst} + \eta X_{icst} + \mu TS_{-icst} + \lambda_{sto_m o_f} + \varepsilon_{icst} \quad (4)$$

In equation (4), *TS* stands for the average ability of grade 8 peers measured by standardized test scores in grade 8. X refers to the individual controls discussed in section 5.3.1. Two issues arise when estimating equation (4). First, peers' ability is correlated with parental ability, which is reflected by parental occupations. Secondly, peers' ability serves as an outcome variable that is influenced by parental occupations, as these occupations impact children's ability and achievements. *TS* is thus a bad control. To solve this problem, one would need an instrument which is orthogonal to parental occupations of grade 8 peers. To instrument for ability, I take advantage of the longitudinal dimension of the data to recover the partially overlapping peers (see De Giorgi et. al, 2010 and Nicoletti et. al, 2023). In this context it translates into 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 peers of pupil *i* are previous classmates of *i*'s peers who were not classmates of *i*. This strategy generates individual-specific groups, thus solving the endogeneity problem[10]. The first stage of equation (4) would be:

$$TS_{-icst} = \alpha + \beta_1 ranklow_{-icst} + \beta_2 rankhigh_{-icst} + \delta Z_{-icst} + \eta X_{icst} + \lambda_{sto_m o_p} + u_{icst} \quad (5)$$

Where the instrument (Z) is the average ESCS[11] (indicator for the socioeconomic background) in grade 5's class. I do not rely on the ability of partially overlapping peers to avoid simultaneity, since I only measure ability after primary school interactions and test score outcomes may be reflected among peers, whereas the instrument that I select is related to class composition, which is set prior to primary school interactions. It can be assumed as predetermined and thus exogenous. Since partially overlapping peers and *i* have never been classmates, they 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 that the exclusion

[10] 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".

[11] 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.

restriction holds. The partially overlapping peers categorization excludes the grade 8 *i*'s peers who were peers of *i* also in grade 5. Thus, the occupations group share in equations (4) and (5) are computed on new peers[12] (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). Table A3 exemplifies the different type of peers discussed in the paper. Column 1 of table 6 presents the estimates of equation (4) and shows that higher ability peers negatively affect the academic track choice. Moving to instrumental variable (IV) estimation, column (2) of the same table reports the first stage (FS) estimates – see equation (5) – and suggests that better peers of my peers (who were not my peers) decrease the average ability of new peers in grade 8. The value of the Kleibergen-Paap F statistic is reassuringly above 10. The mechanism behind this negative first stage is consistent with class ability rank, which refers to the position of a student in the class ability distribution. Hence, for a given individual ability, students have lower rank when surrounded by abler peers (Bertoni and Nisticò, 2023). In this context this means that better peers of *i*'s grade 8 classmates in grade 5 decrease their ability, and in grade 8, *i* is relatively better than their new peers. Thus his- or her ordinal ability rank in grade 8 is increased. In fact, the average new peers class rank in grade 8 is 9 percentile points lower. This explains the negative first stage sign on ESCS. These findings are in line with those of Paffenholz (2024), and Palma, (2024) who report the impacts of socioeconomic status rank on ability rank. Column (3) of table 6 presents the reduced form (RF) estimates which reconcile the sign of the first stage. The direct effect of better peers in terms of parental background on academic choice is indeed positive. The last column of table 6 reports the results of the IV estimation. Overall, looking at column (4) the coefficients on parental occupations are unchanged with respect to the baseline findings, suggesting that parental occupations and peers' ability are two complementary but distinct mechanisms in academic track choice.

6.2. Own ability

Does peers' parental occupation influence track choice directly or via their impact on students' own ability? I investigate this matter by employing own ability as an outcome variable in the baseline specification in equation (2). The estimates of such regression

[12]

produce insignificant results, as it can be seen in table 7. This excludes any impacts of peers' parental occupations through own ability.

6.3. Heterogeneity analysis

Although I estimate linear model, heterogenous effects can still characterize *swing students*. Heterogenous findings imply private gains, which are relevant for policy makers, to understand who is penalized or benefits from peers' exposure. I study whether the results are heterogeneous across different types of high school. Thus, I exploit information on the type of *liceo* that students are enrolled into. In Italy there are broadly six types of *liceo.* Even though they share common features and subjects, they differ in the field of specialization. These are: *scientifico, classico, musicale, scienze umane, artistico* and *linguistico*[13]. *Scientifico* and *classico* are considered the toughest and most prestigious ones. Two interesting results can be found in table 8, where I compare the latter two *licei,* the other *licei,* and the vocational and technical tracks. First, peers' parental occupation matters for the toughest curricula of *licei,* whereas for the others they do not. Second, the sign of peers' parental occupations exposure is reversed for the technical and vocational tracks. This implies that a larger exposure to peers' parents in low occupations group increases the probability of enrolling in vocational and technical tracks, while it decreases with exposure to high occupations group. I interpret these findings as additional evidence of the direct role of occupation through social prestige in shaping students high school choice. As a final investigation I study the role of social mobility. I take the data from Acciari et al., 2022 where they developed a rank based on absolute upward mobility for Italian provinces. In table 9 columns (1) and (2), I show heterogeneous results by median rank, where the provinces below the rank are characterized by greater social mobility. The findings reveal that low occupations group peers are detrimental in provinces characterized by lower social mobility, whereas such peers are irrelevant in provinces with higher social mobility. Considering that I am excluding Southern Italy, where the lowest provinces in terms of social mobility are concentrated, this result is quite powerful. The finding is consistent with the idea that school tracking is closely linked to socioeconomic background. This intuition is confirmed by estimates in columns (3) and (4) where I use the median ESCS. Finally, in

[13] The types and their field of specialization are: scientifico (maths), classico (Latin and Greek), musicale (music), scienze umane (social sciences), artistico (arts) and linguistico (3 foreign languages).

table 10, I divide the sample by the three occupations group. The low group does not contribute to the results, mainly for the small sample size. Both the middle and the high occupations group are affected by peers' parental occupations, by a similar amount but with less precise estimates for the high group (column 3). These results suggest that the middle group is the one that is more responsive to peer exposure, which is consistent since it acts as the counterfactual group.

7. Conclusion

In this paper I evaluate the role of classmates' parental occupation in high school track choice in Italy. I measure one of the channels through which peer effects operate, that is peers' parents. This work contributes to the existing literature in three ways. This paper is the first to assess the effect of peers in track choice prior to college education. Second, it is among the few papers to define the quality of peers based on parental occupations of grademates. This branch of the literature is still scarce if compared to those referring to other peers' characteristics as gender, ability or ethnicity. Third, I show that peers' socioeconomic background may not serve just as measure of peers' quality, but it can have a direct effect on academic outcomes. My analysis builds on the definitions of three categories of occupational groups. Then, I derived from those four different types of shares to avoid picking up some gender or parental role. I found that a one standard deviation increase in the share of high occupations group peers increases the probability of enrolling in the Italian academic school track (*liceo*) by 2 percentage points. The converse is true for the low occupations group. The results are persistent to different definitions and tests throughout the paper. The assumption behind the identification strategy is validated by several balance tests carried out following the recent literature on peer effects. I presented many robustness tests to address potential weaknesses of this work. I instrumented peers' ability through partially overlapping peers group, and I showed that there is not a direct effect of peers' parental occupations on own ability. Hence, I ruled out the impact of both peers' ability and own ability, supporting the notion of a direct effect of peer networks on outcomes. Academic track choice is mediated by peers' ability but by a smaller amount, signalling that peers' parental occupations and peers' ability are two complementary mechanisms. Social capital and social networks are likely to be key determinants in shaping students' preferences for high school track choice. In Italy, pursuing education is regarded as a means

of achieving social recognition, especially for individuals from disadvantaged backgrounds (see figure A1), for this reason parental occupations of classmates may be an important driver of students' decision. The heterogeneity analysis revealed that most affected students come from lower ESCS, provinces characterized by lower social mobility and with parents in the middle of occupations distribution by prestige. Immigrants are not affected since they typically do not enrol in academic high school track, regardless of peer effects. Finally, parental occupations of peers influence the choice of enrolling in the most prestigious curricula of *liceo*, but not to the others. Taken together these findings identify pupils who are exposed the most to peers' parental occupations.

Two policy implications can be derived from this paper. A possible ex-ante policy may be school randomization. Since parents are aware of better schools and try to enrol their pupils in them, one solution could be to randomize between schools close to each other to avoid parents and students' self-selection. Though, this type of policy would require financing students' mobility for example through vouchers (as in Figlio and Rouse, 2006). A more credible ex-post policy is to design remedial interventions. Some students will inevitably have to work alongside with low quality peers, so class reshuffle will not help, as it would just shift peers to other students. However, knowing that some students in terms of parental and socioeconomic background, or by geographical location are the most affected by low quality peers may help schools in designing remedial intervention to controvert such effects. These tailor interventions may be mentoring (Resnjanskij et al., 2024) or summer programs (Azzolini et al., 2023).

More studies are needed to provide a comprehensive understanding of a crucial topic as education pathways. This paper asks and tries to answer to important questions on issues as the importance of school track choice, the role of parental occupations other than parental income in shaping classmates' decision of schooling, and the public debate of class and school design.

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.

Agarwal, L., Brunello, G., & Rocco, L. (2021). The pathways to college. *Journal of Human Capital*, *15*(4), 554-595.

Almadiploma, 2024. https://www.almadiploma.it/indagini/profilo/profilo.aspx. Consulted on 30 July 2024.

Anderson, M. L. (2008): "Multiple inference and gender di↵erences in the e↵ects 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., Brunello, G., & Rocco, L. (2013). When the cat is near, the mice won't play: The effect of external examiners in Italian schools. *Journal of Public Economics*, 104, 65-77.

Bertoni, M., & Nisticò, R. (2023). Ordinal rank and the structure of ability peer effects. *Journal of Public Economics*, *217*, 104797.

Bertoni, M., & Parkam, S. (2025). When do ordinal ability rank effects emerge? Evidence from the timing of school closures. *Economics of Education Review*, *105*, 102633.

Bietenbeck, J. (2024). Do Motivated Classmates Matter for Educational Success?. *The Economic Journal*, ueae060

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.

Bourdieu, P. (1985). The Social Space and the Genesis of Groups. *Theory and Society*, *14*(6), 723–744.

Brunello, G., and Checchi, D., (2007) "Does School Tracking affect equality of opportunity? New international evidence", *Economic Policy*.

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. "Goals and gaps: Educational careers of immigrant children." *Econometrica* 90.1 (2022): 1-29.

Carlana, Michela, Eliana La Ferrara, and Paolo Pinotti. 2022. "Implicit Stereotypes in Teachers' Track Recommendations." *AEA Papers and Proceedings*, 112: 409–14**.**

Carruthers, C. K., & Jepsen, C. (2021). An international perspective1. *The Routledge handbook of the economics of education*, 343.

Cattan, S., Tominey, T., & Salvanes, K.G., (2024). First generation elite: the role of school social networks. *American Economic Review, Forthcoming*.

Checchi, D., Fiorio, C., Leonardi, M., 2013. Intergenerational persistence of educational attainment in Italy. *Economic Letters*, 229–232.

Chise, D., Fort, M., Monfardini, C., 2024. Scientifico! like Dad: On the Intergenerational Transmission of STEM Education in Italy. *Working paper*.

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.

Figlio, D. and Rouse, C. (2006). Do Accountability and Voucher Threats Improve Low-Performing Schools? *Journal of Public Economics*.

Fruehwirth, J. C., & Gagete-Miranda, J. (2019). Your peers' parents: Spillovers from parental education. *Economics of Education review*, *73*, 101910.

Gagete-Miranda, J., Meschi, E., & Pagani, L. (2024). Non-cognitive peer effects in early education: The influence of children born to teenage mothers on peers' behavioral problems. *Economics Letters*, *236*, 111580.

Gibbons, S., & Telhaj, S. (2016). Peer effects: Evidence from secondary school transition in England. *Oxford Bulletin of Economics and Statistics*, *78*(4), 548-575.

Goulas, S., Megalokonomou, R., Zhang, Y., (2024). Female Classmates, Disruption, and STEM Outcomes in Disadvantaged Schools: Evidence from a Randomized Natural Experiment. *Under review.*

Hanushek, E. A., & Woessmann, L. (2011). The economics of international differences in educational achievement. *Handbook of the Economics of Education*, *3*, 89-200.

Hanushek, E. A., Schwerdt, G., Woessmann, L., & Zhang, L. (2017). General education, vocational education, and labor-market outcomes over the lifecycle. *Journal of human resources*, *52*(1), 48-87.

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.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, *60*(3).

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.

Mertz, M., Ronchi, M., Salvestrini, V., Female Representation and Talent Allocation in Entrepreneurship: The Role of Early Exposure to Entrepreneurs. *Working paper*.

Meschi, E., & Pavese, C. (2023). Ability composition in the class and the school performance of immigrant students. *Labour Economics*, 85, 102450.

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.

OECD, 2024, Population with tertiary education. https://www.oecd.org/en/data/indicators/population-with-tertiary-education.html, consulted on 30 July 2024.

Paffenholz, M. (2024) Adolescents' Mental Health and HumanCapital: The Role of Socioeconomic Rank

Palma, M. (2024). Academic rank, socioeconomic rank, and educational outcomes: longitudinal evidence from Chile. *Education Economics*, 1–22.

Polipciuc, M., Cörvers, F., & Montizaan, R. (2023). Peers' race in adolescence and voting behavior. *Economics of Education Review*, 97, Article 102486.

Porreca, E., Rizzica, L., Tonello, M., (2023). A Great Depression? Students' aspirations during the Great Recession. *Working paper*.

Quintano, C., Castellano, R., & Longobardi, S. (2009). *A fuzzy clustering approach to improve the accuracy of italian student data: An experimental procedure to correct the impact of outliers on assessment test scores* (p. 149). Vita e pensiero.

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.

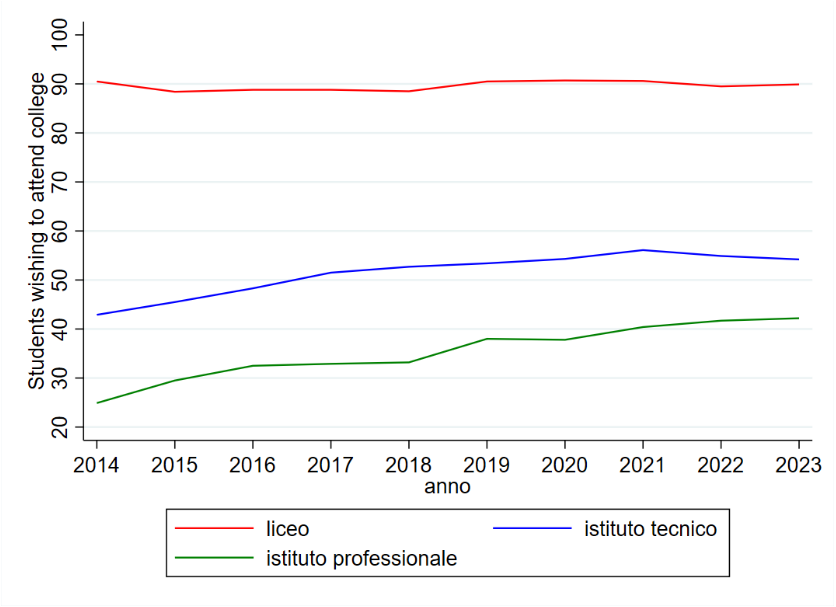
Sacerdote, B. (2014). Experimental and quasi-experimental analysis of peer effects: two steps forward?. *Annu. Rev. Econ.*, *6*(1), 253-272.

Sojourner, A. (2013). Identification of peer effects with missing peer data: Evidence from Project STAR. *The Economic Journal*, 123(569), 574-605.

Terrin, É., & Triventi, M. (2023). The Effect of School Tracking on Student Achievement and Inequality: A Meta-Analysis. *Review of Educational Research*, *93*(2), 236-274.

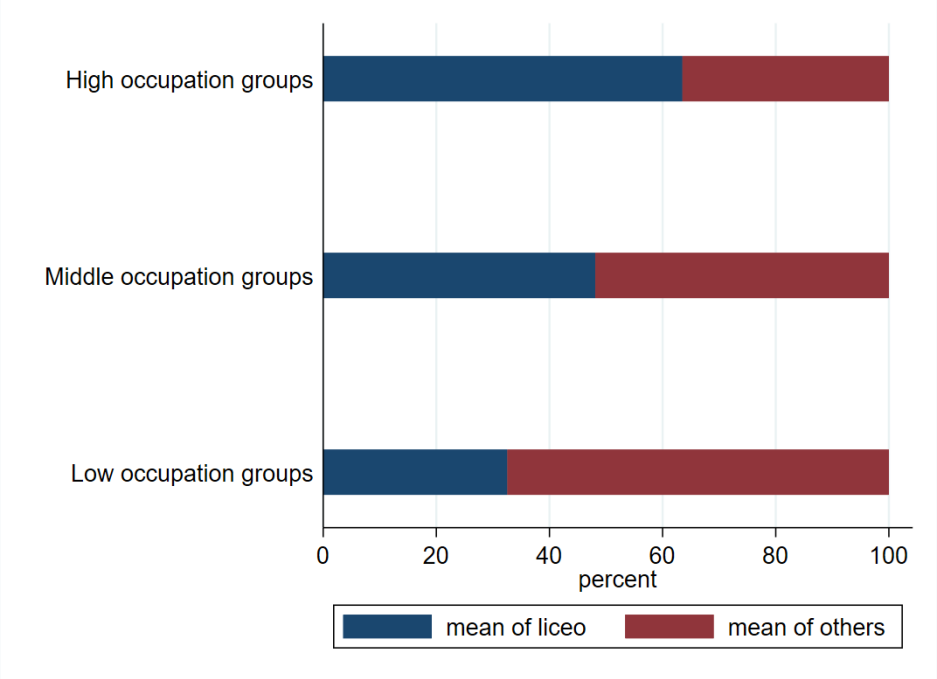
Figures

Figure 1. Percentage of last year high school students reporting they wish to attend college next year, by different high school tracks. Source Almadiploma, 2024.



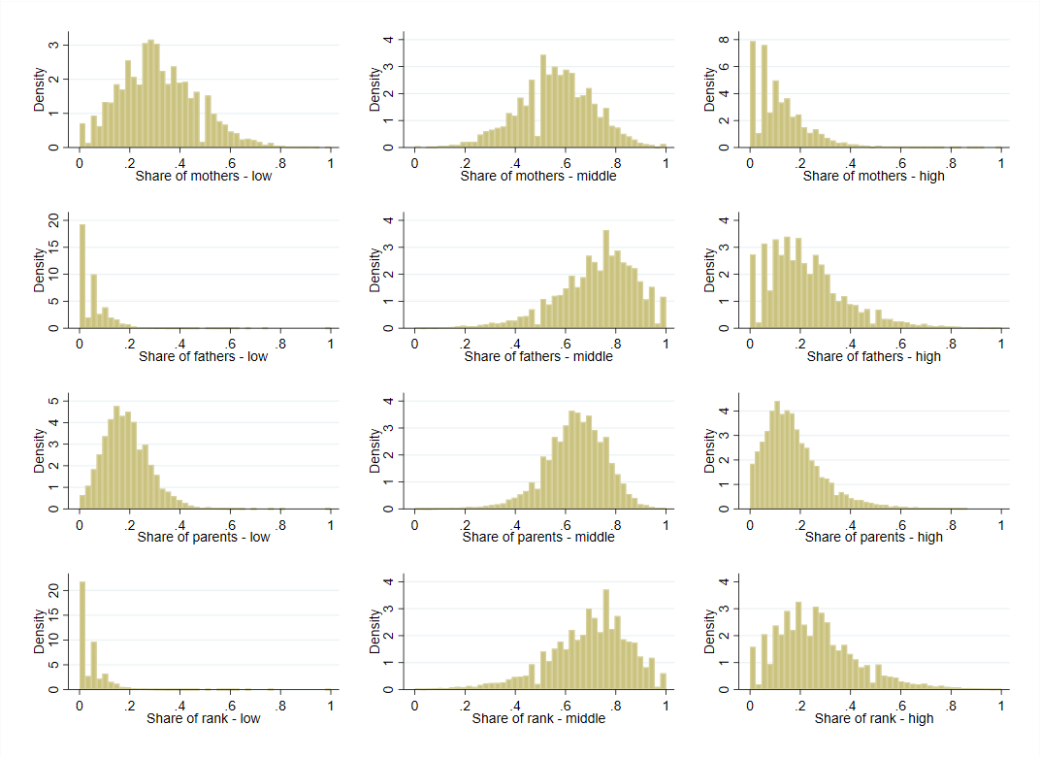
Notes: Almadiploma data are collected during high school last year. Students are asked what their future plans are. They can choose among i) intend to continue education. ii) Undecided about to intend to continue education ii) do not intend to continue education. *Liceo, instituto tecnico, istituto professionale,* refer respectively to academic, technical and vocational high school track in Italy.

Figure 2. Percentage of students going to liceo by occupations group. Source INVALSI.



Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). *Liceo* is the academic high school track in Italy. *Others* refers to technical and vocational tracks. *Liceo* refers to academic high school track in Italy. *Others* refer to technical and academic high school track in Italy.

Figure 3. Density of occupations group shares.



Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Density of the occupations group shares by classes. Number of classes: 26,284.

Tables

Table 1. Summary statistics of the sample.

	N	Mean	Standard deviation
Occupations group shares			
Share of mothers - low	401,760	0.304	0.151
Share of mothers - middle	401,760	0.564	0.148
Share of mothers - high	401,760	0.132	0.113
Share of fathers - low	401,760	0.050	0.061
Share of fathers - middle	401,760	0.709	0.162
Share of fathers - high	401,760	0.241	0.163
Share of parents - low	401,760	0.178	0.088
Share of parents - middle	401,760	0.636	0.123
Share of parents - high	401,760	0.186	0.126
Share of rank - low	401,760	0.038	0.053
Share of rank - middle	401,760	0.677	0.168
Share of rank - high	401,760	0.285	0.171
Share of mothers - low/middle	401,760	0.868	0.113
Share of fathers - low/middle	401,760	0.759	0.163
Share of parents - low/middle	401,760	0.814	0.126
Share of rank - low/middle	401,760	0.715	0.171
Share of mothers - high/middle	401,760	0.696	0.151
Share of fathers - high /middle	401,760	0.950	0.061
Share of parents - high /middle	401,760	0.822	0.088
Share of rank - high /middle	401,760	0.962	0.053
Covariates			
Female	401,760	0.509	0.500
Repeating grade 8	401,760	0.001	0.033
14 years old	401,760	0.912	0.283
Immigrant	401,760	0.097	0.295
Immigrant – first generation	401,760	0.033	0.179
Immigrant – second generation	401,760	0.064	0.244
Early starters	401,675	0.004	0.065
First quarter	401,758	0.236	0.425
Second quarter	401,758	0.252	0.434
Third quarter	401,758	0.262	0.440
Fourth quarter	401,758	0.250	0.433
Outcome of parental backgrounds			
Liceo	401,760	0.559	0.496
Average test scores (percentiles)	401,760	58.766	24.551
Class average test scores (percentiles)	401,760	54.433	9.061

Notes: share refers to the class share of occupations groups. Occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). *Liceo* is the academic high school track in Italy. *Others* refers to technical and vocational tracks. *Liceo* refers to academic high school track in Italy. *Others* refer to technical and academic high school track in Italy.

Table 2. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*).

Dependent variable	*Liceo*
Share of rank – low occupations group	-0.0044***
	(0.0012)
	[0.001]
Share of rank – high occupations group	0.0227***
	(0.0016)
	[0.001]
P-value joint significance	0
Observations	401,760
Clusters	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates of equation (2) using *rank* shares as regressors. The regression includes school-year-mother's occupation-father's occupation fixed effects. 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. Balance test.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables	*Female*	*Repeating grade 8*	*14 years old*	*Pre-school*	*Month of birth*	*Ahead year*	*Immigrant*
Share of rank –	0.0008	-0.0000	-0.0010	-0.0006	-0.0044	0.0001	0.0030***
low occupations group	(0.0011)	(0.0001)	(0.0007)	(0.0005)	(0.0076)	(0.0001)	(0.0008)
	[1.000]	[1.000]	[0.403]	[0.304]	[1.000]	[1.000]	[0.001]
Share of rank –	0.0016	-0.0001	0.0010	0.0004	0.0047	0.0001	-0.0051***
high occupations group	(0.0015)	(0.0001)	(0.0009)	(0.0007)	(0.0106)	(0.0002)	(0.0009)
	[1.000]	[1.000]	[0.403]	[0.488]	[1.000]	[1.000]	[0.001]
P-value joint significance	0.456	0.833	0.202	0.326	0.720	0.897	0
Observations	401,760	401,760	401,760	401,757	358,948	401,672	401,760
Clusters	3738	3738	3738	3738	3569	3738	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Each regression includes school-year-mother's occupation-father's occupation fixed effects. Dependent variables are displayed in the first row. 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 4. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*). Including individual controls, corrected for attrition, and corrected measurement error.

Dependent variable	(1) *Liceo*	(2) *Liceo*	(3) *Liceo*
Share of rank – low occupations group	-0.0029*** (0.0011) [0.001]	-0.0031*** (0.0011) [0.001]	-0.0028*** (0.0009) [0.001]
Share of rank – high occupations group	0.0138*** (0.0014) [0.001]	0.0140*** (0.0015) [0.001]	0.0140*** (0.0013) [0.001]
P-value joint significance	0	0	0
Observations	401,760	401,760	492,256
Clusters	3738	3738	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates of equation (2) using *rank* shares as regressors. Regressions in columns (1), (2) and (3) include school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education. In column (2) for observations that cannot be followed in grade 10 *liceo* is coded equal to 0. In column (3) each share is weighted by the share of non-missing values on parental occupations in the class. 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 5. 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.

Dependent variable	*Liceo*		*Liceo*	
	(1) *Natives*	(2) *Immigrants*	(3) *All*	(4) *Natives*
Share of rank – low occupations group	-0.0022* (0.0011) [0.009]	-0.0045 (0.0035) [0.148]	-0.0021** (0.0011) [0.005]	-0.0016 (0.0011) [0.033]
Share of rank – high occupations group	0.0140*** (0.0015) [0.001]	0.0074 (0.0054) [0.148]	0.0121*** (0.0014) [0.001]	0.0123*** (0.0014) [0.001]
P-value joint significance	0	0.109	0	0
Observations	360,539	29,981	401,748	360,530
Clusters	3736	2846	3738	3736

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates of equation (2) using *rank* shares as regressors by immigrant status. Regressions in column (1), (2), (3) and (4) include school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, mother's education, father's education. In column (3) and (4) the shares computed considering only on natives as regressors. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education 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 6. IV estimates and decomposition of the total effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*). New peers' test score in grade 8.

	(1)	(2)	(3)	(4)
	OLS	FS	RF	IV
Dependent variable	*Liceo*	*New peers' average test score grade 8*	*Liceo*	*Liceo*
Average ESCS		-0.307***	0.0169***	
(partially overlapping peers in grade 5)		(0.0724)	(0.0022)	
Share of rank – low occupations group	0.0022**	-0.481***		-0.0287***
	(0.0010)	(0.0410)		(0.0077)
Share of rank – high occupations group	0.0052***	1.017***		0.0615***
	(0.0011)	(0.0557)		(0.0161)
New peers' average test score grade 8	-0.0010***			-0.0561***
	(0.0001)			(0.0151)
Kleibergen-Paap F statistic		18.02		
P-value joint significance	0			0.001
Observations	347,073	347,073	347,073	347,073
Clusters	3697	3697	3697	3697

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates equations (4) and (5). The instruments: *Average ESCS, Share of immigrants, Share of females* are 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*. Each regression includes school-year-mother's occupation-father's occupation fixed effects and controls for gender, immigrant status, mother's education, father's education. Clustered standard errors at school level are displayed in parentheses. "P-value joint significance" indicates the p-value for the joint significance of the coefficients of "*Share of rank – low occupations group*" and "*Share of rank – high occupations group*". *** p <0.01, ** p <0.05, and * p <0.10.

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

Dependent variable	*Own average test score grade 8*
Share of rank – low occupations group	-0.0174
	(0.0494)
Share of rank – high occupations group	-0.0997
	(0.0671)
P-value joint significance	0.322
Observations	322,106
Clusters	3039

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates equations (4) and (5). *Own average test score grade 8* refers to the average percentile positions in standardized national test scores (reading and maths). Each regression includes school-year-mother's occupation-father's occupation fixed effects and controls for gender, immigrant status, repeating grade 8age, mother's education, father's education. Clustered standard errors at school level are displayed in parentheses. "P-value Hansen test" indicates the p-value for the overidentification test. "P-value joint significance" indicates the p-value for the joint significance of the coefficients of "*Share of rank – low occupations group*" and "*Share of rank – high occupations group*". *** p <0.01, ** p <0.05, and * p <0.10.

Table 8. Heterogenous effect of classmates' parental occupations on the probability of enrolling in different high school tracks.

	(1)	(2)	(3)	(4)	(5)
Dependent variables	*Liceo*	*Scientifico-classico*	*Other types of liceo*	*Technical*	*Vocational*
Share of rank –	-0.0029***	-0.0009	-0.0020**	0.0013	0.0016**
low occupations group	(0.0011)	(0.0010)	(0.0010)	(0.0010)	(0.0008)
	[0.001]	[0.012]	[0.488]	[0.107]	[0.002]
Share of rank –	0.0138***	0.0155***	-0.0016	-0.0110***	-0.0028***
high occupations group	(0.0014)	(0.0015)	(0.0015)	(0.0014)	(0.0009)
	[0.001]	[0.001]	[0.488]	[0.001]	[0.001]
P-value joint significance	0	0	0.058	0	0
Observations	401,393	401,393	401,393	401,393	401,393
Mean outcome	0.496	0.463	0.432	0.464	0.331
Clusters	3738	3738	3738	3738	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors by high school types. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education. Dependent variables are displayed in the first row. 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 9. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by provinces with high/low social mobility and by ESCS (socioeconomic background).

Dependent variable	*Liceo*		*Liceo*	
	(1)	(2)	(3)	(4)
	Below median AUM rank: higher social mobility	*Above median AUM rank: lower social mobility*	*Below median ESCS*	*Above median ESCS*
Share of rank – low occupations group	0.0001	-0.0050***	-0.0039***	-0.0021
	(0.0017)	(0.0014)	(0.0017)	(0.0016)
	[0.452]	[0.001]	[0.002]	[0.116]
Share of rank – occupations group	0.0116***	0.0158***	0.0141***	0.0105***
	(0.0021)	(0.0020)	(0.0023)	(0.0021)
	[0.001]	[0.001]	[0.001]	[0.001]
P-value joint significance	0	0	0	0
Observations	199,120	202,640	167,395	172,395
Clusters	1921	1817	3694	3679

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates equation (2) using *rank* shares as regressors by provinces with low/high social mobility. AUM rank is decreasing in AUM see (Acciari, 2022); the lower the rank the greater the social mobility. ESCS 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. The regressions in column (1), (2), (3), and (4) include school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education. 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. Heterogenous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by parental occupations.

Dependent variable	*Liceo*		
	(1)	(2)	(3)
	Low occupations group	*Middle occupations group*	*High occupations group*
Share of rank – low occupations group	-0.0008	-0.0030**	-0.0025
	(0.0062)	(0.0012)	(0.0022)
	[0.307]	[0.003]	[0.053]
Share of rank – high occupations group	0.0181*	0.0145***	0.0122***
	(0.0108)	(0.0017)	(0.0023)
	[0.209]	[0.001]	[0.001]
P-value joint significance	0.115	0	0
Observations	8,275	275,206	118,279
Clusters	1805	3735	3633

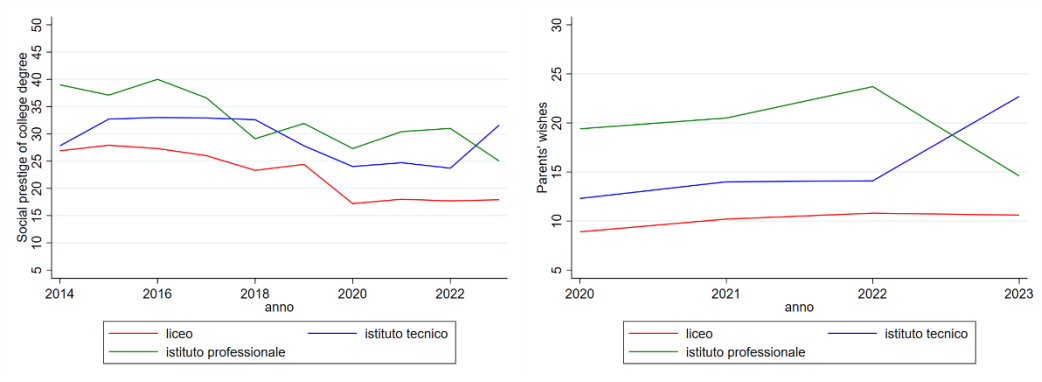
Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates equation (2) using *rank* shares as regressors by provinces with low/high social mobility. AUM rank is decreasing in AUM see (Acciari, 2022); the lower the rank the greater the social mobility. 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. The regressions in column (1), (2), (3), and (4) include school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education. 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.

Appendix

A. Supplementary descriptive statistics

Figures

Figure A1. Percentage of students who wish to attend college because of its social prestige (left) and to fulfil parents' wishes. Source Almadiploma, 2024.



Notes: Almadiploma data are collected during high school last year. Students are asked what the most important motivations for those are who intend to enrol 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. *Liceo, instituto tecnico, istituto professionale,* refer respectively to academic, technical and vocational high school track in Italy.

Figure A2. Correlations among earnings and occupations.

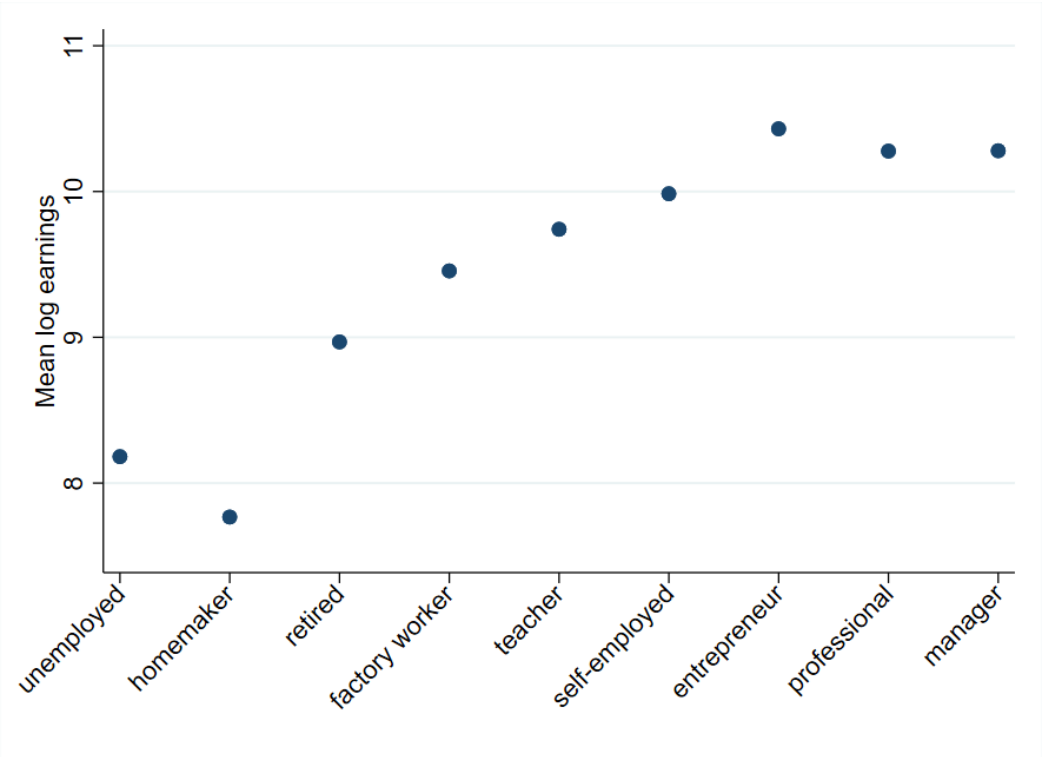


Figure A5. Share of people with tertiary education, by occupation.

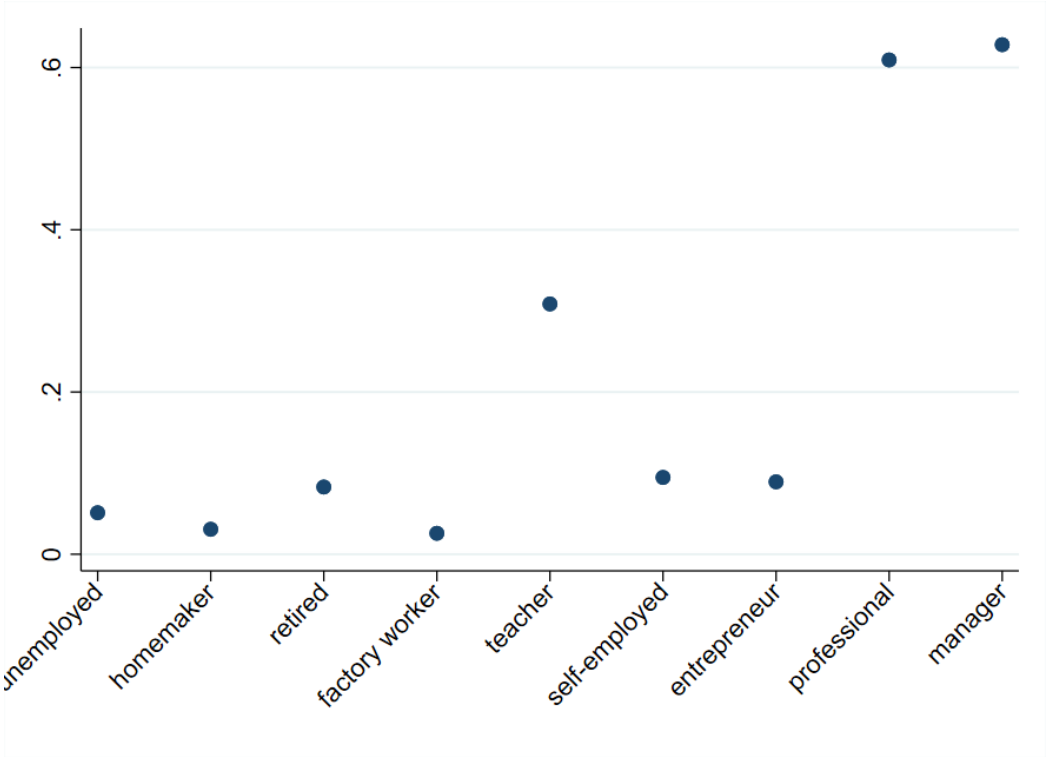
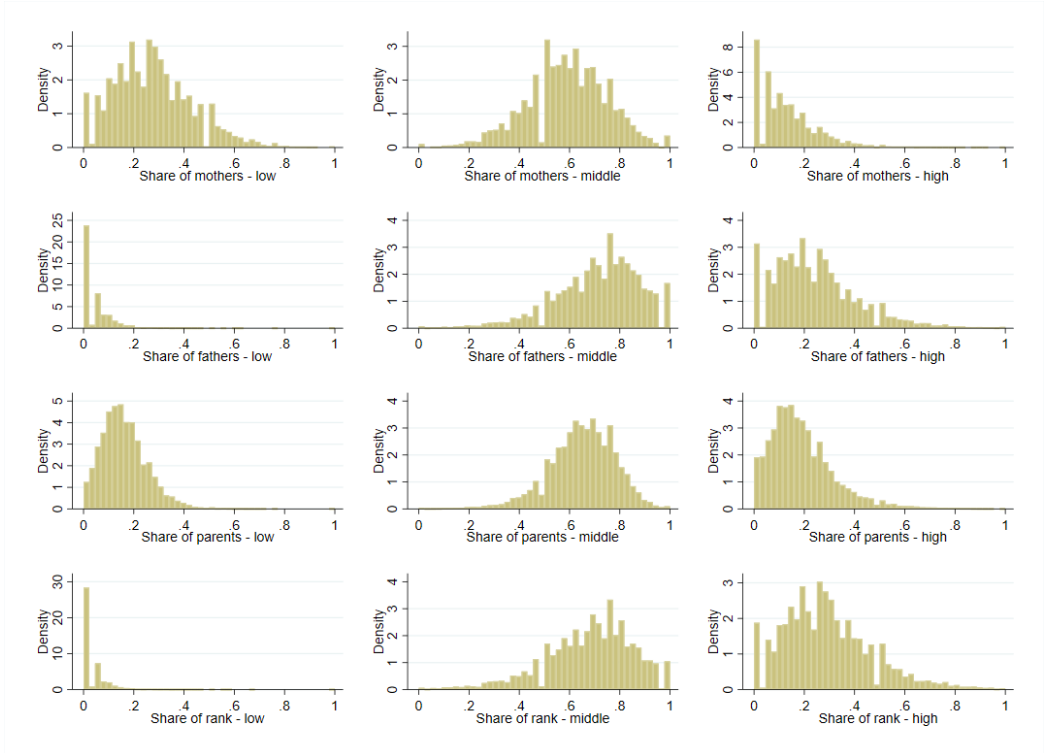


Figure A4. Density of only-native shares occupations group shares.



Notes: Occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Density of the occupations group native-specific shares by classes. Number of classes: 26,280.

Tables

Table A1. College premium. Probability of going to college. Liceo wage premium. (SHIW 2016).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Employed	Ever not employed	Log wage	Work fulltime	Bonus	College	Log wage
College	0.1257***	-0.0799***	0.3716***	0.0714***	0.1086***		
	(0.0159)	(0.0128)	(0.03379)	(0.0193)	(0.0251)		
Liceo						0.5173***	0.2002***
						(0.0321)	(0.0398)
Observations	10,089	10,089	5,137	4,383	4,383	4,391	2,997

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.

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

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

Table A3. 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		Class H: a b c; Class J: d e f	Class H: a b c; Class J: d e f
Grade 8	Class K: a b c	Class K: a b d	Class K: a b d
	a, b and c are *peers*	a and d are *new peers*	e and f are *partially overllaping peers* of a

Other occupations group shares

B1. Alternative shares definitions

In this section I present the complementary estimates for each table using the other three occupations group definitions: namely, the share of *mothers*, *fathers* and *parents* in each occupations group (see section 3.2.2.).

Tables

Table B1. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*). Other shares.

	(1)		(2)		(3)
Dependent variable	*Liceo*		*Liceo*		*Liceo*
Panel A		*Panel B*		*Panel C*	
Share of mothers – low occupations group	-0.0101***	Share of fathers – low occupations group	-0.0038***	Share of parents – low occupations group	-0.0103***
	(0.0014)		(0.0011)		(0.0014)
	[0.001]		[0.001]		[0.001]
Share of mothers – high occupations group	0.0112***	Share of fathers – high occupations group	0.0205***	Share of parents – high occupations group	0.0211***
	(0.0014)		(0.0016)		(0.0017)
	[0.001]		[0.001]		[0.001]
P-value joint significance	0	P-value joint significance	0	P-value joint significance	0
Observations	401,760	Observations	401,760	Observations	401,760
Clusters	3738	Clusters	3738	Clusters	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C show estimates equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors. The regression includes school-year-mother's occupation-father's occupation fixed effects. 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. The corresponding table in the main text is table 2.

Table B2. Balance test. Other shares.

Dependent variables	(1) *Female*	(2) *Repeating grade 8*	(3) *14 years old*	(4) *Pre-school*	(5) *Month of birth*	(6) *Early starters*	(7) *Immigrant*
Panel A							
Share of mothers – low occupations group	-0.0018	0.0001	-0.0002	-0.0003	0.0055	0.0001	0.0018**
	(0.0013)	(0.0001)	(0.0008)	(0.0006)	(0.0090)	(0.0002)	(0.0009)
	[0.188]	[0.530]	[1.000]	[0.737]	[1.000]	[0.862]	[0.019]
Share of mothers – high occupations group	0.0019	0.0000	-0.0003	0.0006	0.0020	0.0003	-0.0015*
	(0.0013)	(0.0001)	(0.0008)	(0.0006)	(0.0097)	(0.0002)	(0.0008)
	[0.188]	[0.530]	[1.000]	[0.732]	[1.000]	[0.230]	[0.019]
P-value joint significance	0.059	0.392	0.530	0.381	0.868	0.215	0
Panel B							
Share of fathers – low occupations group	-0.0005	-0.0000	-0.0007	-0.0009*	-0.0131*	-0.0000	0.0035***
	(0.0010)	(0.0001)	(0.0006)	(0.0005)	(0.0075)	(0.0001)	(0.0007)
	[0.739]	[1.000]	[1.000]	[0.052]	[1.000]	[1.000]	[0.001]
Share of fathers – high occupations group	0.0019	-0.0000	0.0004	0.0001	-0.0044	0.0001	-0.0043***
	(0.0014)	(0.0001)	(0.0009)	(0.0007)	(0.0103)	(0.0002)	(0.0009)
	[0.739]	[1.000]	[1.000]	[0.732]	[1.000]	[1.000]	[0.001]
P-value joint significance	0.402	0.808	0.638	0.076	0.617	0.689	0
Panel C							
Share of parents – low occupations group	-0.0018	0.0001	-0.0006	-0.0008	-0.0027	0.0000	0.0030***
	(0.0013)	(0.0001)	(0.0008)	(0.0006)	(0.0090)	(0.0002)	(0.0009)
	[0.182]	[1.000]	[0.294]	[0.729]	[1.000]	[0.889]	[0.001]
Share of parents – high occupations group	0.0024	0.0000	0.0005	0.0004	-0.0019	0.0003	-0.0040***
	(0.0015)	(0.0001)	(0.0010)	(0.0008)	(0.0113)	(0.0003)	(0.0010)
	[0.182]	[1.000]	[0.918]	[0.796]	[1.000]	[0.530]	[0.001]
P-value joint significance	0.061	0.686	0.278	0.349	0.995	0.371	0
Observations	401,760	401,760	401,760	358,948	401,757	401,672	401,760
Clusters	3738	3738	3738	3569	3738	3738	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C show estimates of equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors. Each regression includes school-year-mother's occupation-father's occupation fixed effects. Dependent variables are displayed in the first row. 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. Dependent variables are displayed in the first row. The corresponding table is table 3.

Table B3. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*). Including individual controls. Other shares.

	(1)		(2)		(3)
Dependent variable	*Liceo*		*Liceo*		*Liceo*
Panel A		*Panel B*		*Panel C*	
Share of mothers –	-0.0059***	Share of fathers –	-0.0019*	Share of parents –	-0.0059***
low occupations group	(0.0013)	low occupations group	(0.0011)	low occupations group	(0.0013)
	[0.001]		[0.003]		[0.001]
Share of mothers –	0.0065***	Share of fathers –	0.0125***	Share of parents –	0.0127***
high occupations group	(0.0013)	high occupations group	(0.0015)	high occupations group	(0.0015)
	[0.001]		[0.001]		[0.001]
P-value joint significance	0	P-value joint significance	0	P-value joint significance	0
Observations	401,760	Observations	401,760	Observations	401,760
Clusters	3738	Clusters	3738	Clusters	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C show estimates of equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education. 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. The corresponding table in the main text is table 4, column (1).

Table B4. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*), corrected for attrition. Other shares.

	(1)		(2)		(3)
Dependent variable	*Liceo*		*Liceo*		*Liceo*
Panel A		*Panel B*		*Panel C*	
Share of mothers –	-0.0066***	Share of fathers –	-0.0023**	Share of parents –	-0.0066***
low occupations group	(0.0011)	low occupations group	(0.0009)	low occupations group	(0.0011)
	[0.001]		[0.001]		[0.001]
Share of mothers –	0.0062***	Share of fathers –	0.0129***	Share of parents –	0.0128***
high occupations group	(0.0012)	high occupations group	(0.0013)	High occupations group	(0.0014)
	[0.001]		[0.001]		[0.001]
P-value joint significance	0	P-value joint significance	0	P-value joint significance	0
Observations	492,256	Observations	492,256	Observations	492,256
Clusters	3738	Clusters	3738	Clusters	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education. For observations that cannot be followed in grade 10 *liceo* is coded equal to 0. Clustered standard errors at school level are displayed in parentheses 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. The corresponding table in the main text is table 4, column (2).

Table B5. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*), corrected for measurement error. Other shares.

	(1)		(2)		(3)
Dependent variable	*Liceo*		*Liceo*		*Liceo*
Panel A		*Panel B*		*Panel C*	
Share of mothers – low occupations group	-0.0060***	Share of fathers – low occupations group	-0.0022***	Share of parents – low occupations group	-0.0031***
	(0.0013)		(0.0011)		(0.0011)
	[0.001]		[0.001]		[0.001]
Share of mothers – high occupations group	0.0065***	Share of fathers – high occupations group	0.0120***	Share of parents – high occupations group	0.0140***
	(0.0013)		(0.0014)		(0.0015)
	[0.001]		[0.001]		[0.001]
P-value joint significance		P-value joint significance		P-value joint significance	
Observations	401,760	Observations	401,760	Observations	401,760
Clusters	3738	Clusters	3738	Clusters	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C show estimates equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education. Each share is weighted by the share of non-missing values on parental occupations in the class. 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. The corresponding table in the main text is table 4, column (3).

Table B6. 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 shares.

Dependent variable	*Liceo*		*Liceo*	
	(1)	(2)	(3)	(4)
	Natives	*Immigrants*	*All*	*Natives*
Panel A				
Share of mothers –	-0.0061***	-0.0014	-0.0054***	-0.0053***
low occupations group	(0.0014)	(0.0043)	(0.0012)	(0.0013)
	[0.001]	[1.000]	[0.001]	[0.001]
Share of mothers –	0.0068***	0.0004	0.0054***	0.0059***
high occupations group	(0.0013)	(0.0052)	(0.0012)	(0.0013)
	[0.001]	[1.000]	[0.001]	[0.001]
P-value joint significance	0	0.937	0	0
Observations	360,539	29,981	401,748	360,530
Clusters	3736	2846	3738	3736
Panel B				
Share of fathers –	-0.0017	-0.0005	-0.0006	-0.0007
low occupations group	(0.0011)	(0.0035)	(0.0010)	(0.0011)
	[0.012]	[0.430]	[0.003]	[0.021]
Share of fathers –	0.0125***	0.0077	0.0112***	0.0111***
high occupations group	(0.0015)	(0.0053)	(0.0014)	(0.0015)
	[0.001]	[0.352]	[0.001]	[0.001]
P-value joint significance	0	0.350	0	0
Observations	360,539	29,981	401,748	360,530
Clusters	3736	2846	3738	3736
Panel C				
Share of parents –	-0.0059***	-0.0008	-0.0049***	-0.0050***
low occupations group	(0.0013)	(0.0043)	(0.0012)	(0.0013)
	[0.001]	[0.637]	[0.001]	[0.001]
Share of parents –	0.0128***	0.0059	0.0112***	0.0114***
high occupations group	(0.0016)	(0.0058)	(0.0014)	(0.0015)
	[0.001]	[0.637]	[0.001]	[0.001]
P-value joint significance	0	0.538	0	0
Observations	360,539	29,981	401,748	360,530
Clusters	3736	2846	3738	3736

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C show estimates equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors by immigrant status. Regressions in column (1), (2), (3) and (4) include school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, mother's education, father's education. In column (3) and (4) the shares computed considering only on natives as regressors. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education 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 <0.1. The corresponding table in the main text is table 5.

Table B7. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in different high school tracks. Other shares.

Dependent variables	(1) *Liceo*	(2) *Scientifico-classico*	(3) *Other types of liceo*	(4) *Technical*	(5) *Vocational*
Panel A					
Share of mothers – low occupations group	-0.0059*** (0.0013) [0.001]	-0.0040*** (0.0013) [0.001]	-0.0019 (0.0012) [0.788]	0.0028*** (0.0009) [0.001]	0.0028*** (0.0009) [0.001]
Share of mothers – high occupations group	0.0065*** (0.0013) [0.001]	0.0082*** (0.0013) [0.001]	-0.0016 (0.0013) [0.788]	-0.0018** (0.0008) [0.001]	-0.0018** (0.0008) [0.001]
P-value joint significance	0	0	0.146	0	0
Observations	401,393	401,393	401,393	401,393	401,393
Mean outcome	0.496	0.463	0.432	0.464	0.331
Clusters	3738	3738	3738	3738	3738
Panel B					
Share of fathers – low occupations group	-0.0019* (0.0011) [0.003]	-0.0003 (0.0010) [0.076]	-0.0016 (0.0010) [0.236]	0.0014* (0.0007) [0.224]	0.0014* (0.0007) [0.003]
Share of fathers – high occupations group	0.0125*** (0.0015) [0.001]	0.0147*** (0.0015) [0.001]	-0.0022 (0.0014) [0.236]	-0.0018*** (0.0009) [0.001]	-0.0018** (0.0009) [0.001]
P-value joint significance	0	0	0.107	0	0
Observations	401,393	401,393	401,393	401,393	401,393
Mean outcome	0.496	0.463	0.432	0.464	0.331
Clusters	3738	3738	3738	3738	3738
Panel C					
Share of parents – low occupations group	-0.0059*** (0.0013) [0.001]	-0.0032*** (0.0012) [0.001]	-0.0027** (0.0012) [0.201]	0.0026*** (0.0013) [0.002]	0.0033*** (0.0009) [0.001]
Share of parents – high occupations group	0.0127*** (0.0015) [0.001]	0.0156*** (0.0017) [0.001]	-0.0029* (0.0016) [0.201]	-0.0106*** (0.0015) [0.001]	-0.0021** (0.0010) [0.001]
P-value joint significance	0	0	0.019	0	0
Observations	401,393	401,393	401,393	401,393	401,393
Mean outcome	0.496	0.463	0.432	0.464	0.331
Clusters	3738	3738	3738	3738	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C show estimates equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors by high school types. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education. Dependent variables are displayed In the first row. 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. The corresponding table in the main text is table 8.

Table B8. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by provinces with high/low social mobility and by ESCS (socioeconomic background). Other shares.

Dependent variable	*Liceo*		*Liceo*	
	(1)	(2)	(3)	(4)
	Below median AUM rank: higher social mobility	*Above median AUM rank: lower social mobility*	*Below median ESCS*	*Above median ESCS*
Panel A				
Share of mothers – low occupations group	-0.0038**	-0.0079***	-0.0048**	-0.0081***
	(0.0019)	(0.0018)	(0.0020)	(0.0018)
	[0.005]	[0.001]	[0.006]	[0.001]
Share of mothers – high occupations group	0.0062***	0.0066***	0.0073***	0.0041**
	(0.0019)	(0.0017)	(0.0021)	(0.0017)
	[0.001]	[0.001]	[0.001]	[0.001]
P-value joint significance	0	0	0	0
Observations	199,120	202,640	167,338	172,395
Clusters	1921	1817	3694	3679
Panel A				
Share of fathers – low occupations group	-0.0011	-0.0024*	-0.0018	-0.0020
	(0.0016)	(0.0014)	(0.0016)	(0.0016)
	[0.133]	[0.004]	[0.044]	[0.201]
Share of fathers – high occupations group	0.0099***	0.0149***	0.0133***	0.0098**
	(0.0022)	(0.0020)	(0.0023)	(0.0020)
	[0.001]	[0.001]	[0.001]	[0.001]
P-value joint significance	0	0	0	0
Observations	199,120	202,640	167,338	172,395
Clusters	1921	1817	3694	3679
Panel A				
Share of parents – low occupations group	-0.0039**	-0.0076***	-0.0049**	-0.0078***
	(0.0019)	(0.0017)	(0.0020)	(0.0019)
	[0.004]	[0.001]	[0.003]	[0.001]
Share of parents – high occupations group	0.0108***	0.0143***	0.0141***	0.009***
	(0.0022)	(0.0020)	(0.0023)	(0.0021)
	[0.001]	[0.001]	[0.001]	[0.001]
P-value joint significance	0	0	0	0
Observations	199,120	202,640	167,338	172,395
Clusters	1921	1817	3694	3679

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C show equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors by provinces with low/high social mobility. AUM rank is decreasing in in AUM (see Acciari, 2022); the lower the rank the greater the social mobility. ESCS 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. The regressions in column (1), (2), (3), and (4) include. The included controls are: gender, immigrant status, mother's education, father's education. 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. The corresponding table in the main text is table 9.

Table B9. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by parental occupations.

Dependent variable	*Liceo*		
	(1)	(2)	(3)
	Low occupations group	*Middle occupations group*	*High occupations group*
Panel A			
Share of mothers – low occupations group	-0.0054	-0.0062***	-0.0049**
	(0.0091)	(0.0015)	(0.0022)
	[0.094]	[0.001]	[0.002]
Share of mothers – high occupations group	0.0100	0.0075***	0.0046**
	(0.0102)	(0.0016)	(0.0019)
	[0.094]	[0.001]	[0.001]
P-value joint significance	0.397	0	0
Observations	8,275	275,206	118,279
Clusters	1805	3735	3633
Panel B			
Share of fathers – low occupations group	0.0035	-0.0023*	-0.0013
	(0.0064)	(0.0012)	(0.0019)
	[0.240]	[0.015]	[0.026]
Share of fathers – high occupations group	0.0198*	0.0131***	0.0107***
	(0.0107)	(0.0017)	(0.0022)
	[0.240]	[0.001]	[0.001]
P-value joint significance	0.174	0	0
Observations	8,275	275,206	118,279
Clusters	1805	3735	3633
Panel C			
Share of parents – low occupations group	-0.0023	-0.0063**	-0.0049**
	(0.0087)	(0.0015)	(0.0022)
	[0.086]	[0.001]	[0.001]
Share of parents – high occupations group	0.0205*	0.0139***	0.0100***
	(0.0113)	(0.0018)	(0.0022)
	[0.086]	[0.001]	[0.001]
P-value joint significance	0.115	0	0
Observations	8,275	275,206	118,279
Clusters	1805	3735	3633

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimates equation (2) using *rank* shares as regressors by provinces with low/high social mobility. AUM rank is decreasing in AUM see (Acciari, 2022); the lower the rank the greater the social mobility. 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. The regressions in column (1), (2), (3), and (4) include school-year-mother's occupation-father's occupation fixed effects. The included controls are: gender, immigrant status, mother's education, father's education. 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. The corresponding table in the main text is table 10.

B2. Alternative occupations group definitions

I isolate the two tails of the occupations group distribution and compared them separately with the others. This generates new occupations group. low (unemployed, homemaker, retiree), middle-high (self-employed, teacher, workman, manager, entrepreneur, professional), high (manager, entrepreneur, professional), middle-low (self-employed, teacher, workman, unemployed, homemaker, retiree). The main equation takes the form of:

$$Y_{icst} = \alpha + \beta_l lowrank_{-icst} + \lambda^l_{sto_m o_f} + \varepsilon_{icst} \quad (6)$$

$$Y_{icst} = \alpha + \beta_h highrank_{-icst} + \lambda^h_{sto_m o_f} + \varepsilon_{icst} \quad (7)$$

For the low and high group respectively. The fixed effects λ^l and λ^h are now defined considering these new definitions of occupations group. The estimates are reported in table B10 (panel A and B). Taken together they suggest that changing the counterfactual group does not alter the main findings.

Table B10. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*). Alternative occupations group definitions.

Dependent variable	*Liceo*		*Liceo*
Panel A		*Panel B*	
Share of rank - low occupations group	-0.0080*** (0.0012)	Share of mothers - low occupations group	-0.0146*** (0.0014)
Observations	401,760	Observations	401,760
Clusters	3738	Clusters	3738
Panel C		*Panel D*	
Share of fathers - low occupations group	-0.0070*** (0.0011)	Share of parents - low occupations group	-0.0165*** (0.0014)
Observations	401,760	Observations	395,896
Clusters	3738	Clusters	3239

Dependent variable	*Liceo*		*Liceo*
Panel E		*Panel F*	
Share of rank - high occupations group	0.0247*** (0.0016)	Share of mothers - high occupations group	0.0151*** (0.0014)
Observations	401,760	Observations	401,760
Clusters	3738	Clusters	3738
Panel G		*Panel H*	
Share of fathers - high occupations group	0.0220*** (0.0016)	Share of parents - high occupations group	0.0255*** (0.0017)
Observations	401,760	Observations	401,760
Clusters	3738	Clusters	3738

Notes: occupations group panel A-B-C-D: low (unemployed, homemaker, retiree), middle -high (self-employed, teacher, workman, manager, entrepreneur, professional). Occupations group panel E-F-G-H: high (manager, entrepreneur, professional), middle -low (self-employed, teacher, workman, unemployed, homemaker, retiree). Panels A-B-C-D-E-F-G-H show estimates of equations (6) and (7) using respectively: *rank, mothers, fathers,* and *parents* shares computed considering only natives as regressors. 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.

C. Additional robustness tests

C.1 Threat to identification strategy

I perform a final test for non-random clustering of class shares. Although some variability between cohorts would not be detrimental, if I find that shares for classes in one of the two cohorts are significantly higher on average, I might suspect that the variation is due to systematic trends that could confound my results (see Merlino et al., 2019). To guard against this threat, I collapse the data to the school level and calculate the change in proportion between classes in the same school but in different cohorts. I plot the distribution of this variable in Figure C1, where there are no obvious asymmetries.

Most papers on peer effects use idiosyncratic variation in peers across cohorts within the same school (Hoxby, 2000). I have not relied on this estimation strategy for two reasons. First, this strategy is better appreciated with panel data, where one can compare many cohorts from the same school and assume that the assignment of peers to classes is driven by demographic factors which is as good as random. Unfortunately, I only have two cohorts, so this strategy may not be the most appropriate for my scenario. Second, the within-school cohort strategy allows me to compare students who have been exposed to the same environment and local business cycle by exploiting the variation across classes, the network of interest. In table C1, I present the balance test using within-school and across-cohort strategy, which are fully comparable to the baseline ones in table 3.

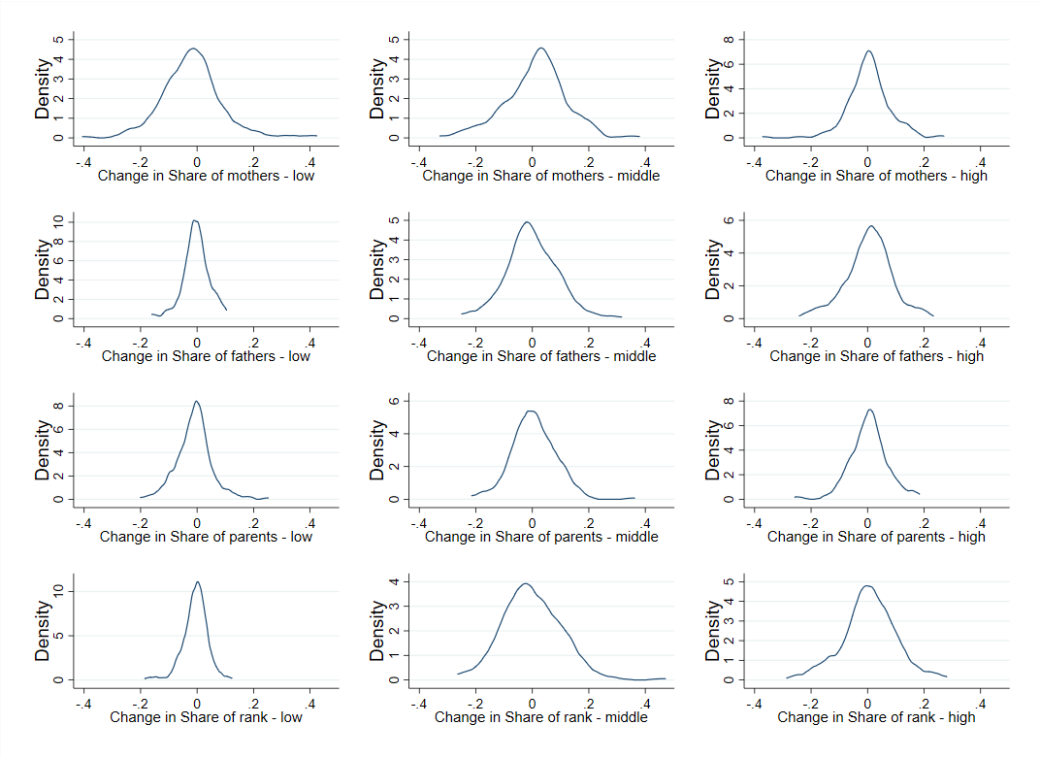
Following Bietenbeck, 2024 Table C2 shows estimates from separate regressions of students' socio-demographic characteristics on the parental occupation of their peers. Most of the coefficients on the peer variables are close to zero and statistically insignificant at conventional levels, confirming that there are no clear relationships between the predetermined characteristics and peers' parental occupations.

Table C3 reports the results of a test proposed by Porreca et al, 2023 and Goulas et al, 2024. The rationale of this exercise is to regress the peer variables on the set of observable characteristics of the students. I repeat the same exercise in Table C3.

To guard against manipulation in class formation, Goulas et al, 2024 suggest conducting a class level equilibrium test. To perform this test, I stack the data at the class level and regress the covariates on an indicator for the different classes. The reference category is the first class for each school, so the regressors in Table C4 can be interpreted as the difference

between the n-th class and the first. This provides additional evidence that there are no systematic differences within classes in the observed characteristics of the pupils.

Figure C1. Year-to-year variation in standardized shares.



Notes: Occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Kernel density of change in occupations group class shares between cohorts

Table C1. Balance test, Hoxby method.

Dependent variables	(1) *Female*	(2) *Repeating grade 8*	(3) *14 years old*	(4) *Pre-school*	(5) *Month of birth*	(6) *Early starters*	(7) *Immigrant*
Panel A							
Share of rank –	-0.0007	-0.0000	-0.0011*	-0.0006	-0.0029	0.0000	0.0036***
low occupations group	(0.0010)	(0.0001)	(0.0006)	(0.0005)	(0.0074)	(0.0001)	(0.0008)
	[0.321]	[1.000]	[0.150]	[1.000]	[1.000]	[1.000]	[0.001]
Share of rank –	0.0019	-0.0001	0.0010	0.0003	0.0008	0.0001	-0.0054***
high occupations group	(0.0014)	(0.0001)	(0.0009)	(0.0007)	(0.0100)	(0.0002)	(0.0009)
	[0.253]	[1.000]	[0.351]	[1.000]	[1.000]	[1.000]	[0.001]
P-value joint significance	0.331	0.717	0.087	0.371	0.912	0.800	0
Panel B							
Share of mothers –	-0.0008	0.0001	-0.0007	-0.0004	0.0048	-0.0000	0.0019**
low occupations group	(0.0012)	(0.0001)	(0.0007)	(0.0006)	(0.0085)	(0.0002)	(0.0009)
	[0.190]	[1.000]	[1.000]	[0.698]	[1.000]	[1.000]	[0.015]
Share of mothers –	0.0026**	0.0000	0.0001	0.0007	0.0033	0.0003	-0.0019**
high occupations group	(0.0013)	(0.0001)	(0.0008)	(0.0006)	(0.0091)	(0.0002)	(0.0008)
	[0.188]	[1.000]	[1.000]	[0.720]	[1.000]	[1.000]	[0.012]
P-value joint significance	0.058	0.478	0.630	0.266	0.832	0.322	0
Panel C							
Share of fathers –	-0.0010	-0.0001	-0.0008	-0.0009*	-0.0144**	-0.0000	0.0034***
low occupations group	(0.0010)	(0.0001)	(0.0006)	(0.0005)	(0.0071)	(0.0001)	(0.0007)
	[0.720]	[1.000]	[1.000]	[0.051]	[0.058]	[1.000]	[0.001]
Share of fathers –	0.0019	-0.0001	0.0004	0.0002	-0.0078	0.0001	-0.0047***
high occupations group	(0.0013)	(0.0001)	(0.0009)	(0.0007)	(0.0098)	(0.0002)	(0.0009)
	[0.720]	[1.000]	[1.000]	[0.711]	[1.000]	[1.000]	[0.001]
P-value joint significance	0.192	0.677	0.385	0.168	0.122	0.634	0
Panel D							
Share of parents –	-0.0013	0.0001	-0.0010	-0.0009	-0.0040	0.0000	0.0032***
low occupations group	(0.0012)	(0.0001)	(0.0007)	(0.0006)	(0.0085)	(0.0002)	(0.0009)
	[0.182]	[1.000]	[0.280]	[0.768]	[1.000]	[0.870]	[0.001]
Share of parents –	0.0029*	-0.0000	0.0003	0.0003	-0.0037	0.0004	-0.0045***
high occupations group	(0.0015)	(0.0001)	(0.0009)	(0.0007)	(0.0107)	(0.0003)	(0.0010)
	[0.105]	[1.000]	[0.855]	[0.785]	[1.000]	[0.546]	[0.001]
P-value joint significance	0.041	0.717	0.331	0.371	0.870	0.383	0
Observations	401,760	401,760	401,760	358,948	401,757	401,672	401,760
Clusters	3738	3738	3738	3569	3738	3738	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C show estimates of equation (2) using respectively: *mothers, fathers,* and *parents* shares as regressors. Each regression includes mother's occupation-father's occupation fixed effect, school fixed effect, year fixed effect and school-year linear specific trend. Dependent variables are displayed in the first row. 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. Dependent variables are displayed in the first row.

Table C2. Balance test for each single regressor (see Bietenbeck, 2024).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables	*Female*	*Repeating grade 8*	*14 years old*	*Pre-school*	*Month of birth*	*Early starters*	Immigrant
Share of rank -	0.0006	-0.0000	-0.0009	-0.0007	-0.0050	0.0001	0.0036***
low occupations group	(0.0010)	(0.0001)	(0.0007)	(0.0005)	(0.0075)	(0.0001)	(0.0008)
Share of rank -	0.0014	-0.0001	0.0013	0.0006	0.0057	0.0000	-0.0057***
high occupations group	(0.0014)	(0.0001)	(0.0009)	(0.0007)	(0.0104)	(0.0002)	(0.0009)
Share of mothers -	-0.0023*	0.0001	-0.0003	-0.0004	0.0049	-0.0000	0.0022**
low occupations group	(0.0012)	(0.0001)	(0.0007)	(0.0006)	(0.0086)	(0.0002)	(0.0009)
Share of mothers -	0.0024*	-0.0000	0.0003	0.0007	0.0004	0.0003	-0.0021***
high occupations group	(0.0013)	(0.0001)	(0.0008)	(0.0006)	(0.0093)	(0.0002)	(0.0008)
Share of fathers –	-0.0008	-0.0000	-0.0007	-0.0009*	-0.0126*	-0.0000	0.0040***
low occupations group	(0.0010)	(0.0001)	(0.0006)	(0.0005)	(0.0074)	(0.0001)	(0.0007)
Share of fathers -	0.0020	-0.0000	0.0006	0.0003	-0.0015	0.0001	-0.0051***
high occupations group	(0.0014)	(0.0001)	(0.0009)	(0.0007)	(0.0102)	(0.0002)	(0.0009)
Share of parents -	-0.0024*	0.0001	-0.0007	-0.0007	-0.0050	-0.0000	0.0039***
low occupations group	(0.0013)	(0.0001)	(0.0008)	(0.0005)	(0.0075)	(0.0002)	(0.0009)
Share of parents -	0.0030**	-0.0000	0.0007	0.0006	-0.0010	0.0002	-0.0050***
high occupations group	(0.0015)	(0.0001)	(0.0009)	(0.0007)	(0.0110)	(0.0003)	(0.0010)
Observations	401,760	401,760	401,760	358,948	401,757	401,672	401,760
Clusters	3738	3738	3738	3569	3738	3738	3738

Notes: Occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Each regression includes school-year-mother's occupation-father's occupation fixed effects. Clustered standard errors at school level are displayed in parentheses. Each coefficient comes from a separate regression of the outcome indicated in the column header on the peer variable indicated in the row. I grouped them in the same column to save space. Dependent variables are displayed in the first row. Clustered standard errors at school level are displayed in parentheses. *** p <0.01, ** p <0.05, and * p <0.10.

Table C3. Balance test. Covariates as independent variable (see Porreca, 2023).

Dependent variables	Share of rank		Share of mothers	
	Low occupations group	High occupations group	Low occupations group	High occupations group
Panel A			*Panel B*	
Female	0.0019	0.0024	-0.0035	0.0033
	(0.0025)	(0.0020)	(0.0023)	(0.0022)
Repeating grade 8	-0.0011	-0.0090	0.0483	-0.0059
	(0.0440)	(0.0327)	(0.0431)	(0.0394)
14 years old	0.0014	-0.0010	0.0032	-0.0004
	(0.0056)	(0.0042)	(0.0048)	(0.0045)
Pre-school	-0.0110	0.0024	-0.0055	0.0081
	(0.0113)	(0.0089)	(0.0101)	(0.0093)
Month of birth	-0.0002	0.0002	0.0002	0.0002
	(0.0004)	(0.0003)	(0.0004)	(0.0003)
Early starters	-0.0030	0.0064	0.0118	0.0126
	(0.0214)	(0.0183)	(0.0190)	(0.0202)
Immigrant	0.0253***	-0.0220***	0.0128**	-0.0111**
	(0.0063)	(0.0043)	(0.0053)	(0.0045)
Observations	358,867	358,867	358,867	358,867
Clusters	3569	3569	3569	3569
Dependent variables	Share of fathers		Share of parents	
	Low occupations group	High occupations group	Low occupations group	High occupations group
Panel C			*Panel D*	
Female	-0.0006	0.0032	-0.0032	0.0036*
	(0.0026)	(0.0019)	(0.0023)	(0.0019)
Repeating grade 8	-0.0179	-0.0131	0.0341	-0.0097
	(0.0442)	(0.0314)	(0.0407)	(0.0314)
14 years old	0.0002	0.0006	0.0026	0.0002
	(0.0055)	(0.0043)	(0.0048)	(0.0040)
Pre-school	-0.0162	-0.0008	-0.0107	0.0033
	(0.0116)	(0.0093)	(0.0102)	(0.0086)
Month of birth	-0.0004	-0.0000	0.0000	0.0000
	(0.0004)	(0.0003)	(0.0004)	(0.0003)
Early starters	-0.0026	0.0090	0.0113	0.0117
	(0.0212)	(0.0190)	(0.0187)	(0.0179)
Immigrant	0.0294***	-0.0195***	0.0219***	-0.0176***
	(0.0059)	(0.0043)	(0.0054)	(0.0041)
Observations	358,867	358,867	358,867	358,867
Clusters	3569	3569	3569	3569

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Each regression includes school-year-mother's occupation-father's occupation fixed effects. Clustered standard errors at school level are displayed in parentheses. Each column refers to different regressions where the regressors are: female, repeating grade 8th, 14 years old, pre-school, month of birth, ahead year, immigrant, mother's education, father's education. Clustered standard errors at school level are displayed in parentheses. *** p <0.01, ** p <0.05, and * p <0.10.

Table C4. Balance test at class level.

Dependent variable	*Female*	*Repeating grade 8*	*14 years old*	*Pre-school*	*Month of birth*	*Early starters*	*Immigrant*
2nd class	-0.0019	0.0000	0.0009	-0.0004	-0.0027	0.0000	-0.0010
	(0.0017)	(0.0001)	(0.0013)	(0.0010)	(0.0140)	(0.0002)	(0.0014)
3rd class	0.0015	0.0001	-0.0001	0.0010	0.0053	-0.0005*	-0.0008
	(0.0020)	(0.0001)	(0.0015)	(0.0013)	(0.0161)	(0.0003)	(0.0017)
4th class	-0.0011	0.0001	-0.0036**	-0.0011	-0.0416**	-0.0004	0.0025
	(0.0023)	(0.0002)	(0.0018)	(0.0015)	(0.0180)	(0.0003)	(0.0019)
5th class	-0.0021	0.0000	-0.0001	0.0001	0.0104	0.0005	0.0001
	(0.0028)	(0.0002)	(0.0021)	(0.0017)	(0.0209)	(0.0004)	(0.0024)
6th class	-0.0058*	-0.0001	-0.0033	-0.0030	-0.0394	-0.0004	0.0032
	(0.0034)	(0.0002)	(0.0026)	(0.0024)	(0.0258)	(0.0005)	(0.0030)
7th class	-0.0118***	-0.0002	0.0004	-0.0003	-0.0027	0.0007	-0.0035
	(0.0044)	(0.0003)	(0.0032)	(0.0037)	(0.0355)	(0.0007)	(0.0036)
8th class	0.0024	0.0000	0.0055	-0.0045	0.0211	0.0001	-0.0004
	(0.0061)	(0.0003)	(0.0046)	(0.0031)	(0.0432)	(0.0010)	(0.0053)
9th class	0.0021	-0.0005	-0.0024	-0.0004	0.0132	-0.0006	-0.0060
	(0.0087)	(0.0004)	(0.0063)	(0.0070)	(0.0675)	(0.0012)	(0.0068)
10th class	-0.0117	0.0010	-0.0058	-0.0061	-0.0251	0.0019	0.0077
	(0.0089)	(0.0010)	(0.0081)	(0.0083)	(0.0714)	(0.0022)	(0.0090)
11th class	-0.0022	-0.0007**	0.0030	-0.0113	-0.0716	0.0018	0.0008
	(0.0122)	(0.0003)	(0.0098)	(0.0140)	(0.1006)	(0.0018)	(0.0112)
12th class	0.0084	0.0004	0.0210	0.0239**	-0.1947	0.0003	-0.0332**
	(0.0150)	(0.0013)	(0.0142)	(0.0096)	(0.1269)	(0.0034)	(0.0130)
13th class	-0.0006	-0.0012*	-0.0306	-0.0000	-0.1386	0.0027	0.0060
	(0.0211)	(0.0007)	(0.0187)	(0.0098)	(0.1692)	(0.0055)	(0.0228)
14th class	0.0058	-0.0007	-0.0075	0.0010	-0.2655	0.0025	-0.0009
	(0.0241)	(0.0007)	(0.0207)	(0.0268)	(0.1963)	(0.0028)	(0.0210)
15th class	-0.0083	-0.0013	-0.0015	0.0028	-0.6719***	-0.0014	0.0542
	(0.0492)	(0.0011)	(0.0275)	(0.0072)	(0.1968)	(0.0010)	(0.0464)
16th class	-0.0080	-0.0002	0.0252	0.0194**	-0.7765***	-0.0025	-0.0764
	(0.0582)	(0.0003)	(0.0439)	(0.0095)	(0.2371)	(0.0016)	(0.0511)
17th class	-0.0372	-0.0002	-0.0410	0.0225*	0.2225	-0.0010	0.0023
	(0.0632)	(0.0003)	(0.0727)	(0.0134)	(0.4383)	(0.0011)	(0.0641)
Observations	26,955	26,955	26,955	26,955	26,955	26,955	26,955
Clusters	3738	3738	3738	3738	3738	3738	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Each regression includes an indicator for class identifier. Class number 1 of each school acts as excluded categories. Clustered standard errors at school level are displayed in parentheses. Dependent variables are displayed in the first row. Clustered standard errors at school level are displayed in parentheses. *** p <0.01, ** p <0.05, and * p <0.10.

C2. Including controls

In table C4, I perform a similar test of table 4, tough the characteristics are computed as average at the class level. Finally, As suggested by Goulas et al., 2024, I estimated the model with both the individual control and the class level controls (table C4). Those results are in line with the estimates in table 4 and this confirms the interpretation.

Table C4. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*). Including class controls.

Dependent variable			*Liceo*		
	Class controls	*Individual and class controls*		*Class controls*	*Individual and class controls*
Panel A			*Panel B*		
Share of rank - low occupations group	-0.0020* (0.0012) [0.043]	-0.0022** (0.0011) [0.005]	Share of mothers - low occupations group	-0.0051*** (0.0015) [0.001]	-0.0043*** (0.0013) [0.001]
Share of rank - high occupations group	0.0188*** (0.0016) [0.001]	0.0124*** (0.0015) [0.001]	Share of mothers - high occupations group	0.0102*** (0.0014) [0.001]	0.0059*** (0.0013) [0.001]
P-value joint significance	0	0	P-value joint significance	0	0
Observations	371,924	371,924	Observations	371,924	371,924
Clusters	3714	3714	Clusters	3714	3714
Panel C			*Panel D*		
Share of fathers - low occupations group	-0.0018 (0.0012) [0.007]	-0.0016 (0.0011) [0.011]	Share of parents - low occupations group	-0.0053*** (0.0015) [0.001]	-0.0046*** (0.0013) [0.001]
Share of fathers - high occupations group	0.0173*** (0.0016) [0.001]	0.0114*** (0.0015) [0.001]	Share of parents - high occupations group	0.0188*** (0.0017) [0.001]	0.0118*** (0.0015) [0.001]
P-value joint significance	0	0	P-value joint significance	0	0
Observations	371,924	371,924	Observations	371,924	371,924
Clusters	3714	3714	Clusters	3714	3714

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C-D show estimates equation (2) for the four occupations group shares. Each regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are the class average of: gender, immigrant status, mother's education, father's education. 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.

C3. Investigating gender roles

I check for potential confounders that operate at gender level. For example, Chise et al., 2024 found that in the transmission of STEM education in Italy the influence of fathers outweighs the one of mothers. In addition, their results are stronger for sons than for daughters. In table C6 I compute the gender-specific share for parents and highest rank parents following Merlino et al., 2019. In table C7 I present estimates a cross-specifications, where equation (2) has now four regressors of interest (two for mothers' occupations and two for fathers' occupations). The idea of this test is to absorb the correlation among parental occupations within family. No gender role channels seem to play important role in school tracking choice based on those estimates. This interpretation is confirmed by the fact that the results are not driven by any specific gender as can be seen in table D1.

Table C6. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*), using gender-specific shares.

Dependent variable	*Liceo*		*Liceo*
Panel A		*Panel B*	
Share of rank - low occupations group	-0.0013 (0.0009) [0.018]	Share of parents - low occupations group	-0.0046*** (0.0010) [0.001]
Share of rank - high occupations group	0.0080*** (0.0013) [0.001]	Share of parents - high occupations group	0.0071*** (0.0012) [0.001]
P-value joint significance	0	P-value joint significance	0
Observations	401,760	Observations	401,760
Clusters	3738	Clusters	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B show estimate of equation (2) using respectively *rank* and *parents* shares as regressors. Shares are gender-specific. Each regression includes school-year-mother's occupation-father's occupation fixed effects and controls for gender, immigrant status, mother's education, father's education. 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.

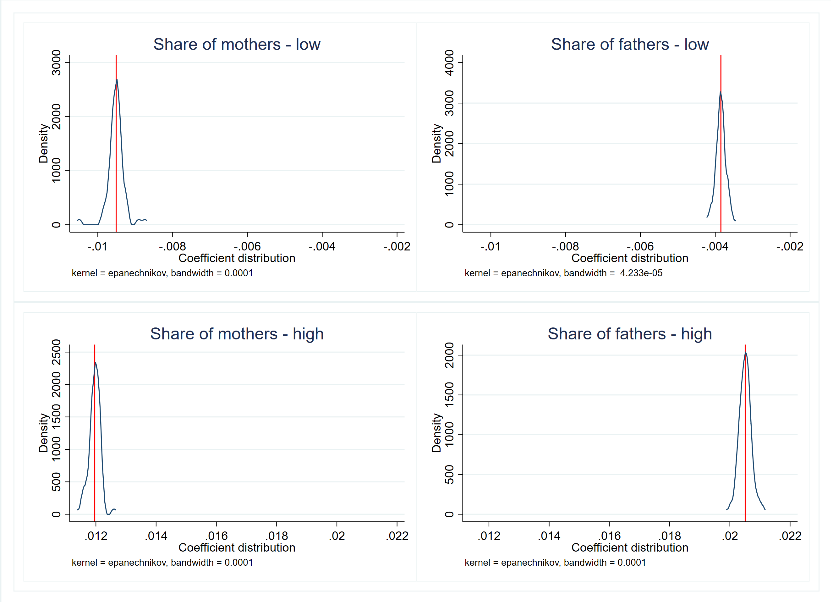
Table C7. Effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*), using cross-gender specification.

Dependent variable	*Liceo*		
	All	*Males*	*Females*
Share of mothers -	-0.0056***	-0.0056***	-0.0050***
low occupations group	(0.0013)	(0.0019)	(0.0018)
	[0.001]	[0.001]	[0.001]
Share of mothers -	0.0032**	0.0028	0.0043**
high occupations group	(0.0014)	(0.0020)	(0.0019)
	[0.001]	[0.003]	[0.001]
Share of fathers -	-0.0015	-0.0022	-0.0015
low occupations group	(0.0011)	(0.0015)	(0.0015)
	[0.006]	[0.009]	[0.0013]
Share of fathers -	0.0106***	0.0136***	0.0075***
high occupations group	(0.0016)	(0.0023)	(0.0021)
	[0.001]	[0.001]	[0.001]
P-value joint significance	0	0	0
Observations	401,760	188,246	195,530
Clusters	3738	3735	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The table shows estimate of equation (2) using both *mothers* and *fathers* shares as regressors in a single regression. Each regression includes school-year-mother's occupation-father's occupation fixed effects and controls for, gender, immigrant status, mother's education, father's education. 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.

D. Heterogeneous effects

Figure D1. Distribution of the estimated effects when one province at a time is dropped from the sample.



Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). The red vertical lines are the estimated coefficients in table 4. The blue curves are the kernel density of the estimated effects when one province at a time is dropped from the sample. Number of provinces: 69.

Table D1. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by gender.

Dependent variable			*Liceo*		
	Males	*Females*		*Males*	*Females*
Panel A			*Panel B*		
Share of rank -	-0.0023	-0.0031**	Share of mothers -	-0.0061***	-0.0053***
low occupations group	(0.0016)	(0.0015)	low occupations group	(0.0019)	(0.0018)
	[0.043]	[0.005]		[0.043]	[0.001]
Share of rank -	0.0157***	0.0119***	Share of mothers -	0.0070***	0.0067***
high occupations group	(0.0021)	(0.0020)	high occupations group	(0.0019)	(0.0018)
	[0.001]	[0.001]		[0.001]	[0.001]
P-value joint significance	0	0	P-value joint significance	0	0
Observations	188,246	195,530	Observations	188,246	195,530
Clusters	3735	3738	Clusters	3735	3738
Panel C			*Panel D*		
Share of fathers -	-0.0026*	-0.0019	Share of parents -	-0.0061***	-0.0057***
low occupations group	(0.0019)	(0.0015)	low occupations group	(0.0021)	(0.0018)
	[0.007]	[0.011]		[0.001]	[0.001]
Share of fathers -	0.0153***	0.0098***	Share of parents -	0.0152***	0.0107***
high occupations group	(0.0021)	(0.0020)	high occupations group	(0.0023)	(0.0021)
	[0.001]	[0.001]		[0.001]	[0.001]
P-value joint significance	0	0	P-value joint significance	0	0
Observations	188,246	195,530	Observations	188,246	195,530
Clusters	3735	3738	Clusters	3735	3738

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C-D show estimates equation (2) using respectively: *rank*, *mothers, fathers,* and *parents* shares as regressors by gender. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: immigrant status, mother's education, father's education. 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 D2a. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by gender and immigrant status (females: natives and non-natives).

Dependent variable			*Liceo*		
Females	*Natives*	*Immigrants*		*Natives*	*Immigrants*
Panel A			*Panel B*		
Share of rank -	-0.0028*	0.0000	Share of mothers -	-0.0056***	0.0027
low occupations group	(0.0016)	(0.0059)	low occupations group	(0.0019)	(0.0074)
	[0.008]	[1.000]		[0.008]	[1.000]
Share of rank -	0.0115***	0.0110	Share of mothers -	0.0064***	0.0046
high occupations group	(0.0021)	(0.0092)	high occupations group	(0.0018)	(0.0084)
	[0.001]	[1.000]		[0.001]	[1.000]
P-value joint significance	0	0.471	P-value joint significance	0	0.843
Observations	174,390	13,091	Observations	174,390	13,091
Clusters	3734	2083	Clusters	3734	2083
Panel C			*Panel D*		
Share of fathers -	-0.0011	-0.0003	Share of parents -	-0.0055***	0.0022
low occupations group	(0.0016)	(0.0059)	low occupations group	(0.0019)	(0.0075)
	[0.005]	[1.000]		[0.001]	[1.000]
Share of fathers -	0.0097***	0.0072	Share of parents -	0.0106***	0.0079
high occupations group	(0.0021)	(0.0090)	high occupations group	(0.0022)	(0.0098)
	[0.001]	[1.000]		[0.001]	[1.000]
P-value joint significance	0	0.709	P-value joint significance	0	0.721
Observations	174,390	13,091	Observations	174,390	13,091
Clusters	3734	2083	Clusters	3734	2083

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C-D show estimates equation (2) using respectively: *rank*, *mothers, fathers,* and *parents* shares as regressors by immigrant status and gender. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: mother's education, father's education. 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 D2b. Heterogeneous effect of classmates' parental occupations on the probability of enrolling in the academic high school track (*liceo*) by gender and immigrant status (males: natives and non-natives).

Dependent variable				*Liceo*	
Males	*Natives*	*Immigrants*		*Natives*	*Immigrants*
Panel A			*Panel B*		
Share of rank -	-0.0010	-0.0165***	Share of mothers -	-0.0062***	-0.0068
low occupations group	(0.0017)	(0.0053)	low occupations group	(0.0020)	(0.0069)
	[0.311]	[0.01]		[0.001]	[1.000]
Share of rank -	0.0161***	0.0041	Share of mothers -	0.0072***	-0.0027
high occupations group	(0.0022)	(0.0088)	high occupations group	(0.0020)	(0.0082)
	[0.001]	[0.175]		[0.001]	[1.000]
P-value joint significance	0	0.004	P-value joint significance	0	0.616
Observations	169,359	11,213	Observations	169,359	11,213
Clusters	3730	2044	Clusters	3730	2044
Panel C			*Panel D*		
Share of fathers -	-0.0028*	-0.0057	Share of parents -	-0.0092***	-0.0077
low occupations group	(0.0016)	(0.0053)	low occupations group	(0.0022)	(0.0067)
	[0.018]	[0.119]		[0.001]	[0.289]
Share of fathers -	0.0151***	0.0087	Share of parents -	0.0251***	0.0044
high occupations group	(0.0024)	(0.0089)	high occupations group	(0.0027)	(0.0095)
	[0.001]	[0.119]		[0.001]	[0.289]
P-value joint significance	0	0.269	P-value joint significance	0	0.355
Observations	169,359	11,213	Observations	169,359	11,213
Clusters	3730	2044	Clusters	3730	2044

Notes: occupations group: low (unemployed, homemaker, retiree), middle (self-employed, teacher, workman), high (manager, entrepreneur, professional). Panels A-B-C-D show estimates equation (2) using respectively: *rank*, *mothers, fathers,* and *parents* shares as regressors by immigrant status and gender. The regression includes school-year-mother's occupation-father's occupation fixed effects. The included controls are: mother's education, father's education. 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.