Labour Economics 85 (2023) 102450

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Ability composition in the class and the school performance of immigrant students☆

Elena Meschi [a,*], Caterina Pavese [b]

[a] University of Milan-Bicocca - Department of Economics, Management and Statistics DEMS; Center for European Studies (CefES), Italy
[b] Ifo Institute – Leibniz Institute for Economic Research at the University of Munich, Germany

ARTICLE INFO

JEL classification:
J15
I21

Keywords:
Peer effects
Immigrant students
Education

ABSTRACT

Using longitudinal data from the Italian National Institute for the Evaluation of the Education System (INVALSI), this paper investigates whether the ability of peers affects the educational attainment of immigrant students. We not only focus on the average quality of peers in the class, but also further investigate which part of the ability distribution of peers drives the effect, by assessing the role played by the extreme tails of the ability distribution. Our empirical strategy addresses students' endogenous sorting into classes by exploiting the within-student across-subjects variation in achievements and the simultaneity problem by using predetermined measures of peers' ability. We show that peers' ability matters. While native students are mostly influenced by the average quality of their peers, immigrant children are detrimentally affected by the share of very low achievers in the classroom. Our findings provide valuable guidance to policymakers concerning the allocation of students to classes in order to foster immigrant students' integration and learning.

1. Introduction

Over the past decades, many OECD countries, especially in Europe, have witnessed a sharp increase in the number of immigrants. In 2020, immigrants accounted for 10% of the total population in the European Union, up from 6.3% in 1990 and 8.2% in 2000 (UN Population Division). This surge gave rise to extensive empirical literature in economics that investigated the effects of immigration on host countries, focusing in particular on the labour markets (e.g., Dustmann et al. (2013), Cohen-Goldner and Paserman (2011) and Card (2001)) and more recently on the educational systems. In particular, a number of studies investigated the impact of immigrant students on natives' school achievements (see Figlio and Özek (2019); Frattini and Meschi (2019), Ballatore et al. (2018), Feld and Zölitz (2017), Ohinata and Van Ours (2013) among others). In most European countries, such interest was motivated by the concern that immigrant students would undermine the performance of native peers in the class. In fact, immigrant children generally tend to be disadvantaged students. Many come from lower socio-economic backgrounds, learn in a foreign language, and enter unfamiliar institutions; thus, they typically exhibit large gaps in performance.

In this paper, we challenge the existing approach of assessing the impact of immigrants on the educational outcomes of natives by instead focusing on how immigrant students are themselves affected by their classroom environment, and by the achievement of their native and immigrant peers. This study builds on the burgeoning literature on peer effects in education by investigating how and to what extent the ability of native and immigrant classmates affects their immigrant peers' educational outcomes. In particular, we estimate the impact of the average quality of peers in the classroom, and further investigate which part of the ability distribution of peers drives the effect, by assessing the role played by extreme tails of the ability distribution. In other words, we ask whether it is the average quality of peers that matters for immigrant students' achievement, or rather the presence of very high-achieving and very low-achieving peers.

Our empirical analysis is based on administrative data collected by the Italian National Institute for the Evaluation of the Education System (INVALSI henceforth) on four cohorts of students who completed lower secondary education in the academic years 2015/16, 2016/17, 2017/18 and 2018/19. Given that data have a longitudinal structure, it is possible to track students over time, and observe their

☆ We would like to thank Viola Angelini, Orazio Attanasio, Raphael Brade, Dalit Contini, Martin Fernandez-Sanchez, Andreas Steinmayr, Michela Tincani and participants of the University of Turin and University of Venice seminars, the V INVALSI workshop and the *Immigration in OECD countries — 11th edition* conference for helpful comments. Special thanks go to Patrizia Falzetti and Michele Cardone (INVALSI) for giving us access to INVALSI microdata. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

* Corresponding author.
E-mail address: elena.meschi@unimib.it (E. Meschi).

https://doi.org/10.1016/j.labeco.2023.102450
Received 1 April 2022; Received in revised form 5 September 2023; Accepted 10 September 2023
Available online 16 September 2023

prior achievements at the end of primary school. This allows us to construct indicators of peer quality in lower secondary schools based on predetermined measures of ability that are not simultaneously determined by students' own achievements. Our identification strategy follows Lavy et al. (2012) and exploits the within-pupil variation in achievement across subjects, as a way to solve the non-random sorting of students across schools and classes, which may bias the estimation of peer effects.

Our study contributes to the existing literature along several dimensions. First, in contrast to the extant literature drawing attention to the potential negative effect of immigrants on native students' learning and behaviour, we focus on understanding the impact of peers' quality on *immigrant* children's performance. Shifting the focus to immigrant children, we provide new insights into how class composition may help narrow the gap between immigrant and native children, and thus improve immigrants' integration into the school systems of their host countries. We believe this is a particularly relevant issue considering that adequate education is a prerequisite for socio-economic success and for overcoming the disadvantages facing immigrants in European societies. Providing a better school environment for immigrant children is an important social investment over the long term. Sustained positive growth effects may, in fact, arise from skilled immigrants fostering innovation through enhanced diversity, entrepreneurship, or international investment and trade (Bonin, 2017). Moreover, government budgets could improve, since education usually generates positive net fiscal returns. In addition, enhancing the educational outcomes of immigrants may help reduce income inequality and make host countries' societies more inclusive. Targeted policies to foster the school attainment of immigrants are thus crucial for reaping the potential of economic growth and social integration that comes with the challenges of immigration. Second, our identification strategy addresses some of the most serious problems in peer effects identification, such as students' endogenous sorting and peers' ability measurement, by exploiting both the within-pupil variation across subjects and the longitudinal structure of the data that provides predetermined measures of peers' ability, which help to solve the reflection problem.[1] Third, differently from Lavy et al. (2012), we are able to very precisely define the peer group because our data provides *class* identifiers (rather than *school* identifiers). This unique feature enables us to draw up a more accurate picture of students' interactions in comparison to studies carried out at a broader level (i.e., using peers' measures at the school level) that might fail to capture some relevant effect. As Carrell et al. (2009) point out, peer effects estimates can, in fact, vary greatly depending on the accuracy with which the set of relevant peers is identified.

Our results suggest the basic but important finding that peers' quality matters. We also show that immigrant and native children are differently affected by their peers. While immigrant children are detrimentally influenced by the proportion of lowest-performing peers in the class, native students are more strongly affected by the average quality of peers. Moreover, the negative impact of low-performing peers appears to be especially detrimental for the most fragile and less integrated immigrants (those who do not speak the Italian language at home and both parents were born abroad). Additionally, we find that peer effects are stronger within groups: natives exert a greater influence on native students and immigrant students on other immigrants. This evidence aligns with findings in the literature on racial peer effects, wherein the largest impacts observed are intra-racial and have little or no spillover into other racial groups (Fruehwirth, 2013; Hoxby, 2000).

The remainder of this paper proceeds as follows. Section 2 reviews the relevant literature on peer effects in education, discusses the primary identification challenges, and highlights our contribution to the existing literature. Section 3 provides the institutional background for our analysis, explaining the main characteristics of the Italian education system (Section 3.1) and of the immigrant population at schools (Section 3.2). Section 4 outlines the empirical analysis and discusses our identification strategy and possible threats to identification (Section 4.1). Section 5 describes the data and provides some relevant descriptive statistics. Section 6 presents results, while Section 7 discusses a number of robustness and sensitivity checks. Finally, Section 8 concludes with a discussion of the potential policy implications of this research.

2. Literature review

The importance of peer effects in educational outcomes has attracted considerable attention in the literature. The idea motivating this line of research is that peers matter when determining students' performance and behaviour. Students with higher incoming ability may, in fact, improve their classmates' performance, for example, by motivating them (through competition) to work harder or by enabling their teachers to teach at a higher level. By the same token, low-ability or disruptive students may harm their classmates' educational attainment by demanding more of the teacher's attention. Overall, the empirical literature tends to substantiate the claim that students generally benefit from proximity to better-performing peers around (see Sacerdote (2011) for a review), however, this effect is likely to be heterogeneous. High-performing students benefit from other high-achievers, while students with lower achievements benefit from peers who are performing slightly better than them (see, for example, Lavy et al. (2011); Imberman et al. (2012), Burke and Sass (2013)). On the other hand, high-performing peers do not always exert a positive influence on their peers. They may decrease their classmates' self-confidence, and in turn worsen their performance. A recent and growing stream of literature has analysed the role of relative comparison in the educational context and confirmed the existence of rank effects motivated by the impact of peers on self-esteem, conscientiousness, and expectations (see, for example, Elsner and Isphording (2017); Murphy and Weinhardt (2020), Bertoni and Nisticò (2023), Pagani et al. (2021)).

Empirically, the identification of both the size and nature of peer effects engenders a number of difficulties (Angrist, 2014, Manski, 1993). The two primary challenges for the identification of peer effects derive from the endogenous selection of students into a specific group (non-random sorting), and the simultaneous determination of the outcomes of students belonging to the same group (simultaneity).

The former problem relates to the fact that students are not randomly allocated to schools and classes. They self-select into schools and classrooms based on their observable and unobservable characteristics. In the context of this study, for instance, immigrant children tend to have a lower socio-economic background, and less access to information about the characteristics of local schools. This means they are more likely to be clustered into lower-quality schools than their native counterparts.

Previous studies adopted various empirical strategies to deal with this endogenous sorting of students across schools. One method is to rely on some form of exogenous variation in student assignment to schools or classrooms. Duflo et al. (2011), for example, exploit the variation in peer composition generated by actual randomisation, while Angrist and Lang (2004) rely on a substantial increase in the number of disadvantaged black or other minority students in the schools in Boston's rich suburbs due to the Metropolitan Council for Educational Opportunity's (Metco) desegregation programme. Similarly, Gould et al. (2009) rely on variation in the number of immigrant students caused by the exogenous immigration waves to Israel in the early 1990s, while Ballatore et al. (2018) use the exogenous variation in the number of natives and immigrants owing to the compulsory cap of 25 students per class in Italian primary schools. Figlio et al. (2023) exploit the variation in test scores of siblings who experience

[1] As many studies in the peer effects literature, we do not aim to separately identify endogenous and exogenous peer effects (Manski, 1993).

Table 1
Guiding principles followed by school principals' in the allocation of students to classes.

Criterion	Percentage
Guarantee that classes are internally heterogeneous in terms of learning levels	63.70
Guarantee that classes are internally heterogeneous in terms of socio-demographic characteristics	17.52
Assign students to classes at random	0.03
Other criteria	18.78

Notes: The table reports the guiding principles underlying the allocation of students to lower secondary school classes as reported by a nationally representative sample of school principals surveyed by INVALSI in the school year 2013/14.

different cumulative exposure to school-cohort-specific peers' characteristics, holding the heterogeneity in family life-cycle fixed. Another common method for dealing with endogenous sorting of students across schools is to use school fixed effects models to control for the unavoidable self-selection into schools and exploit the idiosyncratic within-school variation in peer characteristics across adjacent cohorts (Hoxby, 2000; Ammermueller and Pischke, 2009; Lavy et al., 2011, Tonello, 2016, Gibbons and Telhaj, 2016) or across different classes (Frattini and Meschi, 2019; Ohinata and Van Ours, 2013; Contini, 2013).

Even if controlling for school fixed effect accounts for most unobserved heterogeneity, sorting of students within schools across classes or cohorts, could still be potentially non-random. In this case, the estimation of peer effects would be biased. Lavy et al. (2012) improve on this strategy by applying within-pupil regressions and exploiting the variation in achievements across three compulsory subjects tested.

The second empirical challenge in the estimation of peer effects lies in the fact that individual and peers' achievements are simultaneously determined, which means causal inference is only possible if the peer group's predetermined ability measures are available. Some papers overcome this problem by measuring peer quality by fixed (and therefore predetermined) students' characteristics, such as the socio-economic background (see for example Hoxby (2000) and Ammermueller and Pischke (2009)). However, as underlined by Hoxby and Weingarth (2005) the attributes of a student's background have little or no effect on students' outcomes once peers' achievements are properly controlled for. When longitudinal data are available, it is possible to use past achievements to measure peer ability. For example, Lavy et al. (2012) define the ability of 8th grade peers using their predetermined 5th grade test score. Taking advantage of the compulsory transition between primary and lower secondary school, which leads to a substantial reshuffling of peers, these measures are conceivably exogenous to 8th grade outcomes (see also Lavy et al. (2011), Gibbons and Telhaj (2016) that use a similar strategy).

In terms of findings, the existing literature tends to conclude that peer effects, even when statistically significant, are rather modest: most studies find that a one standard deviation increase in average peers' quality raises outcomes by less than 10 percent of a standard deviation (see Lavy et al. (2011)). More recent studies have further demonstrated that the effect of peers is not constant, but greatest when peers are students either at the very bottom or the very top of the academic ability distribution. For example, Lavy et al. (2011) find that a high proportion of low-achieving students reduces classmates' educational outcomes, the effect of which leads to focusing teachers' attention on struggling students and increasing the level of violence and disruption in the class. Similarly, Lavy et al. (2012) provide evidence that a large proportion of very low-achieving peers decreases the educational performance of the other classmates. Fruehwirth (2013) illustrates how high-achieving students benefit the most from high-achieving peers. Similarly, Gibbons and Telhaj (2016) find that test scores of low-achieving students in the U.K. are undermined by the presence of high-achieving students, while upper-middle achieving students benefit from the presence of high-achieving students. Feld and Zölitz (2017) show that while students benefit from better peers on average, low-achieving students are negatively impacted by high-achieving peers. Overall, recent papers provide evidence of non-linearities in peer effects and suggest that different students react differently to the ability of their peers. In this study, we investigate whether immigrants' performance is affected by the ability of their peers, and whether these effects are non-linear and concentrated in the tails of the ability distribution.

Our paper contributes to the existing literature in three important and distinct ways. First, it focuses on how peers' ability affects how immigrant students perform in school. In spite of extensive literature analysing the impact of immigrants and ethnic minorities concentration in the class on native students, little is known about the impact that peers have on migrants and ethnic minorities (Schneeweis, 2015; Jensen and Rasmussen, 2011; Ohinata and Van Ours, 2013). To the best of our knowledge, this is the first study on ability peer effects specifically focusing on the educational achievement of immigrant children. Irrespective of the increasing number of immigrant students in all OECD countries' educational systems, there is scarce evidence about which policies to implement in order to foster their integration and learning.[2] In this paper, we contribute to the debate by showing how class composition impacts the educational achievement of immigrant students. Second, following Lavy et al. (2012), we adopt an identification strategy that addresses some of the most severe problems in peer effects identification, such as students' endogenous sorting across school and classes and peers' ability measurement, by exploiting both within-pupil variation across subjects and the longitudinal dimension of the data that provides predetermined measures of peers' ability. Third, in deviation from Lavy et al. (2012) we are able to precisely define the peer group since our data provide *class* identifiers (rather than *school* identifiers). Finally, we further investigate whether peer effects stem from the average quality of classmates, or from the very bright and very poor achievers in the class.

3. Institutional background

3.1. The Italian education system

Education in Italy begins at age 6 and is compulsory until age 16. The school system is organised in two cycles. The first cycle comprises primary and lower secondary education. Primary education lasts for five years, while lower secondary education, the focus of our analysis, starts at age 11 and lasts for three years. When moving from primary to lower secondary school, students change schools, and classes are reshuffled. Without having to pass any kind of examination, students enter lower secondary school with virtually all new classmates. Once assigned to these classes, students remain with this set of peers from 6th grade until the end of 8th grade.

Each school's principal oversees the allocation of children to their classes. Formation criteria are centrally established, and each school provides relevant information in official online documents. Generally, these criteria establish that students should be equally distributed by ability, gender, and economic and social background. Table 1 reports

[2] For other educational policies affecting the school performance of immigrant children, see, among others, Corazzini et al. (2021) on the impact of early childcare on immigrant children's cognitive outcomes and Carlana et al. (2022) on the impact of providing tutoring and career counselling to high-ability immigrant students.

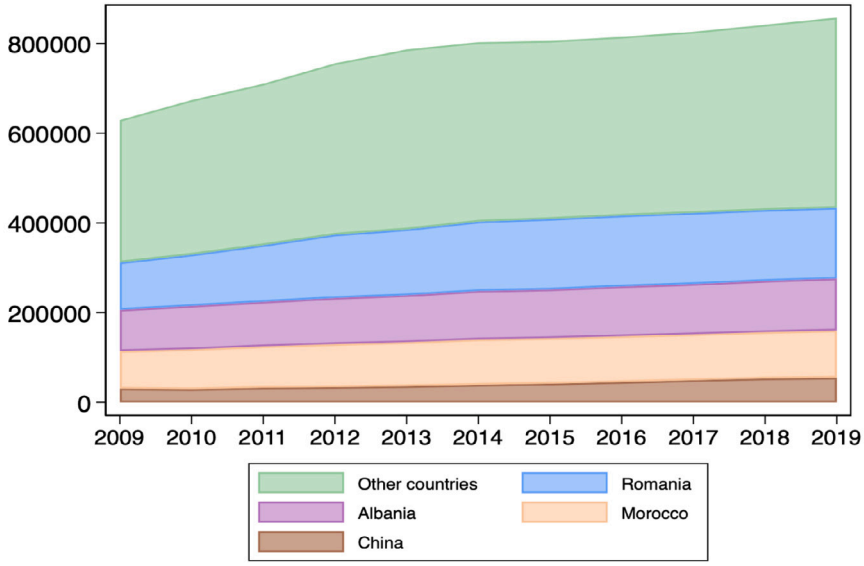

Fig. 1. Number of foreign students in Italy by country of origin, 2009–2019.
Source: MIUR data.

the main guiding principles followed by school principals when allocating students to lower secondary school classes, and confirms that the most relevant aspects in the class formation process are comparability across classes and heterogeneity within classes in the same school.

The Italian school system is mainly public and does not allow for students to be tracked by ability (Eurydice, 2013). Pupils with similar abilities or educational attainment cannot be grouped together for specific subjects, such as mathematics or science. In fact, children are assigned to the same class for all subjects and taught by the same teachers. Class size ranges from a minimum of 18 students to a maximum of 27. The weekly class schedule is of 30 h, with reading and maths occupying the most instruction hours. The academic disciplines, time of instruction, educational programmes and their content are centrally defined by the Ministry of Education and therefore consistent across all Italian schools.

3.2. Immigrant students in Italian schools

Immigration is a relatively recent phenomenon in Italy. Between 1990 and 2019, the number of immigrants (foreign-born) in Italy increased from 780,000 to 5.2 million, meaning that the share of immigrants in the Italian population has risen from 1.4% to 8.7% over the last 30 years (Frattini and Vigezzi (2018) and www.demo.istat.it). The majority of immigrants in Italy come from low and middle-income countries and, on average, are characterised by a lower socio-economic background than natives (Carlana et al., 2022). The rapid surge in immigration to Italy also implies rising shares of students with an immigrant background in the education system. In the school year 2018/19, immigrant students accounted for 10% of the student population in Italy. As shown in Fig. 1, over the last decade, the number of immigrant students increased by 27.3%, from 673,000 in 2009/10 to 857,000 in 2018/19, with immigrant students from Romania being the most represented (18.4%) followed by Albanian (13.5%), Moroccan (12.3%) and Chinese (6.4%).

As reported in Fig. 2, immigrants tend to concentrate in the Centre-North of the country (87%), with Lombardy and Emilia-Romagna regions hosting 25.4% and 11.9% respectively (Ministero Della Pubblica Istruzione, 2020). Given the described distribution of immigrant students in the Italian territory, our analysis does not include the South and the Islands where on average immigrants account for less than 3% of the student population. In Section 7, we show that this choice does not alter the external validity of our empirical analysis.

The left panel of Fig. 3 reports the share of first- and second-generation immigrant students in lower secondary schools, which are the focus of our paper. The figure confirms the rise of students with an immigrant background and reveals that second-generation migrants (children born in Italy to non-native parents) account for the majority of immigrant children in lower secondary schools. Despite we do not observe the country of origin of children in our sample, in the school years 2018 and 2019 students were asked to report the language most frequently spoken at home.[3] The graph in the right panel of Fig. 3 shows the percentage of migrants by language most frequently spoken at home. It is clear to see that the language is a relevant source of heterogeneity: more than 60% of immigrant children speak a language at home that is different from Italian. This may explain the educational disadvantage of immigrant children that we will discuss in Section 6.2.

4. Empirical strategy and identification

The aim of our analysis is to estimate the impact of peer academic quality on students' test scores in 8th grade, at the end of lower secondary schools. The main challenge in estimating peer effects is the non-random sorting of students across schools and in some cases also between classes within schools (see also Angrist (2014)). Our identification strategy follows Lavy et al. (2012) and relies on within-pupil regressions that exploit the variation in attainments across two compulsory subjects tested at grades 8 and 5.[4] This strategy allows us to determine whether within-student variation in reading and maths test scores is systematically associated with differences in peers' ability across subjects. In other words, this analytical framework examines whether a student's exposure to high- or low-achieving peers in a given subject translates into an improvement or a decline in their own performance in that subject. An advantage of this approach is that by including individual fixed effects, we are able to control for a student's own unobservable average ability across the two subjects as well as for unmeasured family and class characteristics.

[3] In all school years children are asked to report the language most frequently spoken at home but only in the school years 2018 and 2019 are they required to provide the specific language rather than simply reporting whether it is a language different from Italian.

[4] See also Lavy (2015) and Dee (2007) for similar strategies exploiting within-student across subjects variation.

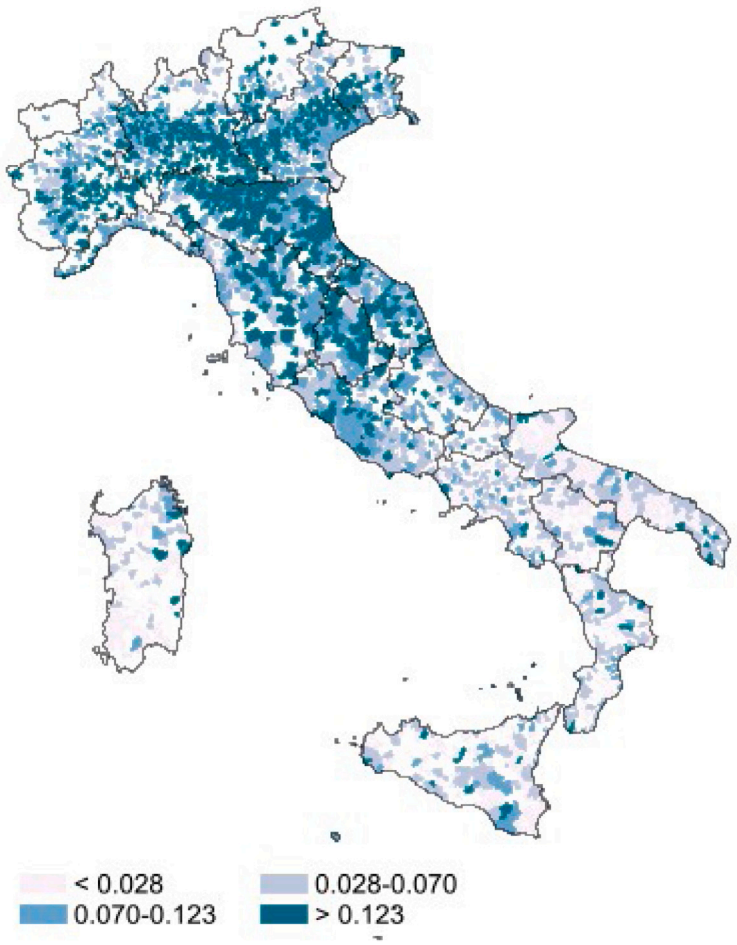

Fig. 2. Distribution of the share of migrants in Italy, year 2018.
Note: The map shows the share of 8th grade immigrant students across Italian municipalities in 2018. Break points are quartile intervals in the share of migrants.
Source: INVALSI data.

More specifically, following the notation in Angrist (2014), we estimate the following regression equation:

$$Y_{isct} = \alpha_i + \gamma_{st} + \beta_s xGender + \delta_1 P_{(i)sct} + \delta_2 P^h_{(i)sct} + \delta_3 P^l_{(i)sct} + \epsilon_{isct} \quad (1)$$

where the dependent variable Y_{isct} measures 8th grade achievements of child i in subjects s in lower secondary school class c in cohort t. Specifically, each child is tested on two compulsory subjects: reading and maths. Our main variables of interest are $P_{(i)sct}$, $P^h_{(i)sct}$, and $P^l_{(i)sct}$ that respectively capture the average ability in 5th grade in subjects s of 8th grade peers in class c, and the proportion of very high and very low ability peers in class c. In particular, we define high-ability and low-ability peers as students at the top and bottom 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. Crucially, all peer measures related to student i are calculated leaving i him or herself out of the calculation. Individual and subject-by-cohort fixed effects are captured respectively by α_i and γ_{st}. Furthermore, we include subject-by-gender fixed effects, $\beta_s xGender$, to account for gender-specific differences across subjects that might potentially affect the sorting of students into lower secondary schools (Fryer and Levitt, 2010). Finally, ϵ_{isct} is the error term. Standard errors are clustered at the class level. The parameters of interest are δ_1, δ_2 and δ_3 which respectively capture the effect of the average ability of 8th grade peers and the impact of the proportion of high- and low-achieving peers on students' outcomes.[5]

All peers' ability measures are computed based on 5th grade test scores, which are predetermined and thus do not suffer from the reflection problems. This is especially true in our setting where children at the end of 5th grade make a compulsory transition between primary and lower secondary school, implying a significant change to the class group composition. On average, in our data, 71% of 8th grade students in the same class did not attend the same primary school, and thus that they could not have mutually affected their 5th grade test scores.

We exploit this high inter-school mobility to build separate measures of peer quality for new and old peers.[6] New peers are defined as peers who are in a student's 8th grade class but were not in their 5th grade class. Old peers are those who are together in the 8th grade, and were also in the same 5th grade class. In our empirical analysis, we mainly focus on new peer quality measures because they are certainly immune to the reflection problem. In fact, student i 5th grade test score is predetermined, and thus not affected by 8th grade new peers' outcomes. Differently from Lavy et al. (2012), we define peers at the *class* level. In Italy, peer group composition is constant throughout the lower secondary school period, and most of the students' interactions take

[5] In our baseline specification we do not control for students' own test score in 5th grade, but as shown in columns 3–6 of Table 8, our results are mostly unaltered when it is included among controls.

[6] Gibbons and Telhaj (2016) are the first to make this empirical distinction.

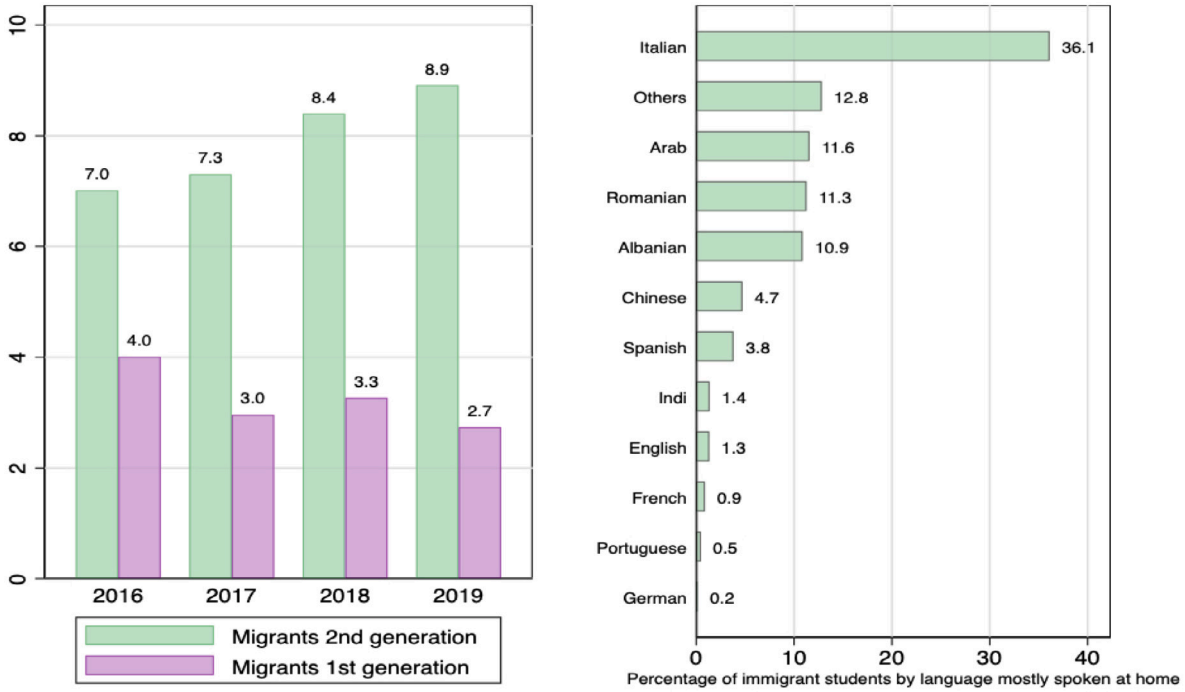

Fig. 3. Immigrant students in our sample.
Note: Panel A shows the percentage of first- and second-generation migrants over time in lower secondary school. Panel B reports the percentage of immigrant students by language most frequently spoken at home in the s.y. 2018 and 2019 (for the s.y. 2016 and 2017 the information is not available).
Source: INVALSI data.

place within classes more than across classes (students have no choice but to spend time together and stay in the same class together for the three years during lower secondary school). Finally, it is worth noting that our empirical approach imposes some restrictions on the types of peer effects that we can identify. First, peer effects are assumed to be the same for the two subjects. Second, the effect we are considering excludes potential spill overs across subjects. For instance, having high-achieving peers in maths could affect students' reading test scores and vice versa. However, if this were the case, we would be bound to find no effects.

4.1. Threats to identification

Our empirical strategy exploits within-student variation in attainment across subjects and allows to control for students' own unobservable average ability across subjects and unmeasured individual and school or class characteristics. This approach should therefore achieve a clean identification of peers' effects, under the main identifying assumption that peers' subject-specific ability is unrelated to unobserved determinants of individual students' subject-specific skills.

A first concern that may threaten this assumption is the subject-specific sorting of students across schools and classes. This could happen, for example, if a specific school or class specialises in maths or languages and attracts students based on their ability in those specific subjects. In this case, we would observe a correlation between individual variation in grade 5 test scores and within-student across-subject quality of peers. In the Italian context, this is unlikely because ability grouping is not permitted, and educational programmes are all centrally drawn up by the Ministry of Education. In other words, lower secondary schools are not allowed to tailor their educational offer or specialise in a specific subject. Furthermore, following Lavy (2015), schools that do not use ability as an admission criterion are unlikely to select students based on subject-specific considerations. However, in order to exclude the possibility of subject-specific sorting, in Table 8 of Section 7, we perform a falsification test where we regress the ability of peers in grade 8 on students' own ability in grade 5. Reassuringly, the estimates reveal that there is no significant correlation between within-student across-subjects variation in prior achievements, and the variation in peers' ability across subjects. This evidence is also corroborated by the fact that when we augment our specification including students' own test scores at grade 5 (see columns 4–6 of Table 8), our coefficients of interest remain remarkably stable.

A second related concern is the non-random allocation of teachers to classes. This could happen, for instance, if highly experienced teachers prefer to teach better classes: even if grouping students by ability is not feasible, a class that ends up being better in a specific subject may be allocated to a better teacher in that subject. Or, on the contrary, school principals may allocate the best teachers in a specific subject to classes where students are particularly low-performing in that subject. In order to deal with this concern, we exploit detailed and novel data on teachers' characteristics provided by INVALSI for a nationally representative sample of schools as described in Section 7. Using this data, we perform a balancing check to illustrate how teachers are in fact randomly allocated to classes within schools (see Table 9, Section 7).

Finally, in Tables 10 and 11, we further check the robustness of our findings to alternative samples and estimation strategies by leveraging variation within schools across adjacent cohorts and classes that are commonly used in the peer effect educational literature (see for example Ammermueller and Pischke (2009)). To conclude, in Table 12, we test the sensitivity of our results to changes in thresholds used to define high(low)-achieving peers.

5. Data and descriptive statistics

Our analysis relies on administrative data collected by the INVALSI, the independent public agency in charge of evaluating the Italian school system and monitoring student achievement in reading and maths. Student attainment is evaluated on an annual basis upon completion of 2nd, 5th, 8th, and 10th grade. All students in these grades are required to take part in the INVALSI assessment. The tests are administered on the same day, and assessments are carried out externally following a predetermined marking scheme. This feature of the data makes student

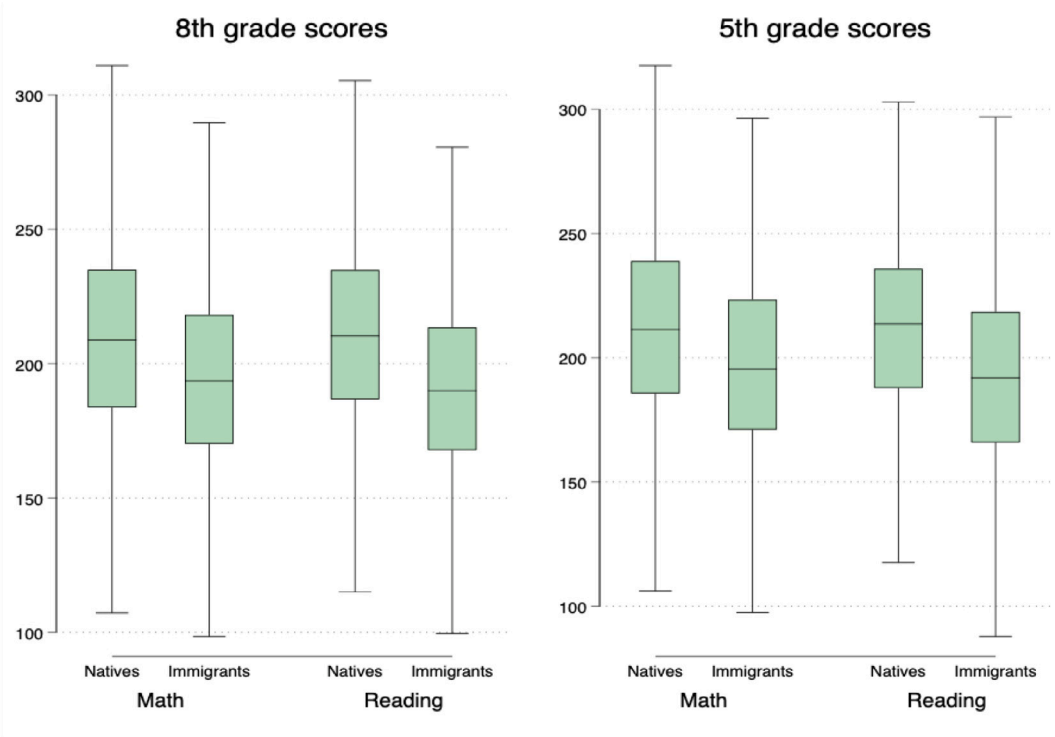

Fig. 4. Immigrant students' gap in 8th and 5th grade by subject.
Note: The figure shows the boxplots of the standardised test scores by subject in 8th grade (left graph) and 5th grade (right graph) across immigrant and native students. The black line shows the median values, while the boxes represent the interquartile range. All test scores are standardised to have a mean equal to 200, and a standard deviation equal to 40 for each subject and cohort.

performances wholly comparable across all Italian schools and classes. Testing comprises both multiple-choice and open-ended questions to assess students' key competencies in reading and maths. In particular, the reading test evaluates mastery of grammar and reading comprehension, while the maths test measures skills in problem-solving, logic and interpretation of quantitative phenomena. All test scores are standardised to have a mean equal to 200, and a standard deviation equal to 40 for each subject and cohort.

This analysis focuses on the four cohorts of students who completed the INVALSI test when they had finished 8th grade in the school years from 2015/16 to 2018/19. The longitudinal structure of the data and the availability of class identifiers at each school stage allow us to relate each 8th grade student to their 5th grade test scores, as well as those of their peers. We exploit this feature of the data to compute predetermined measures of students' ability and peer quality for lower secondary school peers who either attended (i.e., *old peers*) or did not attend (i.e., *new peers*) the same primary school class.

INVALSI data also includes information on a number of demographic characteristics of children and their families, including their gender, ethnicity, their parents' educational levels and working conditions and the index of Economic, Social and Cultural Status (ESCS).[7] We use this information to define students' immigrant status. A child is determined to be an immigrant if both parents are non-Italian citizens, regardless of their place of birth (in Italy, citizenship is acquired according to the *Ius Sanguinis* principle). Native children are those who have at least one parent with Italian citizenship.

Our sample is restricted to students attending a public school in the Northern and Central regions of Italy. Southern regions are excluded because most of the immigrants live in the North and Centre (Frattini and Vigezzi, 2018), Southern test scores are considered less reliable due to the higher incidence of cheating behaviours (Lucifora and Tonello, 2015, Bertoni et al., 2013) and finally because many classes in the South did not take part in the INVALSI test days in the 2014/15 school year due to a strike in the school sector.[8] This means that 5th grade test scores cannot be linked with 8th grade test scores, a necessary condition for inclusion in this study sample.[9] The sample is further restricted to students with no missing observations on the variables used in the analysis. Our final sample includes 218,688 and 1,740,210 observations for immigrant and native students respectively.

Table A.7 in the Appendix lists and describes all variables included in our empirical analysis, while Table 2 presents key summary statistics (mean and standard deviation). Column (1) reports statistics for "regular" students, Column (2) for "top" students and Column (3) for "bottom" students. Regular students are defined as those with age 11 (i.e., grade 5) test scores in both reading and maths above the 5th percentile, and below the 95th percentile of the grade 5 test score distribution. Top and bottom students are those who achieved above the 95th percentile and below the 5th percentile, respectively, in at

[7] The ESCS (Economic Social and Cultural status) index describes the socio-economic and cultural status of students' families. It is developed using information provided by the students' and the schools' questionnaires regarding parents' educational level and working status, as well as the material possession of some specific goods including books, internet connections, and a personal computer. The index is calculated using principal component analysis and, by construction, has mean zero and unitary variance.

[8] In Campania, Apulia and Sardinia, less than 75% of the classes took part in the INVALSI test. In Sicily, it was less than 30%. See INVALSI (2015) for additional details.

[9] The average attrition rate (share of students observed in grade 5 but not in grade 8) in our sample is about 8%. This rate is higher for the sample of immigrant students (23%). In Table A.1, we investigate if attrition is systematically correlated with individual characteristics for the sample of immigrants (column 1) and natives (column 2) separately. The estimates reveal that, for both natives and immigrants, females, students with highly educated fathers, students who attended pre-primary education, and students with higher test scores at grade 5 are slightly less likely to dropout.

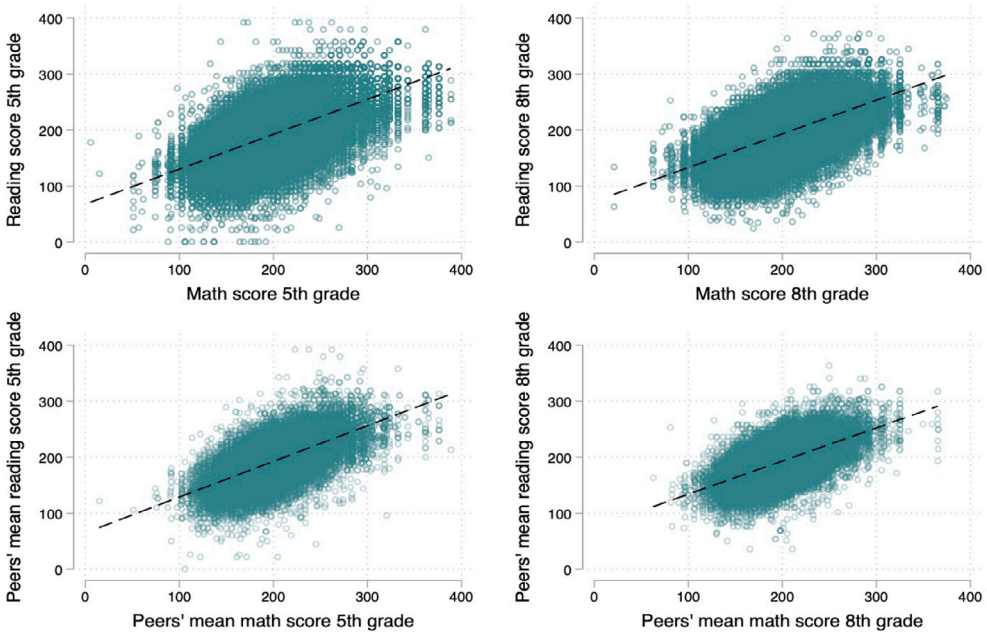

Fig. 5. Correlations between students' own test scores and peer's test scores across subjects.
Note: The top right (left) panel of the figure shows the correlation between student's *i* test score in maths and reading in 5th (8th) grade. The bottom right (left) panel shows the correlation between peers' average test scores in maths and reading in 5th (8th) grade. Dashed lines are the lines of best fit.

Table 2
Descriptive statistics.

	(1)	(2)	(3)	(4)	(5)	(6)
	Immigrants			Natives		
	Regular	Top	Bottom	Regular	Top	Bottom
Panel A: Test scores						
8th grade reading test scores	195.72	227.06	156.38	210.13	244.38	167.36
	(31.85)	(38.32)	(28.77)	(32.46)	(34.71)	(28.17)
5th grade reading test scores	198.69	259.45	138.04	211.98	271.01	150.3
	(29.47)	(37.96)	(27.84)	(29.46)	(33.21)	(28.52)
8th grade maths test scores	199.59	239.7	166.44	209.74	251.02	167.48
	(33.5)	(42.58)	(28.8)	(34.57)	(39.7)	(27.85)
5th grade maths test scores	203.164	283.177	153.58	211.58	282.02	150.98
	(30.62)	(34.63)	(28.63)	(30.79)	(34.65)	(26.94)
Panel B: Demographic characteristics						
Female	0.52	0.47	0.53	0.5	0.46	0.51
	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)
ESCS	−0.38	−0.13	−0.64	0.22	0.55	−0.16
	(0.83)	(0.87)	(0.81)	(0.88)	(0.87)	(0.87)
Immigrant first-generation	0.27	0.26	0.37			
	(0.2)	(0.44)	(0.48)			
Immigrant second-generation	0.72	0.74	0.63			
	(0.2)	(0.44)	(0.48)			
Observations	87,616	4310	17,446	742,469	76,482	51,234

Notes: The table shows the mean and standard deviation in parenthesis. Panel A displays students' educational outcomes, Panel B students' demographics. We restrict the sample to students that have INVALSI test scores in both the 5th and 8th grade. All test scores are standardised to have a mean equal to 200 and a standard deviation equal to 40 for each subject and cohort. For the definition of the variables, see Table A.7 in the Appendices.

least one of the two subjects tested at grade 5. By construction, regular students' 5th and 8th grade test scores are concentrated around 200, while those of pupils in the top and bottom 5 percent will either outperform or fall behind any other student. Concerning students' demographic characteristics, stark differences become apparent across the three groups. For instance, the lowest-performing students are least likely to have highly educated and working parents. They are also more likely to have lower ESCS values and an immigrant background (in either the first- or second-generation). The opposite picture emerges for the highest-performing students, who generally have a higher ESCS, better-educated parents, and are more likely to be native citizens.

To substantiate our claim on the existing educational gap between native and immigrant children, in Fig. 4 we present 8th grade (left panel) and 5th grade (right panel) average students' test scores by immigrant status and subject. The figure makes it clear that immigrant students perform worse than natives in reading and maths both in the 5th and 8th grades.

In this study, the identification of peer effects is based on a comparison of the same student's performance in two different subjects. Hence, identification is possible only if, for each student, there is enough variation across subjects in their own and in their peers' test scores. In the upper panel of Fig. 5, we show the correlation between

students' own test scores in reading and maths in grades 5 and 8, while the bottom panel reports the correlation between class average test scores in reading and math, measured in grades 5 and 8. The figure reveals that despite all correlations being high and positive, there is still substantial variation that we can exploit in our regressions.

Similarly, the top panel of Fig. A.1 in the Appendix reports the distribution of the difference in subjects' averages within classes, and confirms that there is a significant amount of within-pupil across-subject dispersion in average peers' 5th grade test scores. In addition, when analysing the role of top and bottom peers, our strategy requires classes with enough variation in the shares of bottom (top) performers between the two subjects. In the bottom panel of Fig. A.1, we present the distribution of the difference between the shares of top (left panel) and bottom (right panel) peers across the two subjects. Despite being concentrated around zero, the figure illustrates how we can leverage a fair amount of variation in the share of top (bottom) peers across subjects to identify the impact of the quality of peers.

To further corroborate the fact that all test scores exhibit quite a wide range of variation, Table A.2 in the Appendix reports means and standard deviations (overall, between- and within-students) of 5th and 8th grade test scores and peers' quality measures, for immigrant (Panel A) and native (Panel B) students. Although, as expected, most of the variation in test scores is explained by the between-students variation, Table A.2 shows that there is substantial within-student variation across subjects: this evidence confirms that test scores for the same student are not perfectly correlated across subjects and there is enough variation for us to leverage. Similarly, there is some relevant variation in the quality of peers across the two subjects, too. The latter holds both for immigrants and for native students. Overall, the descriptive evidence in this section supports our identification strategy by showing that there is substantial within-student variation in test scores and peers' quality.

6. Results

This section reports our results on the causal effect of peer quality on students' educational outcomes in 8th grade. Section 6.1 presents our baseline estimates, while Section 6.2 explores various dimensions of heterogeneity in the effect of peers' quality and discusses potential mechanisms at work. Section 6.3 further discusses our findings and compares the results for immigrant and native students.

6.1. Baseline estimates

Table 3 presents our results on how peers' quality affects the test scores of immigrant students in the 8th grade. Columns (1) and (3) report Ordinary Least Square (OLS) estimates while columns (2) and (4) report Within-Student (WS) estimates described in Section 4. Columns (1) and (2) present OLS and WS estimates of the effect of average peers' quality, while columns (3) and (4) report OLS and WS estimates of the impact of the percentage of bottom and top 5% peers. All specifications control for subject-by-cohort and subject-by-gender fixed effects. In OLS regressions, we also include cohort fixed effects and controls for students' demographic characteristics (gender, immigrant status, quarter of birth, and the ESCS indicator measured in 5th grade). The outcome variables are reading and maths test scores, standardised to have a mean of 200 and a standard deviation of 40 for each subject and cohort. Peer quality is predetermined and measured as the average of 5th grade test scores of peers in the 8th grade (mean score), as well as by the proportion of peers in each class that are below the 5th percentile (bottom) or above the 95th percentile (top) of the cohort-specific national distribution of 5th grade test scores.

The OLS estimates in columns (1) and (3) of Table 3 indicate a positive and significant correlation between subject-specific peer quality and student achievement. In particular, the coefficient of peers' mean scores is about 0.15, while the proportion of top students in the class is associated with an increase in test scores of about 9 points (corresponding to about 22 percent of the overall standard deviation) and the proportion of bottom peers is associated with a reduction in test scores of about 30 points (which is some 75 percent of a standard deviation). Clearly, these estimates are biased by non-random sorting of students across schools: the positive association that we observe could simply be due to better students choosing better schools. When we address this issue, including student fixed effects in columns (2) and (4) and exploiting the random variation in peer quality across different subjects, as expected the coefficients of peers' variables shrink substantially. In particular, once we add student fixed effects, the coefficient of peers' mean score changes from 0.15 to 0.09, and the coefficients of the share of top and bottom peers diminish respectively to 4.4 and 10.8. This drop in the point estimates' size indicates that the inclusion of student fixed effects is effective for controlling for within- and between-schools endogenous sorting by eliminating unobserved student and school characteristics. The direction of the implied bias suggests that there is a positive selection: students with higher unobserved ability are more likely to be assigned to schools (and classes within schools) with higher average peer quality. The reduction in size of the estimated coefficients may also be due to the fact that WS estimates net-out spill overs across subjects. The rest of the paper will focus only on our preferred and most stringent specification that includes student fixed effects.

Table 3
Impact of peer quality on 8th grade outcomes of immigrant students. OLS and WS estimates.

	(1) OLS	(2) WS	(3) OLS	(4) WS
Peers' mean score	0.149***	0.092***		
	(0.006)	(0.013)		
Share of top 5% students			9.905***	4.356**
			(1.357)	(1.879)
Share of bottom 5% students			−29.826***	−10.806***
			(1.322)	(2.180)
Observations	217,338	218,688	217,338	218,688

Notes: The table reports the regressions' coefficients of the average peer quality (columns 1–2) and the proportion of top and bottom peers (columns 3–4) on immigrant students' 8th grade standardised test scores. Columns (1) and (3) report OLS estimates. Columns (2) and (4) report within-student (WS) across-subjects estimates. The dependent variables are the INVALSI 8th grade test scores in reading and maths. All test scores are standardised to have a mean equal to 200 and a standard deviation equal to 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include gender-by-subject fixed effects and cohort-by-subject fixed effects. Columns (1) and (3) additionally control for cohort fixed-effect, gender, immigrant status, quarters of birth, and the ESCS index measured in 5th grade. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

Results in Table 3 document that the quality of peers matters for immigrant students' performance. However, the peer measures in Table 3 are computed based on the abilities (measured in grade 5) of all classmates in grade 8. These estimates might suffer from reflection problems if some peers in grade 8 attended the same primary school as student *i* (old peers). In this case, 5th grade test scores of the student *i* and those of her old peers would be simultaneously determined. Therefore, in Table 4, we construct peer quality measures separately for new peers and old peers and focus on the effect of *new peers* on students' achievements.

In our main specifications, reported in columns (1), (2), and (3) of Table 4, we control for the average test scores of old peers, but as shown in columns (4)-(6), our estimates are not sensitive to the inclusion of this variable.[10]

[10] In Table A.3, we also augment our specification by controlling for the class share of top and bottom old peers. The results suggest that our main coefficients of interest are essentially unaffected when including these controls.

Table 4
Impact of new peers' quality on 8th grade outcomes of immigrant students.

	(1)	(2)	(3)	(4)	(5)	(6)
Peers' average score (new peers)	0.005		−0.008	0.046***		0.032**
	(0.011)		(0.012)	(0.011)		(0.013)
Share of top 5% students (new peers)		0.028	0.758		3.653	0.624
		(2.303)	(2.61)		(2.356)	(2.68)
Share of bottom 5% students (new peers)		−6.936***	−7.499***		−10.349***	−8.012***
		(2.526)	(2.626)		(2.685)	(2.748)
Control for old peers' quality	yes	yes	yes	no	no	no
Observations	218,688	218,688	218,688	218,688	218,688	218,688

Notes: The table shows WS estimated coefficients of the average quality of new peers, the proportion of top and bottom new peers on students' 8th grade standardised test scores. "New peers" refer to students in grade 8 in a given cohort that do not come from the same primary school. All test scores are standardised to have a mean equal to 200 and a standard deviation equal to 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include student fixed effects, as well as gender-by-subject fixed effects and cohort-by-subject fixed effects. Columns (1)–(3) additionally control for the average quality of old peers. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

Table 5
Impact of peer quality on 8th grade outcomes of immigrant students. Heterogeneity by gender, socioeconomic background and initial ability.

	(1)	(2)	(3)	(4)	(5)	(6)
	Gender		ESCS		Initial ability	
	Females	*Males*	*High*	*Low*	*High*	*Low*
Peers' average score (new peers)	−0.016	0.001	−0.031	−0.000	−0.005	−0.009
	(0.017)	(0.017)	(0.024)	(0.014)	(0.022)	(0.015)
Share of top 5% (new peers)	1.663	−0.302	8.499*	−1.987	−3.654	3.786
	(3.529)	(3.597)	(4.919)	(2.974)	(4.209)	(3.177)
Share of bottom 5% (new peers)	−7.064**	−7.972**	−5.157	−7.798***	−1.793	−9.475***
	(3.470)	(3.513)	(4.752)	(3.022)	(4.633)	(3.023)
Observations	113,774	104,914	53,936	163,402	69,024	149,664

Notes: The table reports WS estimates of the effect of the average quality of new peers, proportion of top and bottom new peers on students' 8th grade standardised test scores by gender (male/female), socio-economic background (ESCS higher/lower than the median) and initial ability (average test score across subjects higher/lower than the median). The dependent variables are the INVALSI test scores in reading and maths. All test scores are standardised to have a mean equal to 200 and a standard deviation equal to 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include the average quality of old peers, gender-by-subject fixed effects and cohort-by-subject fixed effects. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.001$.

Table 6
Impact of peer quality on 8th grade outcomes of immigrant students. Heterogeneity by immigrant status, language spoken at home and parents' nationality.

	(1)	(2)	(3)	(4)	(5)	(6)
	Immigrant status		Language spoken at home		Parents' nationality	
	1st gen	*2nd gen*	*Italian*	*Other language*	*Both foreign*	*One Italian*
New peers' average score	−0.002	−0.010	−0.010	−0.010	−0.007	−0.031
	(0.022)	(0.014)	(0.019)	(0.016)	(0.013)	(−0.056)
Share of top 5% (new peers)	−3.070	2.417	1.154	0.330	0.462	6.669
	(4.662)	(2.985)	(3.823)	(3.45)	(2.674)	(10.833)
Share of bottom 5% (new peers)	−9.022**	−6.754**	−3.379	−10.266***	−7.643***	−3.881
	(4.363)	(3.135)	(3.844)	(3.355)	(2.685)	(11.613)
Observations	63,360	155,328	82,546	127,898	208,846	9842

Notes: The table reports WS estimates of the effect of the average quality of new peers, the proportion of top and bottom new peers on students' 8th grade standardised test scores by immigrant status (1st/2nd generation) (column 1–2), language most frequently spoken at home (column 3–4), and parents' nationality (column 5–6). The dependent variables are the INVALSI test scores in reading and maths. All test scores are standardised to have a mean of 200 and a standard deviation of 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include the average quality of old peers, gender-by-subject fixed effects and cohort-by-subject fixed effects. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.001$.

Different columns include different peer quality measures as treatments. In columns (1) and (4), we only include (new) peers' average ability; in columns (2) and (5), we only include only the proportion of new peers at the top and the bottom 5 percent of the ability distribution, while in columns (3) and (6), we include all the three treatments together.

Estimates results reported in columns (1)-(3) of Table 4 indicate that, once we only focus on new peers, the positive impact of average peer quality disappears: the coefficients of mean score of new peers in columns (1) and (3) shrink in magnitude (compared to the results presented in Table 3) and become insignificant. Similarly, column (2) shows that the effect of the top 5% peers is positive, but small, and not statistically different from zero. Only the effect of the bottom 5% peers remains considerable and significantly negative after focusing on new peers only. This result also applies when we do not control for the quality of old peers (columns 4–6). In particular, the coefficient of 7.5 in column (3) implies that a 10 percentage point increase in the share

This reassures us that there is no correlation between the proportion of high-low achievers in the class among old and new peers.

Table 7
Impact of peer quality on 8th grade outcomes of immigrant and native students by the immigrant status of peers.

	(1)	(2)	(3)	(4)	(5)	(6)
	Immigrants			Natives		
New peers' average score (natives)	0.017*		0.023*	0.036***		0.040***
	(0.010)		(0.013)	(0.005)		(0.006)
New peers' average score (immigrants)	0.022***		0.008	0.001		0.002
	(0.007)		(0.008)	(0.003)		(0.004)
Share of top 5% (new peers — natives)		0.802	1.736		1.552**	3.829***
		(1.408)	(1.760)		(0.638)	(0.798)
Share of top 5% (new peers — immigrants)		0.756	2.890		−0.095	0.339
		(1.141)	(2.054)		(0.528)	(1.060)
Share of bottom 5% (new peers — natives)		−2.479	−5.059**		−2.291***	−3.401***
		(1.592)	(2.016)		(0.750)	(0.974)
Share of bottom 5% (new peers — immigrants)		−1.659***	−2.884**		−0.289	0.600
		(0.642)	(1.241)		(0.292)	(0.572)
Observations	218,688	218,688	190,188	1,740,210	1,740,210	1,566,979

Notes: The table shows WS estimates of the effect of the average quality of new peers, and the proportion of top and bottom new peers on students' 8th grade INVALSI test scores by immigrant status. Columns (1)–(3) refer to immigrant students, while Columns (4)–(6) to native students. The dependent variables are the INVALSI test scores in reading and maths. All test scores are standardised to have a mean of 200 and a standard deviation of 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. In columns (3) and (6), the average scores of peers are computed, leaving out the scores of top and bottom achievers in the class. Peer measures are reported separately for immigrant and native children. Peer measures are developed using the same ability distribution for immigrant and native students. All specifications include the average quality of old peers, gender-by-subject fixed effects, and cohort-by-subject fixed effects. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

Table 8
Identification tests: impact of peers' ability on students ability in grade 5 (col 1–3) and on students ability in grade 8 controlling for grade 5 test scores (col. 4–6)

Dep variable:	(1)	(2)	(3)	(4)	(5)	(6)
	Test scores at grade 5			Test scores at grade 8		
New peers' average score	−0.002		−0.004	0.006		−0.007
	(0.007)		(0.008)	(0.010)		(0.012)
Share of top 5% (new peers)		−0.208	0.128		0.095	0.717
		(1.401)	(1.585)		(2.269)	(2.567)
Share of bottom 5% (new peers)		−0.287	−0.546		−6.843***	−7.323***
		(1.350)	(1.462)		(2.474)	(2.568)
Own test score at grade 5	–	–	–	0.322***	0.322***	0.322***
				(0.006)	(0.006)	(0.006)
F-test for joint significance	–	*0.967*	*0.967*	–	–	–
Observations	218,688	218,688	218,688	218,688	218,688	218,688

Notes: The table shows WS estimated coefficients of the average quality of new peers, the proportion of top and bottom new peers on students' test scores in 5th grade (columns 1–3) and 8th grade standardised test scores (columns 4–6). "New peers" refer to students in grade 8th in a given cohort that do not come from the same primary school. All test scores are standardised to have a mean of 200 and a standard deviation of 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include students' fixed effects, gender-by-subject fixed effects and cohort-by-subject fixed effects. Columns (4)–(6) additionally control for the student's own test score at 5th grade. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

of very low-achieving students in the class (which is like adding two bottom students in an average class of 20 students) reduces immigrant children's performance by 0.75 points; this roughly corresponds to roughly 2 percent of a standard deviation (0.75/40 = 0.0185). The implied effect size is thus fairly small, but in line with the findings of other studies. For example, Gibbons and Telhaj (2016) find that a one standard deviation increase in peer group prior achievement is associated with a 0.02 standard deviation increase in student achievement, while results in Lavy et al. (2012) indicate that an increase in the class share of bottom peers from 0% to 20% (minimum and maximum values in their sample) would lead to a reduction in students' test scores of around 9% of a standard deviation. In general, specifications in Table 4 focusing on new peers detect smaller effects than those estimated when focusing on all peers, which suggests that looking at new peers helps overcome the residual reflection problem. Together, our results indicate that immigrant students are most affected by the quality of peers at the very bottom of the ability distribution. This finding is consistent with a "bad-apple"-type model of peer effects in which a small number of very weak students adversely affects the learning of all others (Lavy et al., 2012; Lazear, 2001) and is in line with previous papers documenting a detrimental impact of very low-achieving pupils (see Burke and Sass (2013); Lavy et al. (2011), Hoxby and Weingarth (2005)). Academically weak students may generate negative externalities: they are likely to require more of the teacher's attention at the expense of other pupils (see Auestad (2018)) and to distract both the teacher and students from productive tasks, hence reducing the quality of the instruction time in the class.

To further understand how peers' ability shapes immigrant students' educational outcomes, the next section tests for the presence of heterogeneous effects across several characteristics, which will help identify possible channels at work.

6.2. Heterogeneous effects

In this section, we investigate whether the impact of class composition on students' achievements is heterogeneous along a number of relevant dimensions. In Table 5, we start by unpacking our results to allow peer effects to differ according to gender, socio-economic background, and initial level of ability. The first two columns of Table 5 show estimates by gender and indicate that there is no significant heterogeneity across this dimension. Both males and females are equally negatively affected by low-achieving students. We also find a weak indication that females benefit more than males from interactions with very bright peers: in fact, the coefficient of the share of top-achieving peers

Table 9
Identification tests: random allocation of teachers.

Teachers' characteristics	(1) High education	(2) High training	(3) Long tenure	(4) Permanent contract
Peers' average score	0.000	0.001	−0.002	−0.001
	(0.001)	(0.003)	(0.003)	(0.002)
Share of top 5%	0.151	−0.190	−0.237	0.097
	(0.251)	(0.379)	(0.379)	(0.263)
Share of bottom 5%	−0.095	0.579	−0.170	−0.072
	(0.216)	(0.370)	(0.333)	the (0.252)
Observations	2714	2714	2714	2714

Notes: The table reports coefficients from a within-classroom across subject-specific teachers' model. The dependent variables are defined as binary outcomes. Each of the four columns denotes a different teacher's characteristic: (1) an indicator for having an education level greater than a master's degree, (2) an indicator for having attended more training courses than the average, (3) an indicator for having an experience greater than five year, (4) an indicator for having an open-ended contract. See A.7 for variable definitions. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include class-fixed effects, the share of females per class, class-by-subject fixed effects and cohort-by-subject fixed effects. Regressions include only the classes (teachers) who were sampled to be part of the INVALSI nationally representative sample for the school cohorts 2016 and 2017. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.001$.

in the class is positive for girls and negative for males, despite them both not being statistically significantly different from zero. Columns (3) and (4) report the results by students' socio-economic status, as captured by the ESCS index. Students whose ESCS is above the median are defined as "high ESCS", while students with ESCS below the median are categorised as "low-ESCS". We find that test scores of immigrant children coming from a more advantaged backgrounds are positively affected by the share of top-achieving peers in the class, with the effect of the share of bottom peers proving to be insignificant. On the other hand, the negative impact of the share of bottom peers persists for low-ESCS immigrant students. In columns (5) and (6) we explore whether peer effects differ by the initial level of ability, measured using students' predetermined test scores in grade 5. In particular, we define students as having a high (low) initial ability level if the average 5th grade test score across subjects is above (below) the median. The results portend a sizeable and significant negative impact of bottom peers only for low-ability immigrant students, while we find that peer quality does not affect the test scores of high-ability immigrants. Overall, results in Table 5 suggest that the detrimental effects of low-achieving peers revealed in our baseline estimates for immigrant students (column 3 of Table 4) are driven primarily by the most vulnerable students, namely children coming from a more disadvantaged background both in terms of resources and school readiness.

We have so far explored the heterogeneity of peer effects along some of the dimensions most investigated in the existing literature. We now focus on some novel characteristics that better qualify the immigrant status of the students in our sample. In reality, immigrant students are a much more diverse than homogeneous population. Students with an immigrant background can differ widely in their country of origin, language traditions, and length of time spent in the host country, and these factors may impact the way peers influence their learning. In Table 6, we test if peer effects vary according to (1) whether the immigrant student is foreign-born (first generation) or born in Italy from foreign-born parents (second generation) (columns 1 and 2), (2) whether they usually speak Italian or another language at home (columns 3 and 4), (3) whether they have one or both parents with an immigrant background (columns 5 and 6).

We find no heterogeneous effects for first-versus second-generation immigrant children: both groups are adversely affected by high shares of very low-achieving peers and do not significantly benefit from higher-quality peers in the class. However, the negative effect of a high share of low-performing peers seems to be more pronounced for immigrant students that do not use the Italian language at home (column 4) and whose parents were both born abroad. Negative effects of being exposed to foreign language-speaking peers are also detected by Chuard et al. (2022), demonstrating a reduced probability of pursuing an academic path for ethnic minorities exposed to a higher share of foreign language-speaking peers in the class; particularly if peers speak the same foreign language.

Overall, these results, reinforce our previous finding: a large proportion of "badly performing" peers in the class are detrimental to learning, especially for most vulnerable and less integrated immigrant children. We interpret this finding in the light of the fact that a higher proportion of low-achieving students in the class may lead to a deterioration of teachers' pedagogical practices and to the relationships between teachers and students, and may increase the level of classroom disruptions (see, for example, Lavy et al. (2011)). Vulnerable immigrant students are especially harmed by this because they are likely to be less able to cope with difficult learning environments.

6.3. Discussion

Overall, our estimates indicate that immigrant students' school performance is undermined by the presence of high shares of very low-achieving peers in the class. In contrast, the average peers' quality and the proportion of very bright peers do not affect their educational outcomes. While our paper focuses on the school performance of immigrant students, in order to provide policy implications regarding class formation, it is important to consider the effect of peers' quality on native students, too. In this paragraph, we first replicate our baseline analysis in Table 4 on the sample of native students (see Table A.4 in the Appendix), and then in Table 7 we investigate whether immigrant and native students have different reference groups (Fordham and Ogbu, 1986) by defining our peer measures separately for immigrant and native students. This distinction sheds light upon potential differences in the impact that immigrant and native students' academic quality has on their peers. We use the same ability distribution for immigrant and native students when defining students in the top and bottom 5%. In Table A.5 in the Appendix, as a sensitivity check, we show that our results still apply if we use separate ability distributions for immigrant and native students.[11] Results in Table A.4 show that native students are not negatively affected by the share of bottom peers, but they benefit from a higher average quality of peers in the class (see columns 1, 4, and 6). Moreover, column (5) outlines that the proportion of bright peers positively impacts native students' achievement when the quality of old peers is not controlled for. However, this effect becomes statistically insignificant when all the treatments are jointly analysed in the same regression (column 6). Overall, these findings indicate that the quality of peers matters for both native and immigrant students, but they are affected in different ways: immigrants are detrimentally influenced by the share of bottom peers at the extreme of the ability distribution, while natives are mostly affected by the average quality of peers in the class. This evidence may be explained by the fact that, on average, immigrant children start secondary school with lower initial test scores and would benefit more from teachers' time compared to natives; for that reason, they appear to be more sensitive to potentially disruptive low-achieving peers.

Estimates in Table 7 provide further evidence in line with those in previous tables, and they improve our understanding of the contribution of immigrant and native children to the observed effects. We

[11] This may be relevant considering that, as discussed in Section 5, the distributions of test scores of immigrant and native children differ considerably: in the proportion of top peers there is an under-representation of immigrant children, while they are more heavily concentrated in the lower tail of the ability distribution.

Table 10
Robustness of results to alternative samples.

	(1) Include Southern regions	(2) Schools high share immigrants	(3) Provinces high share new peers	(4) Classes high share new peers	(5) Classes low share new peers
New peers' average score	−0.018	−0.006	−0.004	0.023	−0.020
	(0.011)	(0.015)	(0.016)	(0.022)	(0.015)
Share of top 5% (new peers)	0.523	−0.099	−1.362	−2.194	2.931
	(2.392)	(3.105)	(3.305)	(3.686)	(3.679)
Share of bottom 5% (new peers)	−6.830***	−8.447***	−9.287***	−6.864*	−7.116*
	(2.458)	(2.952)	(3.282)	(3.677)	(3.734)
Observations	246,808	162,164	145,204	109,356	109,332

Notes: The table shows the WS estimated coefficients of the average quality of new peers, the proportion of top and bottom new peers on students' 8th grade standardised test scores. "New peers" refer to students in grade 8th in a given cohort that do not come from the same primary school. All test scores are standardised to have a mean equal to 200 and a standard deviation equal to 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include students' fixed effects, gender-by-subject fixed effects and cohort-by-subject fixed effects. In columns (2) and (3), we define schools and provinces as having a high share of migrants if the share of migrants is higher than the median. In columns (4) and (5), we define classes as having a high (low) share of new peers if the share of new peers is higher (lower) than the median. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

Table 11
Robustness of results to alternative empirical approaches: school fixed effects.

	(1)	(2)	(3)	(4)	(5)	(6)
New peers' average score	0.085***		0.071***	0.076***		0.061***
	(0.007)		(0.010)	(0.007)		(0.009)
Share of top 5% (new peers)		6.999***	−1.014		6.267***	−0.693
		(1.404)	(1.679)		(1.383)	(1.656)
Share of bottom 5% (new peers)		−13.822***	−7.258***		−13.041***	−7.342***
		(1.384)	(1.543)		(1.366)	(1.529)
Demographic controls	*no*	*no*	*no*	*yes*	*yes*	*yes*
School fixed effects	*yes*	*yes*	*yes*	*yes*	*yes*	*yes*
Observations	218,688	218,688	218,688	217,338	217,338	217,338

Notes: The table reports the regressions' coefficients of the average peer quality (columns 1–2), the proportion of top and bottom peers (columns 3–4) and average peer quality, and the proportion of top and bottom peers (columns 5–6) on immigrant students' 8th grade standardised test scores. The dependent variables are the INVALSI 8th grade test scores in reading and maths. All test scores are standardised to have a mean of 200 and a standard deviation of 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include controls for school fixed effects, gender-by-subject fixed effects, and cohort-by-subject fixed effects. Columns (2),(4), and (6) additionally control for cohort fixed effect, gender, immigrant status, and quarters of birth. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

Table 12
Impact of peer quality on 8th grade outcomes of immigrant students — Different thresholds to define the groups of top and bottom peers.

Thresholds	(1) 1%	(2) **5%**	(3) 10%	(4) 20%
Mean score new peers	−0.000	**−0.008**	−0.006	−0.005
	(0.011)	**(0.012)**	(0.014)	(0.016)
Share of top new peers	3.042	**0.758**	0.808	0.362
	(5.274)	**(2.61)**	(2.053)	(1.737)
Share of bottom new peers	−9.792*	**−7.409*** **	−3.256*	−2.032
	(5.851)	**(2.626)**	(1.971)	(1.661)
Observations	218,688	218,688	218,688	218,688

Notes: The table shows WS estimates of the effect of the average quality of new peers, and the proportion of top and bottom new peers on immigrant students' 8th grade standardised test scores by different definitions of treatments. The dependent variables are the INVALSI test scores in reading and maths. All test scores are standardised to have a mean of 200 and a standard deviation of 40 for each subject and cohort. Columns (1)–(4) differ according to the thresholds used to define the top and bottom peers, which are 1% in column (1), 5% in column (2), 10% in column (3), 20% in column (4). All specifications include old peers' average quality, gender-by-subject fixed effects and cohort-by-subject fixed effects. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.001$.

define all peer measures separately for immigrant and native students. In this case, since there are, on average, few migrants per class, the average score of migrants can be greatly influenced by a bottom (or top) achiever. Therefore, in columns (3) and (6), when we include all treatments together, we calculate the average scores leaving the bottom and top achievers out of the computation (see also Lavy et al. (2012)).

Our results suggest that while immigrant students are equally affected by their native and immigrant peers, native students are only impacted by their native peers. In particular, immigrant students benefit from the average ability of both immigrant and native students (column 1), and are negatively affected by the proportion of low-achieving immigrant peers (column 2). When including all treatments together, immigrants' test scores are still affected by the share of low-achieving peers (both immigrant and native) while also being positively influenced by the average ability of their native peers (column 3). In particular, the size of the coefficient in column 3 implies that increasing native peers' average quality by one standard deviation (about 40 points) translates into an increase in migrant students' performance of about 2.3 percent of a standard deviation. When examining native children (see columns 4–6), we find that the average peer ability of native students plays an important role in native students' outcomes: estimates are positive and significant both when analysed separately and together in the same specification. Moreover, we find that the proportion of high-achieving (and low-achieving) native peers has a positive (negative) and significant effect, both when excluding the average peer ability (column 5) and when jointly considering all treatments (column 6). Interestingly, in all specifications, immigrant students' average ability, as well as the share of high- or low-achieving immigrants, do not affect native students' achievements. We interpret these results as evidence of homophily, particularly for native students. This finding is

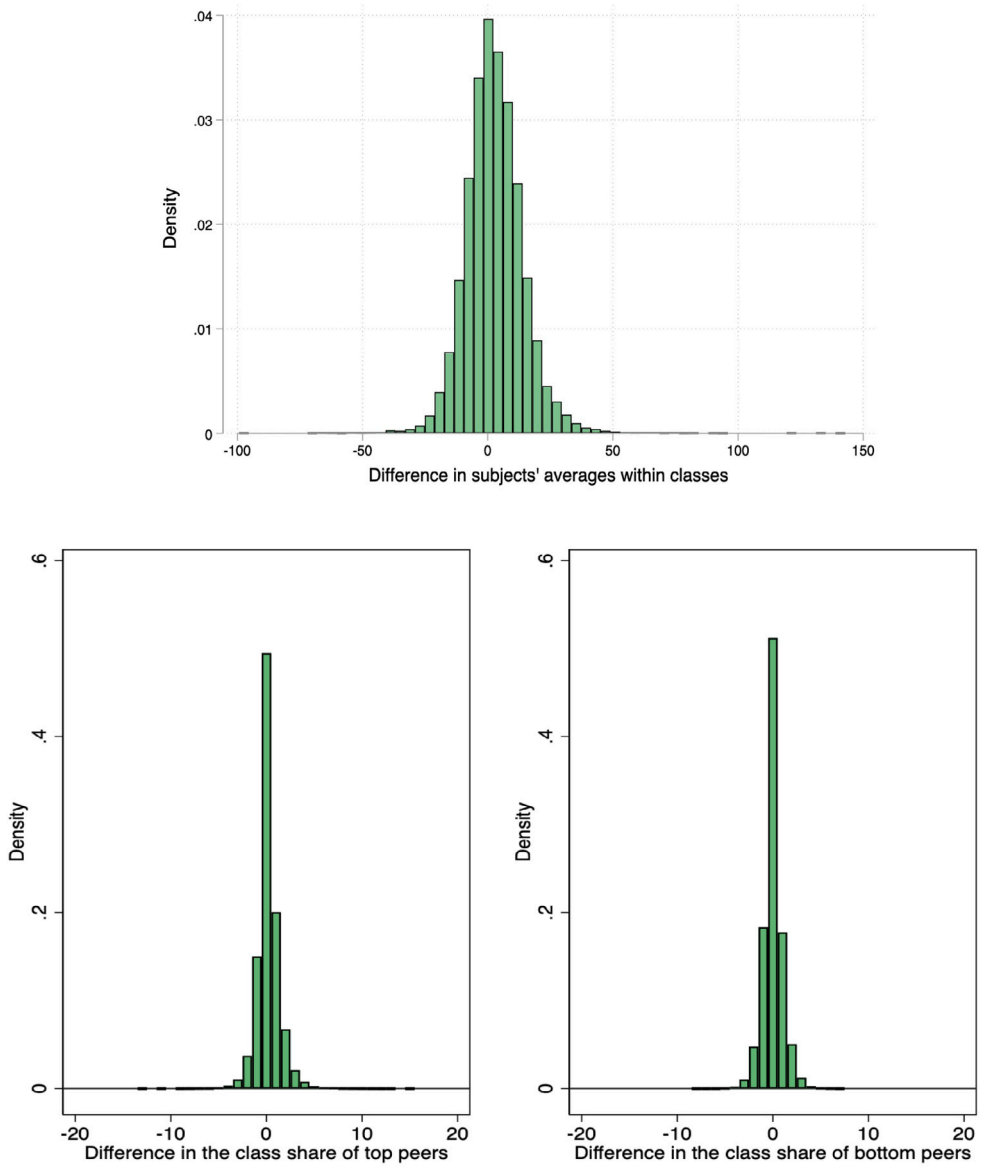

Fig. A.1. Distribution of differences in subjects averages, share of bottom and top peers within classes.
Note: Top Panel of the figure shows the distribution of the differences in subjects averages calculated using students' 5th grade test scores within 8th grade classes. Bottom Panel of the figure show the distribution of the within-classes variation in the shares of top (left bottom panel) and bottom (right bottom panel) peers across subjects. Top (bottom) peers are students that score in the top (bottom) 5% percent of the national ability distribution.

consistent with Figlio and Özek (2019), and Figlio et al. (2023) finding zero to positive effects of immigrants on the educational outcomes of native students and with the existing evidence on racial peer effects, where most studies suggest that peer effects are greater within the same racial group than between groups (Hoxby, 2000; Hanushek et al., 2003, 2009). This could be because many more peer interactions take place within a race than across races (Hanushek et al., 2003). As suggested by Fordham and Ogbu (1986) and Fryer and Levitt (2010), this evidence can be interpreted as students placing different emphasis on peers from diverse backgrounds. In our context, native students might value native peers more than immigrant ones (Fruehwirth, 2013). In other words, stronger within-group spillovers might reflect the fact that students respond more to peers with greater similarities to them. This result is relevant in light of the debate on immigrant students' impact on the educational performance of natives. Despite the concerns in many advanced countries that the increasing shares of immigrants may be detrimental to native students' educational achievement, the empirical evidence of such spill over effects is scant; our analysis seems to confirm the lack of such an effect. In order to test if the effect of specific peers depends on the number of migrants in the class, we replicate the estimates in Table 7 by dividing our sample into two sub-groups according to whether the share of immigrant students in the class is above or below the second tercile. The results are reported in Table A.6 in the Appendix and indicate that the impact of immigrant and native bottom peers' ability on immigrant performance is significant only in classes with a high share of migrants. Results conform with findings in Schneeweis (2015) showing that immigrant students suffer in classes with a high share of migrants while no effect is detected on average for native students. The latter suggests that a potentially important mechanism for our results has to do with peer effects in the class forming along an ethnic dimension that hampers migrants' educational integration by reducing their social interactions and learning opportunities from their native counterparts.

Table A.1
Attrition and individual characteristics.

	(1) Migrants	(2) Natives
Female	−0.070***	−0.024***
	(0.003)	(0.001)
Mother's education — college degree	−0.002	−0.004***
	(0.005)	(0.001)
Father's education — college degree	−0.020***	−0.009***
	(0.004)	(0.001)
Pre-primary school (0–2) attendance	−0.009***	0.003***
	(0.003)	(0.001)
Quarter of birth — 2	0.004	−0.002***
	(0.004)	(0.001)
Quarter of birth — 3	−0.001	−0.002***
	(0.004)	(0.001)
Quarter of birth — 4	−0.003	−0.005***
	(0.004)	(0.001)
Irregular career	0.043***	0.150***
	(0.004)	(0.005)
Maths standardised score — grade 5	−0.001***	−0.000***
	(0.000)	(0.000)
Language standardised score — grade 5	−0.001***	−0.001***
	(0.000)	(0.000)
Cohort fixed effects	*yes*	*yes*
Average dep. variable	*0.23*	*0.06*
Observations	106,755	729,684

Notes: Entries in the Table are the estimated coefficients of individual characteristics measured in grade 5 (gender, parents education, pre-primary school attendance, irregular career, quarter of birth, test scores in maths and language, and cohorts fixed effects) on attrition probability (a dummy variable equal to 1 if a student is observed in grade 5 but not in grade 8). Columns 1 and 2 report separate estimates for immigrants and natives, respectively.

7. Robustness checks and sensitivity

In this section, we present a number of tests on our identifying assumptions and several sensitivity checks to test the robustness of our findings.

First, we present a set of tests to support the causal interpretation of our results. Our approach makes it possible to estimate the causal effect of peers under the main identifying assumption that peers' subject-specific ability is not related to unobserved determinants of individual students' subject-specific skills. Therefore, the first threat to identification is the potential subject-specific sorting of peers across schools and classes, as explained in Section 4. In fact, despite our specification controls for students' average ability across subjects, there could be some residual correlation left between the subject-specific 5th grade within-student across-subjects variation and the variation of peers' quality across subjects in 8th grade. For example, if students are unobservedly more able in reading are systematically assigned to classes with a higher proportion of top peers in reading, our peer quality coefficients would be upward biases. In order to dispel this concern, in Table 8, we first perform a falsification test (or placebo) where we regress peers' ability in grade 8 on students' own ability in grade 5 (see columns 1–3) and then we augment our baseline specification including student's 5th grade test score among the control variables.

The estimates from the falsification test reported in columns (1)-(3) of Table 8 show that the coefficients of the three peer quality measures do not statistically significantly differ from zero, which reassures us that there is no significant correlation between the within-student across-subjects variation in prior achievements and the variation in peers' ability across subjects. For columns (3) and (4), the table also reports the p-values for the F-test testing the joint significance of peers' quality measures that are again strongly rejected. This evidence is also corroborated by the results in columns (4)-(6), which indicate that when we augment our baseline specification, including students' lagged test scores at grade 5, our coefficients of interest remain remarkably stable compared to baseline results reported in Table 4. The effects of the average peer quality and of the share of top 5% peers in the class remain insignificant, while the impact of the bottom 5% peers only marginally drops from 0.75 to 0.73. This suggests that conditional on student fixed effects, peer quality in one subject is balanced with respect to students' own test scores in that subject at the end of primary school (grade 5) (see also Lavy et al. (2012)). This evidence is not surprising in our context where, as explained in Section 3.1, lower secondary schools are comprehensive, and neither setting nor tracking practices are allowed, nor can schools or classes specialise in specific subjects.

Another potential threat to our identification strategy is the non-random allocation of teachers to classes. In this case, variation in students' ability across subjects may be correlated with teachers' characteristics. If, for example, students with higher ability in a specific subject were systematically paired with higher-quality teachers in the same subject, our estimates would inevitably be biased. In order to test for this possibility, we exploit the detailed data on teachers' characteristics provided by INVALSI for a nationally representative sample of teachers and we test whether subject-specific teachers' characteristics are systematically correlated with students' subject-specific ability measures in grade 5, prior to the assignment to lower secondary school classes.[12] We use four characteristics to proxy teachers' quality: (i) educational level, (ii) tenure (years of teaching experience) (iii) number of training courses attended by the teacher and (iv) type of contract relation, whether fixed- or indefinite-term. The use of these characteristics to proxy teachers' quality is well established in the educational literature (Chetty et al. (2011), Hanushek and Rivkin (2006), Figlio and Özek (2019) among others) and descriptive statistics for these variables are reported in Table A.8 of the Appendices. We then regress the four indicators of teachers' quality in each subject on subject-specific peer quality measures. The dependent variables are respectively defined as binary outcomes equal to one when teachers' education is higher than a master's degree, the number of training courses attended is higher than the median, the number of years spent in the school is more than 5, and when teachers have a permanent contract. Intuitively, if high-(low-) ability students in a specific subject are systematically assigned to better teachers, we should find that predetermined students' peer quality measures are strong predictors of teachers' quality. Table 9 reports the results. Encouragingly, they indicate that students' ability in primary school does not predict the type of teacher a student is assigned to in lower secondary school, consistent with the random allocation of teachers to classes. None of the measures of teachers' quality is, in fact, statistically significantly correlated to students' subject-specific ability.

Overall, the above-described tests point to the validity of our identification strategy and confirm that we can attach a causal interpretation to our regression results.

In the following tables, we present additional sensitivity checks to test the robustness of our specification. In Table 10, we test whether our results are robust to alternative samples. In our main analysis, we considered all students in public schools in Northern and Central Italy. In column (1) of Table 10, we run our baseline specification as in Table 4, now including the Southern regions in the estimation sample. In columns (2) and (3), we run our baseline analysis in a sub-sample of schools (column 2) and provinces (column 3) with a higher share of immigrants compared to the median. The estimates reported in the table suggest that our results are robust to the inclusion of Southern regions and indicate that the effect of bottom peers on immigrant students is greater in schools and provinces where most immigrants are concentrated. This result is in line with findings in Table A.6, which highlighted that spill overs are greater in contexts where there

[12] Data on teachers' characteristics are available only for the school years 2016 and 2017; thus, this analysis is only possible for these two school years.

Table A.2
Within and between students' variation of test scores and peers' measures.

	(1) Mean	(2) Overall s.d.	(3) Between s.d.	(4) Within s.d.
Panel A: Immigrant students				
8th grade scores	193.28	36.07	32.35	15.95
5th grade scores	194.90	39.54	35.47	17.47
Peers mean score (8th grade)	206.49	29.88	29.11	6.72
Share of top 5% (8th grade)	0.03	0.05	0.05	0.03
Share of bottom 5% (8th grade)	0.03	0.05	0.04	0.03
N	218,688			
n	109,344			
T	2			
Panel B: Native students				
8th grade scores	210.75	36.90	33.53	15.39
5th grade scores	213.87	38.82	35.08	16.62
Peers mean score (8th grade)	208.42	31.13	30.42	6.60
Share of top 5% (8th grade)	0.03	0.06	0.05	0.03
Share of bottom 5% (8th grade)	0.03	0.05	0.04	0.03
N	1,740,210			
n	870,105			
T	2			

Notes: The table shows means and overall, between- and within-standard deviations for 8th and 5th grade test scores and peers' quality measures (peers mean score, share top and share bottom) for immigrant (Panel A) and native (Panel B) students. We restrict the sample to students that have INVALSI test scores not missing both in 5th and 8th grade. All test scores are standardised to have a mean of 200 and a standard deviation of 40 for each subject and cohort. For the definition of the variables see Table A.7 in the Appendices.

Table A.3
Impact of new peers' quality on 8th grade outcomes of immigrant students — control for old peers' quality.

	(1) Baseline	(2) + top & bot old	(3) Baseline	(4) + top & bot old	(5) Baseline	(6) + top & bot old
New peers' average score	0.005	0.009			−0.008	−0.003
	(0.011)	(0.011)			(0.012)	(0.013)
Share of top 5% (new peers)			0.028	0.394	0.758	0.719
			(2.303)	(2.289)	(2.610)	(2.603)
Share of bottom 5% (new peers)			−6.936***	−7.084**	−7.499***	−7.333***
			(2.526)	(2.516)	(2.626)	(2.618)
Observations	218,688	218,688	218,688	218,688	218,688	218,688

Notes: The table reports WS estimated coefficients of the average quality of new peers, the proportion of top and bottom new peers on students' 8th grade INVALSI test scores of immigrant students. Columns (1), (3) and (5) show our baseline specification (Table 4) and in columns (2),(4) and (6) the same specification is augmented with controls for the quality of top and bottom old peers. "New peers" refer to students in grade 8 in a given cohort that do not come from the same primary school. All test scores are standardised to have a mean equal to 200 and a standard deviation equal to 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include student fixed effects, as well as gender-by-subject fixed effects and cohort-by-subject fixed effects. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

Table A.4
Impact of new peers' quality on 8th grade outcomes of native students.

	(1)	(2)	(3)	(4)	(5)	(6)
New peers' average score	0.011**		0.009	0.040***		0.038***
	(0.005)		(0.006)	(0.005)		(0.006)
Share of top 5% (new peers)		1.681	0.793		4.695***	1.046
		(1.027)	(1.150)		(1.052)	(1.171)
Share of bottom 5% (new peers)		−0.894	−0.205		−2.687**	0.142
		(1.200)	(1.268)		(1.233)	(1.296)
Control for old peers' quality	*yes*	*yes*	*yes*	*no*	*no*	*no*
Observations	1,740,210	1,740,210	1,740,210	1,740,210	1,740,210	1,740,210

Notes: The table shows the estimated coefficients of the average quality of new peers and the proportion of top and bottom new peers on students' 8th grade standardised test scores. "New peers" refer to students in grade 8th in a given cohort that do not come from the same primary school. All test scores are standardised to have a mean equal to 200 and a standard deviation equal to 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. All specifications include students' fixed effects, gender-by-subject fixed effects and cohort-by-subject fixed effects. Columns (1)–(3) also control for the average grade 5th test scores of old peers. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

Table A.5
Impact of peer quality on 8th grade outcomes of immigrant and native students, separate ability distribution.

	(1)	(2)	(3)	(4)	(5)	(6)
	Immigrants			Natives		
New peers' average score (natives)	0.017*		0.017	0.036***		0.043***
	(0.010)		(0.013)	(0.005)		(0.006)
New peers' average score (immigrants)	0.022***		0.018***	0.001		0.002
	(0.007)		(0.008)	(0.003)		(0.004)
Share of top 5% (new peers — natives)		0.782	1.494		2.423***	4.035***
		(1.408)	(1.740)		(0.644)	(0.801)
Share of top 5% (new peers — immigrants)		0.779	1.560		−0.212	0.459
		(0.812)	(1.544)		(0.392)	(0.776)
Share of bottom 5% (new peers — natives)		−2.461	−4.480***		−1.556**	−2.978***
		(1.505)	(1.848)		(0.710)	(0.898)
Share of bottom 5% (new peers — immigrants)		−1.521*	−3.178**		−0.263	−0.446
		(0.812)	(1.496)		(0.378)	(0.724)
Observations	218,688	218,688	194,783	1,740,210	1,740,210	1,571,459

Notes: The table shows WS estimates of the effect of the average quality of new peers, and the proportion of top and bottom new peers on students' 8th grade INVALSI test scores by immigrant status. Columns (1)–(3) refer to immigrant students while Columns (4)–(6) to native students. The dependent variables are the INVALSI test scores in reading and maths. All test scores are standardised to have a mean equal to 200 and a standard deviation equal to 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. In columns (3) and (6), the average scores of peers are computed, leaving out the scores of top and bottom achievers in the class. Peer measures are reported separately for immigrant and native children. Peer measures are developed using the same ability distribution for immigrant and native students. All specifications include old peers' quality measures, gender-by-subject fixed effects and cohort-by-subject fixed effects. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

Table A.6
Impact of peer quality on 8th grade outcomes of immigrant students by the share of immigrants in the class.

Immigrant share in the class	(1) High	(2) Low	(3) High	(4) Low	(5) High	(6) Low
New peers' average score (natives)	0.015	0.025			0.023	0.020
	(0.012)	(0.019)			(0.016)	(0.023)
New peers' average score (immigrants)	0.026***	0.006			0.012	−0.004
	(0.008)	(0.012)			(0.010)	(0.015)
Share of top 5% (new peers — natives)			0.050	3.039	0.494	4.763
			(1.654)	(2.662)	(2.144)	(3.080)
Share of top 5% (new peers — immigrants)			0.674	0.871	2.881	5.101
			(1.3667)	(2.084)	(2.098)	(13.609)
Share of bottom 5% (new peers — natives)			−3.213*	0.340	−5.432**	−3.927
			(1.819)	(3.215)	(2.361)	(3.820)
Share of bottom 5% (new peers — immigrants)			−1.988***	−0.489	−2.740**	−11.887
			(0.751)	(1.177)	(1.272)	(9.200)
Observations	150,772	67,916	150,772	67,916	188,442	60,939

Notes: The table shows WS estimates of the effect of the average quality of new peers, proportion of top and bottom new peers on students' 8th grade INVALSI test scores by the share of immigrants in the class. We define a class to have a high (low) share of migrants if the share of migrants in the class is higher (lower) than the second tercile. The dependent variables are the INVALSI test scores in reading and maths. All test scores are standardised to have mean equal to 200 and standard deviation equal to 40 for each subject and cohort. The proportion of top (bottom) peers is defined as the class share of students at the top (bottom) 5 percent of the cohort-specific national distribution of 5th grade INVALSI test scores. In columns (3) and (6), the average scores of peers are computed, leaving out the scores of top and bottom achievers in the class. Peer measures are reported separately for immigrant and native children. All specifications include old peers' quality measures, gender-by-subject fixed effects and cohort-by-subject fixed effects. Robust standard errors (adjusted for clustering at the class level) are in parenthesis, $^{***}p < 0.01$, $^{**}< 0.05$, $^{*}p < 0.1$.

is a stronger presence of immigrant students. In columns (4) and (5), we investigate the heterogeneity in peer effects by the share of new peers in the class. Any concern about interdependence between old and new peers would be less important as the share of new peers in the classroom gets larger. Hence, we test how the coefficients of new peers change as one considers groups of a different share of new peers. We define a class as having a high (low) share of new peers if the share of new peers is above (below) the median. Our findings in columns (4) and (5) indicate that there is no significant difference in peer effects between classes with high and low shares of new peers. This result is in line with the evidence of old peers (students that were in the same class both in the 5th and 8th grade) representing a relatively small share of a student's peers. Moreover, it reassures us that the reflection issue is not a concern even in classes with a relatively smaller share of new peers.

Next, we test the robustness of our main results to an alternative empirical approach. Our empirical strategy, described in Section 4, exploits the random variation in peer quality across subjects within each student. This strategy makes it possible to absorb individual unobserved heterogeneity and accounts for the non-random sorting of students across classes. However, we are not able to capture ability spill overs across subjects that may arise if, for example, peers' reading ability affects students' performance in maths. In Table 11, we investigate if our main finding regarding the negative impact of very low-achieving peers still applies when using an alternative empirical strategy commonly used in several existing studies, and relying on within-school variation in students' ability across adjacent cohorts or across different classes. Table 11 presents the results from our within-school across cohorts strategy on students' performance for three specifications that progressively add peers' quality measures. Specifications in columns (4)-(6) additionally include indicators for gender, quarters of birth, immigrant status (first- or second-generation immigrant) and the ESCS index measured in 5th grade. In line with our main specification, we find that the share of very weak students worsen students' educational achievements, lending support to our results.

Table A.7
Definition of variables and sources.

Variable	Description	Source
Test scores		
5th grade Reading test score	WLE 200 Rasch Score in Reading test	INVALSI
5th grade Maths test score	WLE 200 Rasch Score in Maths test	INVALSI
8th grade Reading test score	WLE 200 Rasch Score in Reading test	INVALSI
8th grade Maths test score	WLE 200 Rasch Score in Maths test	INVALSI
Student's characteristics (individual level)		
Female	dummy = 1 if female	INVALSI
Immigrant I generation	dummy = 1 if child is immigrant II generation	INVALSI
Immigrant II generation	dummy = 1 if child is immigrant I generation	INVALSI
Age	child year of birth	INVALSI
II quarter month of birth	dummy = 1 if child is born in second quarter	INVALSI
III quarter month of birth	dummy = 1 if child is born in third quarter	INVALSI
Native parents	dummy = 1 if mother & father were born in Italy	INVALSI
Student ESCS	Student's economic social and cultural status	INVALSI
Student language		
mostly spoken at home	dummy = 1 if equal to Italian	INVALSI
Teachers' characteristics		
High education	dummy = 1 if educational level higher than master	INVALSI
High tenure	dummy = 1 if tenure in the school higher than 5 years	INVALSI
High training	dummy = 1 if training courses attended higher than the median	INVALSI
Permanent contract	dummy = 1 if permanent contract	INVALSI

Table A.8
Teachers descriptive statistics.

	(1) Mean	(2) Standard deviation	(3) Observations
Dependent variables			
Low education	0.923	0.267	2714
High education	0.077	0.267	2714
Low training	0.541	0.498	2714
High Training	0.459	0.498	2714
Low tenure	0.379	0.485	2714
High Tenure	0.621	0.485	2714
Fixed-term contract	0.128	0.334	2714
Permanent contract	0.872	0.334	2714

Notes: The table shows descriptive statistics for the dependent variable used in Table 9 on a nationally representative sample of teachers in the school years 2016 and 2017.

Finally, in Table 12 we test the sensitivity of our results to changes in the thresholds defining high(low)-achieving peers. In our baseline estimates, we define the proportion of high- and low-achieving peers as the proportion of top and bottom *5%* students in the cohort-specific national distribution. In principle, other thresholds might have been selected, potentially changing the reported effects of peers' quality.[13] In Table 12, we replicate our baseline specification, using alternative cut-offs to define peers at the top and bottom of the ability distribution. In columns (1)-(4), we report estimates defining the proportion of low(high)-achieving students as those in the (i) 1% (ii) 5% (our baseline) (iii) 10% (iv) 20% of the cohort-specific national distribution.

Estimates in column (1) show that when defining top (bottom) peers as those in the top (bottom) 1% of the distribution, coefficients increase in magnitude compared to those in column (2), our baseline estimates. The share of bottom peers has a larger negative effect and becomes significant for natives as well. The share of top peers has no significant effect, but the magnitude is again larger. On the other hand, columns (3) and (4) show that the more comprehensive the definition of top (bottom) peers, the more the effect fades away. Interestingly, the effect of very good and very bad peers in column (3) drops by more than 50 percent. We interpret this dramatic drop as suggesting that our baseline definitions of top and bottom peers are the most appropriate, capturing the most relevant group of peers at the extreme of the ability distribution.

[13] Despite being reasonable, the choice of the cut-off has also been made to ease the comparison with existing results in the literature (Lavy et al., 2012).

8. Conclusions

Recent years have witnessed growing interest in how immigrants' concentration in the class impacts on native children's outcomes. However, little research has been conducted into empirical support for policies that promote and foster immigrant students' learning. Surprisingly, no studies have yet explicitly explored how the ability composition of a class affects immigrant children's educational outcomes. In fact, it is still an open question as to which circumstances the class environment promotes immigrant students' academic performance. This is a particularly relevant issue considering that successful education is crucial for overcoming the disadvantages of immigrants in European societies and fostering their integration.

This article fills that gap by providing an empirical investigation into the effect of both immigrant and native classmates' academic ability on the educational achievement of immigrant students in Italy. In particular, by adopting an identification strategy that exploits within-student variation in ability across subjects, we explore the roles played by the peers' average academic ability and by the proportion of peers at the extreme tails of the national cohort-specific ability distribution.

Importantly, our results reveal that there are ability peer effects for both natives and immigrants. However, native and immigrant students are differently affected by the performance of their peers. While natives are mainly affected by the average quality of classmates, immigrant students are detrimentally influenced by the share of very low-achieving peers. Our analysis further demonstrates that the share of weak peers is especially detrimental to the most vulnerable and less integrated immigrant children (those who do not use the Italian language at home or whose parents were both born abroad). In addition, we show that immigrants and natives tend to have different reference groups, with immigrants affected by the performance of their native and immigrant peers, and natives only affected by the ability of their native peers. This finding is in line with the literature on racial peer effects emphasising that peer effects are greater intra-race than across races (Hoxby, 2000; Hanushek et al., 2003, 2009). Moreover, it seems to suggest that the widespread perception of immigrant students creating adverse peer effects on their native-born peers may not be empirically grounded.

As policymakers increasingly look for policies to help migrants integrate into host countries' educational systems, we show that taking class ability composition into account might be a promising and low-cost tool to help improve immigrant students' educational performance, especially for the disadvantaged ones. In fact, our results indicate that immigrant students, especially the most vulnerable ones, are particularly susceptible to the composition of the class and the exposure to

academically weak students. Due to their disadvantaged background and lack of home resources, immigrant children might be more dependent on the school context for their learning. One might argue that concentrating more vulnerable children within the same class might help teachers tailor the educational offer and identify the specific needs of struggling students. Instead, our findings suggest that in order to foster immigrant children's learning, it is important to avoid the concentration of very weak students in the same class, especially when these low-achieving peers are migrants. From a policy perspective, our results indicate that it is possible to achieve some gains in immigrant students' performance by reorganising peer groups and increasing the ability mix and heterogeneity in the class. Immigrant students appear to benefit from a more even distribution of low-ability and foreign-born students across schools, which encourages the enforcement of measures such as residential desegregation policies or transport subsidies which promote a more heterogeneous pool of students for class formation. Our findings are also critical in light of the current debate regarding the possible expansion of school choice (see Gibbons and Telhaj (2016)); this may lead to a higher degree of sorting across schools along the lines of prior ability, and thus compromising the school attainment of migrants.

Data availability

We will provide instructions on how to access the data, which we cannot directly share although they are accessible to all researchers upon request to the data provider (INVALSI).

Appendix

See Tables A.1–A.8 and Fig. A.1.

References

Ammermueller, A., Pischke, J.-S., 2009. Peer effects in European primary schools: Evidence from the progress in international reading literacy study. J. Labor Econ. 27 (3), 315–348.

Angrist, J.D., 2014. The perils of peer effects. Labour Econ. 30, 98–108.

Angrist, J.D., Lang, K., 2004. Does school integration generate peer effects? Evidence from Boston's Metco program. Amer. Econ. Rev. 94 (5), 1613–1634.

Auestad, M.L., 2018. The effect of low-achieving peers. Labour Econ. 55, 178–214.

Ballatore, R.M., Fort, M., Ichino, A., 2018. Tower of babel in the classroom: Immigrants and natives in Italian schools. J. Labor Econ. 36 (4), 885–921.

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. J. Public Econ. 104, 65–77.

Bertoni, M., Nisticò, R., 2023. Ordinal rank and the structure of ability peer effects. J. Public Econ. 217, 104797.

Bonin, H., 2017. The Potential Economic Benefits of Education of Migrants in the EU. IZA Research Reports 75, Institute of Labor Economics (IZA).

Burke, M.A., Sass, T.R., 2013. Classroom peer effects and student achievement. J. Labor Econ. 31 (1), 51–82.

Card, D., 2001. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. J. Labor Econ. 19 (1), 22–64.

Carlana, M., La Ferrara, E., Pinotti, P., 2022. Goals and gaps: Educational careers of immigrant children. Econometrica 90 (1), 1–29.

Carrell, S.E., Fullerton, R.L., West, J.E., 2009. Does your cohort matter? Measuring peer effects in college achievement. J. Labor Econ. 27 (3), 439–464.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., Yagan, D., 2011. How does your kindergarten classroom affect your earnings? Evidence from project STAR. Q. J. Econ. 126 (4), 1593–1660.

Chuard, C., Aerne, A., Balestra, S., Eugster, B., Hodler, R., 2022. Ethnic clustering in schools and early career outcomes. In: CEPR Working Paper, Vol. 17505.

Cohen-Goldner, S., Paserman, M.D., 2011. The dynamic impact of immigration on natives' labor market outcomes: Evidence from Israel. Eur. Econ. Rev. 55 (8), 1027–1045.

Contini, D., 2013. Immigrant background Peer effects in Italian schools. Soc. Sci. Res. 42 (4), 1122–1142.

Corazzini, L., Meschi, E., Pavese, C., 2021. Impact of early childcare on immigrant children's educational performance. Econ. Educ. Rev. 85, 102181.

Dee, T.S., 2007. Teachers and the gender gaps in student achievement. J. Human Resour. 42 (3), 528–554.

Duflo, E., Dupas, P., Kremer, M., 2011. Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. Amer. Econ. Rev. 101 (5), 1739–1774.

Dustmann, C., Frattini, T., Preston, I., 2013. The effect of immigration along the distribution of wages. Rev. Econom. Stud. 80 (1), 145–173.

Elsner, B., Isphording, I.E., 2017. A big fish in a small pond: Ability rank and human capital investment. J. Labor Econ. 35 (3), 787–828.

Eurydice, 2013. Key Data on Education. Technical Report, Indire reports.

Feld, J., Zölitz, U., 2017. Understanding peer effects: On the nature, estimation, and channels of peer effects. J. Labor Econ. 35 (2), 387–428.

Figlio, D.N., Giuliano, P., Marchingiglio, R., Özek, U., Sapienza, P., 2023. Diversity in schools: Immigrants and the educational performance of U.S.-born students. Rev. Econ. Stud. (04).

Figlio, D., Özek, U., 2019. Unwelcome guests? The effects of refugees on the educational outcomes of incumbent students. J. Labor Econ. 37 (4), 1061–1096.

Fordham, S., Ogbu, J.U., 1986. Black students' school success: Coping with the "burden of 'acting white'". Urban Rev. 18 (3), 176–206.

Frattini, T., Meschi, E., 2019. The effect of immigrant peers in vocational schools. Eur. Econ. Rev. 113, 1–22.

Frattini, T., Vigezzi, N., 2018. 2nd Migration Observatory Report: "Immigrant Integration in Europe and in Italy". Technical Report, Centro Studi Luca D'Agliano and Collegio Carlo Alberto.

Fruehwirth, J.C., 2013. Identifying peer achievement spillovers: Implications for desegregation and the achievement gap. Quant. Econ. 4 (1), 85–124.

Fryer, Jr., R.G., Levitt, S.D., 2010. An empirical analysis of the gender gap in mathematics. Am. Econ. J. Appl. Econ. 2 (2), 210–240.

Gibbons, S., Telhaj, S., 2016. Peer effects: Evidence from secondary school transition in England. Oxf. Bull. Econ. Stat. 78 (4), 548–575.

Gould, E.D., Lavy, V., Daniele Paserman, M., 2009. Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence. Econ. J. 119 (540), 1243–1269.

Hanushek, E.A., Kain, J.F., Markman, J.M., Rivkin, S.G., 2003. Does peer ability affect atudent achievement? J. Appl. Econometrics 18 (5), 527–544.

Hanushek, E.A., Kain, J.F., Rivkin, S.G., 2009. New evidence about brown v. Board of education: The complex effects of school racial composition on achievement. J. Labor Econ. 27 (3), 349–383.

Hanushek, E.A., Rivkin, S.G., 2006. Teacher quality. Handbook of the Economics of Education, Vol. 2. Elsevier, pp. 1051–1078.

Hoxby, C., 2000. Peer Effects in the Classroom: Learning from Gender and Race Variation. Technical Report, National Bureau of Economic Research.

Hoxby, C.M., Weingarth, G., 2005. Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects. Technical Report, Working paper.

Imberman, S.A., Kugler, A.D., Sacerdote, B.I., 2012. Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. Amer. Econ. Rev. 102 (5), 2048–2082.

INVALSI, 2015. Rilevazioni nazionali degli apprendimenti 2014-15: Rapporto Risultati.

Jensen, P., Rasmussen, A.W., 2011. The effect of immigrant concentration in schools on native and immigrant children's reading and math skills. Econ. Educ. Rev. 30 (6), 1503–1515.

Lavy, V., 2015. Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. Econ. J. 125 (588), F397–F424.

Lavy, V., Paserman, M.D., Schlosser, A., 2011. Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. Econ. J. 122 (559), 208–237.

Lavy, V., Silva, O., Weinhardt, F., 2012. The good, the bad, and the average: Evidence on ability peer effects in schools. J. Labor Econ. 30 (2), 367–414.

Lazear, E.P., 2001. Educational production. Q. J. Econ. 116 (3), 777–803.

Lucifora, C., Tonello, M., 2015. Cheating and social interactions. Evidence from a randomized experiment in a national evaluation program. J. Econ. Behav. Organ. 115, 45–66.

Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. Rev. Econ. Stud. 60 (3), 531–542.

Ministero Della Pubblica Istruzione, 2020. Gli Alunni con Cittadinanza non Italiana a.s. 2018/2019. Scuole Statali e non Statali. Anno scolastico 2018/2019.

Murphy, R., Weinhardt, F., 2020. Top of the class: The importance of ordinal rank. Rev. Econom. Stud. 87 (6), 2777–2826.

Ohinata, A., Van Ours, J.C., 2013. How immigrant children affect the academic achievement of native dutch children. Econ. J. 123 (570), F308–F331.

Pagani, L., Comi, S., Origo, F., 2021. The effect of school rank on personality traits. J. Hum. Resour. 56 (4), 1187–1225.

Sacerdote, B., 2011. In: Hanushek, E.A., Machin, S., Woessmann, L. (Eds.), Chapter 4 - Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?. In: Handbook of the Economics of Education, vol. 3, Elsevier, pp. 249–277.

Schneeweis, N., 2015. Immigrant concentration in schools: Consequences for mative and migrant students. Labour Econ. 35, 63–76.

Tonello, M., 2016. Peer effects of non-native students on natives' educational outcomes: Mechanisms and evidence. Empir. Econ. 51 (1), 383–414.