



UNIVERSITA' DEGLI STUDI DI PADOVA

**DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.
FANNO"**

**CORSO DI LAUREA MAGISTRALE IN
APPLIED ECONOMICS**

TESI DI LAUREA

**"THE HIDDEN LEGACY OF PEER EFFECTS:
IMMIGRANT SCHOOLMATES AND
LONG-TERM EDUCATIONAL OUTCOMES IN ITALY"**

RELATORE:

CH.MO PROF. Giorgio Brunello

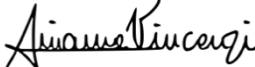
LAUREANDO/A: Arianna Vincenzi

MATRICOLA N. 2125727

ANNO ACCADEMICO 2024-2025

Dichiaro di aver preso visione del “Regolamento antiplagio” approvato dal Consiglio del Dipartimento di Scienze Economiche e Aziendali e, consapevole delle conseguenze derivanti da dichiarazioni mendaci, dichiaro che il presente lavoro non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere. Dichiaro inoltre che tutte le fonti utilizzate per la realizzazione del presente lavoro, inclusi i materiali digitali, sono state correttamente citate nel corpo del testo e nella sezione ‘Riferimenti bibliografici’.

I hereby declare that I have read and understood the “Anti-plagiarism rules and regulations” approved by the Council of the Department of Economics and Management and I am aware of the consequences of making false statements. I declare that this piece of work has not been previously submitted – either fully or partially – for fulfilling the requirements of an academic degree, whether in Italy or abroad. Furthermore, I declare that the references used for this work – including the digital materials – have been appropriately cited and acknowledged in the text and in the section ‘References’.

Firma (signature) 

ABSTRACT

With the rise in immigrant concentration in all European countries, it becomes central to understand the implications of increased heterogeneity in the field of education, among others. Contemporaneous peer effects arising from immigrant exposure have been studied accurately in the last decades, but there is still a gap with respect to whether these effects last over time. We exploit data from INVALSI to retrieve information over one cohort of student in Italy from 5th to 13th grade and try to estimate the causal effect of immigrant concentration in 5th grade on test scores in 5th, 8th, and 13th grades. To address non-random sorting of students across schools and classrooms, we adopt fixed effects at province level and measure the school-level share of immigrants. When controlling for important individual characteristics and for the socio-economic background of children, we find evidence of persistent negative effects arising from early exposure to immigrant peers up that last up to the end of upper secondary education.

Table of contents

1	Introduction	1
2	Why do peer effects matter? A motivating framework	5
2.1	Human capital, education and economic prosperity	5
2.2	Managing diversity: distribution of migrants and school policies to foster their integration	10
2.3	The risk of school segregation	16
2.4	The role of primary education	20
3	Literature review	23
3.1	Understanding peer effects and the challenges of identification	23
3.2	Measuring the invisible: identifying peer effects in schools	27
3.3	Do immigrant peer effects last?	31
3.4	Evidence from Italy and research gaps	33
4	Data	37
4.1	Data sources	37
4.2	Building the dataset and defining the sample	39
4.3	Key variables and descriptive statistics	45
4.4	Immigrant children in Italy	47
4.5	The performance of students in Italian schools	51
4.6	Data limitations	58
5	Empirical strategy	63
5.1	Theoretical framework and challenges in identifying peer effects	63
5.2	Identification strategy and limitations	66
6	Results	71
6.1	Baseline results	71
6.2	Heterogeneous peer effects	89
6.3	Testing nonlinear effects	95

6.4	Exploring potential tipping points	96
7	Robustness and sensitivity check	99
7.1	Placebo regressions.....	99
7.2	Non-random sorting across schools.....	101
7.3	The issue of attrition	109
7.4	Testing different samples	112
7.5	Other educational outcomes.....	118
8	Discussion of results and conclusions	121
	References.....	127
	Appendix.....	135
	A.4.2 Observations in INVALSI dataset without SIDI code and number of observations lost during the merging process.....	135
	A.4.5. Relationship between test scores and share of first-generation immigrants and second-generation immigrants in 5 th grade.....	136
	A.6.1. Distribution of first-generation immigrant children by their age of arrival in Italy.	141

1 Introduction

Recent raises in migration flows have sparked interest in understanding the potential implications for the native population. Italy has faced a substantial increase in its population of non-natives in recent decades, which makes it the ideal setting to investigate the impact of increased ethnic heterogeneity. The effect of growing immigration on the school system has been at the center of a large part of recent research. Education is an important pillar of any society, and there is clear proof of its association with the accumulation of human capital and economic prosperity. As economists and social scientists, understanding the dynamics that affect the development of skills and knowledge allows for a different point of view on policies that regulate the education system. The increase in the numbers of non-native children enrolled in Italian schools enhance the research value of understanding the implications for the performance of native pupils. So far, research has focused mostly on contemporaneous effects, and little is still known about the long-term implications. To our knowledge, there is no research that specifically addresses long-term peer effects arising from immigrant concentration in Italy.

This paper aims at measuring the causal effect of immigrant concentration during primary education on school performance in primary and secondary education. While it has been found proof of a negative association between immigrant exposure and educational outcomes, there is no information as to how persistent these effects can be over time. We use of INVALSI data over one cohort of students, who attended 5th grade between 2014-2015, and retrieve information on 8th grade (2017-2018) and 13th grade (2022-2023). What makes this specific cohort of students particularly useful to the scope of this analysis is that for the first time INVALSI registers each student with the SIDI code, a unique identifier that makes it possible to track the same student over time. We were able to build a sample of students that we follow from elementary school to the end of upper secondary education. On average, immigrant pupils achieve lower scores than native students both in mathematics and reading at all grades under exam. Raw data show that native students who attend primary schools with higher shares of immigrant children tend to perform slightly worse than their peers in schools less exposed to immigrant concentration, suggesting that there might be indeed an ongoing negative correlation between immigrant concentration and natives' performance. However, as long as we don't address specific issues related to students' distribution this is just a spurious correlation. When it comes to peer effects, the greatest challenge is posed by the fact that children – both native and non-native ones – are not randomly allocated across schools and likely even across classrooms within the same school. School choice, especially in early education, is usually

driven by residential proximity; nevertheless, residential choice is often the result of housing costs and job opportunities. Following that, it is common to find that individuals who are similar in terms of socio-economic background live in the same areas, and their children cluster in the same schools. As a result, immigrant students and native pupils from less advantaged socio-economic backgrounds may be disproportionately represented in the same schools. In this context, a naïve model that does not account for selection may produce a downward-biased estimate of the effect of immigrant concentration on educational outcomes. Econometric tools that range from fixed effects to instrumental variable help to address the endogenous placement of students. To deal with this possibility, we exploit fixed effects at province level, the smallest geographical unit available in our dataset, assuming that there is limited degree of within-heterogeneity in terms of immigrant concentration and local economic context. There is a chance as well that students within the same school are not randomly assigned to classrooms. Administrators may decide to place students expected to be high achievers to high-quality teachers, to further enhance their performance. Else, they may choose to assign non-native pupils to more experienced teachers in order to support their integration into the school system. We compute our treatment – the share of immigrants in 5th grade – at the school level rather than at the class level to avoid this issue. Even though most peer interactions take place at classroom level, school-level immigrant concentration still is a valid measure of early exposure to non-native children in the educational setting. We deliberately sacrifice some level of detail to strengthen the robustness of our results. The outcomes of interest are test scores in mathematics and reading obtained in the INVALSI assessment in 5th, 8th, and 13th grades. The identification strategy is based on the assumption that unobserved variables that influence both the distribution of immigrants and the performance of native students are, to some extent, time invariant.

This research needs to address two challenges. Firstly, INVALSI data refer only to *test-takers*, that is student who attended school on the day of the assessments. To avoid any potential issue arising from poor coverage rates, we decided to drop from the sample of students all observations from the regions of Campania, Puglia, Sardegna and Sicilia, who had very low participation rates in the 2014-2015 INVALSI wave. The second concern is given by attrition. The analysis was conducted on the sample of students individually observed in all three grades of interest. During the process, we lost multiple observations from students who, for any reason, were not present in 8th grade and 13th grade dataset. These could be students who at some point in time were held back a grade, who were sick on the day of the test and missed class, who purposefully missed the tests because of poor school attainment or who dropped out of school.

If it was the case that these missing observations were systematically different from those in the final sample, our estimates could be biased. In particular, we would have observed an upward bias if we believe that students who fell out of the sample are the lower-achieving ones who could have been negatively affected by immigrant concentration to the point that, for instance, they abandoned school. Because of the lack of precise information on the reason why some students are not observed in 8th and 13th grade, there was no feasible solution we could adopt to specifically address this issue. Nevertheless, we conduct precise robustness checks to ensure minimum impact on our findings.

We find that a higher share of immigrants in 5th grade negatively affects performance in both mathematics and reading up to the end of upper secondary school. The effects are moderate but significant, and there is no evidence that the adverse effects of early immigrant exposure tend to fade over time. The analysis also yields proof that peer effects are heterogeneous; we find stronger negative effects for male native students and for native children from a more disadvantaged socio-economic background.

The rest of the paper is structured as follows. The next section motivates the economic interest for the subject of this analysis. Section 3 reviews the existing literature on peer effects. Section 4 describes data and their limitations. Section 5 and 6 respectively present the empirical framework and the main results. Section 7 reports a set of robustness and sensitivity checks. Section 8 discuss results and concludes.

2 Why do peer effects matter? A motivating framework

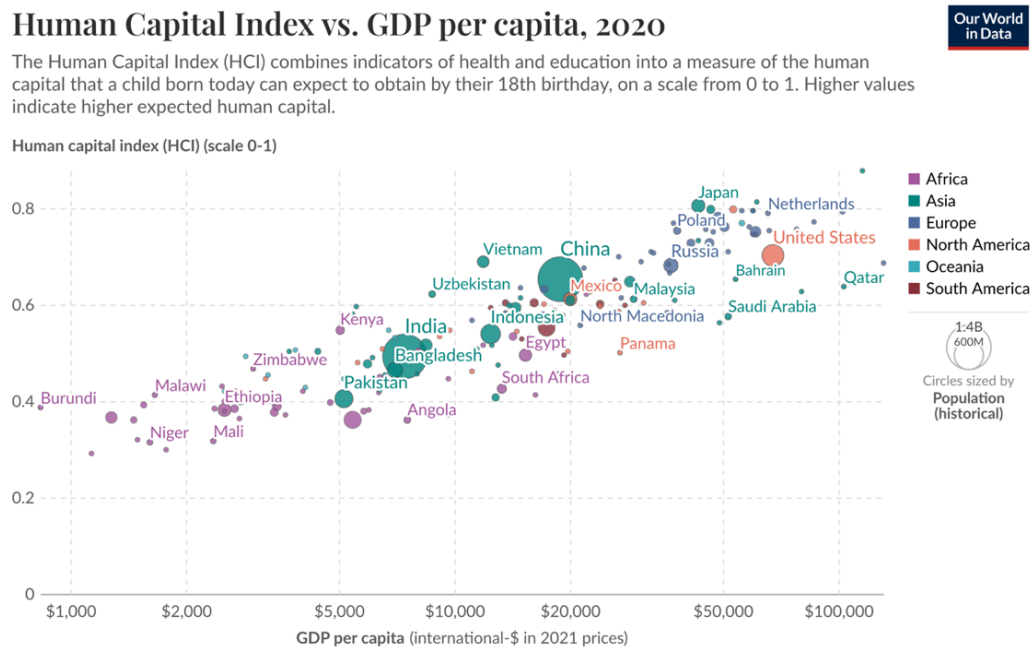
2.1 Human capital, education and economic prosperity

The term *human capital* became popular around 1960s primarily through the contributions of Jacob Mincer (1958) and Gary Becker (1962), who define it as the set of skills, knowledge, experience and other intangible resources that affect individual productivity and, by extension, economic prosperity. Among the various factors that contribute to the accumulation of human capital, both scholars emphasized the pivotal role of education. Another important determinant is on-the-job training, whose impact on productivity and aggregate economic performance is well-documented and definitely evident. However, understanding how formal education influence these outcomes is less straightforward. Becker (1962) proposes an interesting interpretation of schooling, suggesting that it is a form of training akin to that offered by firms, differing only in its timing relative to the production of goods. If schools are conceptualized as if they were a special type of firm, it turns out that schooling steepen the age-earning profile. In other words, it means that the earning growth from year to year is faster for individuals who acquire more education. When considering this, the connection between education and economic prosperity becomes more evident: schooling can enhance individual knowledge and skills, increasing productivity and thus their earnings, ultimately enabling higher level of consumption which directly contribute to broader economic welfare. Figure 2.1 illustrates the correlation between GDP per capita and the Human Capital Index¹ (HCI) in 2020 using cross-country data. GDP per capita is expressed in international dollars to facilitate meaningful comparisons across countries by adjusting for inflation and differences in cost of life. The positive association between human capital and economic prosperity is readily observable when looking at this figure. In particular, it can be observed that for small gaps of HCI, there are very

¹ The HCI was created in 2018 by the Human Capital Project as a measure of the level of human capital that a child born at the time of computation can expect to obtain by their 18th birthday, highlighting how current health and education shape the productivity of the next generation of workers. It underscores its pivotal role in realizing individual potential as productive members of a society. The index incorporates measures of child survival, stunting for children under age 5, adult survival rates, expected years of schooling and international test scores, and compare them to a benchmark, equal across countries, of complete education and full health. Formally, HCI is computed as $HCI = \frac{p}{p^*} * e^{\varphi(s_{NG}-s^*)} * e^{\gamma(Z_{NG}-Z^*)}$, where the first term captures future productivity due to child mortality, the second term reflects loss in future productivity due to children completing less than 14 years of school and the third term reflects the reduction in future worker productivity due to poor health (Kraay, 2018; World Bank, 2020).

large differences in terms of GDP per capita. Nevertheless, it is important to mention that this figure shows the simple association, and other factors that impact both GDP and HCI are omitted. For instance, political stability and infrastructures influence GDP as well as factors that compose the HCI but aren't accounted for, possibly confounding the relationship. There could also be overlooked reverse causality, as higher GDP per capita can lead to better investments in education and health.

Figure 2.1. Correlation between GDP per capita and Human capital index (HCI).



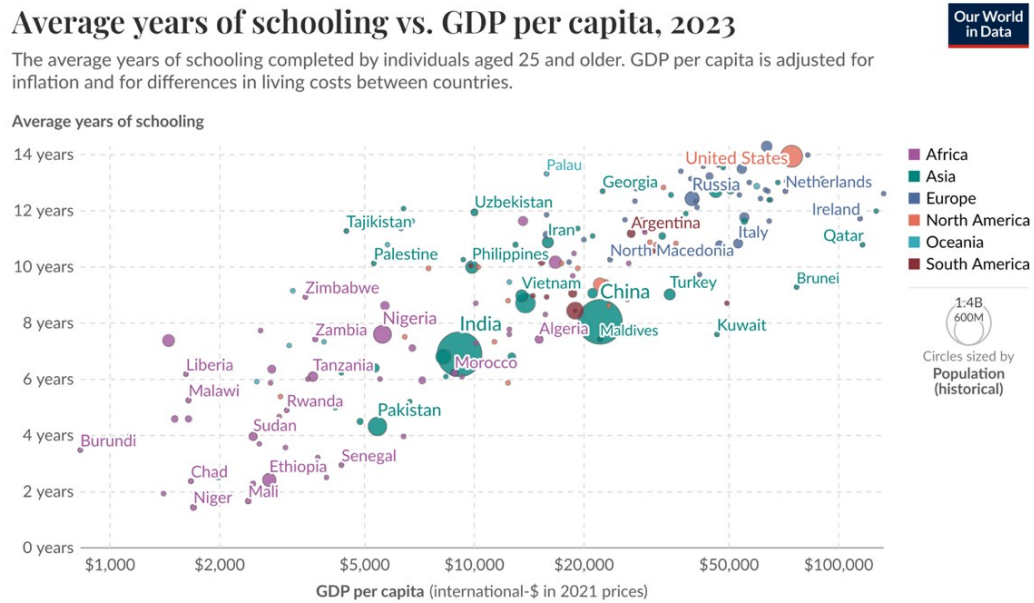
Note: HCI ranges between 0 and 1 in 2020. GDP per capita is expressed in international dollars to allow for meaningful comparison over time, to adjust for inflation and for differences in the cost of living between countries. Data source: World Bank (2025).

The role of education as a driver of human capital and growth has been widely researched in the last decades. E. A. Hanushek & Woessmann (2010) gather three mechanisms through which education can affect economic growth. First, according to neoclassical growth theories, education increases human capital which affects labor productivity, leading to a higher equilibrium level of output. Second, endogenous growth theories suggest that education increases the innovative capacity of an economy, and such novel knowledge of modern technologies and processes ultimately fosters growth. Third, education promotes the diffusion and transmission of knowledge that can be used to successfully implement technologies conceived by others, enhancing economic growth. While theory supports the relevant role of education in the process of economic growth, a clear measurement of what is education is required in order to practically understand its impact. Different possible metrics were proposed and tested over time, spacing from years of schooling to test scores, school quality, school

resources or degree attainment. Originally, the most common measure of education in economic and social science research was years of schooling. There is a vast literature (Barro, 1991, 1996; Barro & Lee, 1993; Mankiw & Weil D. A, 1992) that used cross-country growth regressions, modelling average annual growth in GDP per capita as a function of years of schooling and other variables, that support the positive association between the quantity of schooling and economic growth. The core assumption in these frameworks is that additional years of education improve individual knowledge and skills and directly contribute to human capital accumulation. Figure 2.2 illustrates the positive relation between average years of schooling and GDP per capita in 2023 using cross-country data. Countries in which children complete on average the full 14 years of education have the highest GDP per capita. Again, it must be noted that the simple association between these two measures omits important aspects, for instance that higher GDP may facilitate greater investment into education, allowing children to complete school. In any case, it is evident that there must be – at some degree – a positive relation between the number of years that a child stays in school and their ability and skills. It may be discussed up to which point school can contribute to human capital accumulation, or, in other words, if there exists a threshold after which no significant increase occurs. Interestingly, Sala-i-Martin et al. (2004) finds that primary schooling is the second most robust influence factor for economic growth. They test 67 explanatory variables – retrieved from existing literature – in growth regressions using a sample of 88 countries with data spanning between 1960-1996, and they find that the enrollment in primary schooling is the second most influencing factor in terms of economic growth². The enrollment in higher education – retrieved from Barro & Lee (1993) – is found at the 25th place, possibly suggesting that primary schooling has a stronger influence on the development of individual cognitive skills. This finding is particularly relevant to the scope of this study, as it supports the broader evidence that primary school experience plays a pivotal role in shaping children’s abilities and human capital creation, with long-term implications for individual outcomes.

² The first most influencing factor is a dummy for East Asian country, which reflects the extreme exceptional growth performance of East Asian countries between 1960 and 1996 (Sala-i-Martin et al., 2004).

Figure 2.2. Association between average years of schooling and GDP per capita in 2023.



Note: GDP per capita is expressed in international dollars to allow for meaningful comparison over time, to adjust for inflation and for differences in the cost of living between countries. Data source: UNDP, Human Development Report (2025). Retrieved from: World Bank (2025).

Even if the positive association between quantity of schooling and economic prosperity has been found in multiple contexts, main doubts remain as to whether it is the appropriate measure of education. Affirming that education can be measured exclusively by years of schooling is based on the implicit assumption that one additional year of school affects all individuals the same, regardless of the quality of the education system and of the intrinsic ability that individuals possess, and that all skills and human capital come from formal schooling. Yet there is extensive research showing that other factors such as health, nutrition, family and *peers* affect the development of cognitive skills (E. A. Hanushek & Woessmann, 2010). More recent research has turned a more comprehensive measure of quality of schooling as a parameter of education. E. A. Hanushek & Kimko (2000) find that quality of schooling can be measured in two ways, either through schooling inputs – expenditure, teacher-pupil ratio, teacher salaries – or through direct measures of knowledge and skills of individuals. We are particularly interested in the latter, for which they use data on test scores in mathematics and science. When regressing on initial income per capita, they find that quantity of schooling loses significance as soon as school inputs or quality measures are added to the model, thus supporting the importance of quality of education, rather than quantity, in explaining the accumulation of human capital and the ultimate impact on economic growth. E. A. Hanushek & Woessmann (2012) find similar results using again math and science test scores in a sample of 50 countries. Barro (2001),

E. A. Hanushek & Woessmann (2008), Heller-Sahlgren & Jordahl Henrik (2023) – among others – yield similar results on the impact of schooling quality measured through test scores on growth rate and economic prosperity. Today, there is a general consensus in literature that quality of schooling is a more effective measure of education, as it can capture the actual inputs to schools that shape individual cognitive skills and school outcomes.

Overall, there is large evidence that education plays a crucial role in the accumulation of human capital and, consequently, in fostering economic growth. This explains why both economists and social scientists are deeply interested in investigating the dynamics of education, as well as the underlying mechanisms that influence the acquisition of knowledge and skills, with the aim of informing the development of policies designed to optimize the benefits arising from schooling.

Moving backwards, it becomes necessary to discover what elements influence education and school achievement, so that proper intervention and investment can be done to maximize welfare gains. There are multiple factors that determine schooling quality and the development of cognitive skills: class size (Angrist & Lavy, 1999; Etim et al., 2020; Krueger, 2003), teachers' quality and salary (Darling-Hammond, 2000), student-teacher ratio (Del Boca et al., 2018; Giambona & Porcu, 2018), socio-economic background, family and *peers*, are just a few dimensions that have been examined over time. As already mentioned, peers are a key element to this research. The intrinsic importance of peers in school will be untangled in the next section; what matters at this point is that if peers affect schooling quality and students' performance, which indirectly affect the economic well-being, then policy makers and economists shall be interested in understanding which are the mechanisms that lie behind peer influences and if there exist optimal policies that can take advantage of these effects through manipulation of peer composition. The key takeaway at the end of this section is that peer characteristics and school composition affect students' achievement and performance, contributing to human capital accumulation. Among other aspects, the role of ethnicity and immigrant concentration has become an interesting niche in peer effects studies in the last two decades.

2.2 Managing diversity: distribution of migrants and school policies to foster their integration

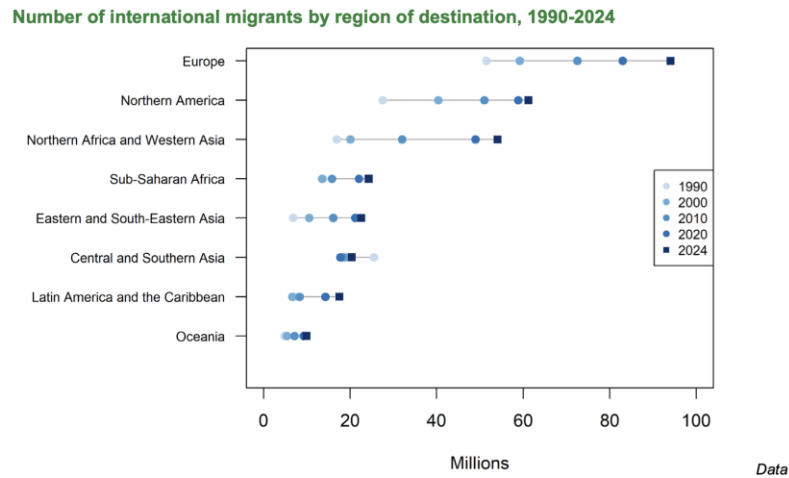
Under the assumption that peer characteristics affect school performance, it follows that policies targeted at shaping school and classroom composition can directly affect students' achievement and their development of cognitive skills. The way in which individual characteristics – such as ability or, in this case, ethnicity – influence educational outcomes deserves substantial consideration in the development of school policies. Besides that, education systems should guarantee equal opportunities to all children, in proportion to each one's own attributes and possibilities. Ideally, the same level of investment in education should be ensured for all students within the same system, so that all children have access to the same resources and receive the same amount of attention from their teachers. This has been an especially challenging task over the past several years, due to the large migration waves that hit European countries, leading to a sharp increase in the number of non-native students with different needs and abilities, in terms of the host-country language. There exist two main schools of thought regarding the increased presence of migrant pupils in schools. The first one emphasizes the potential challenges, arguing that a higher share of non-native children can create language barriers³, strain resources and complicate teaching, potentially lowering overall student performance. This perspective highlights the need for targeted support and integration policies to mitigate any possible detrimental effects. The second school of thought focuses on the benefits of diversity and increased heterogeneity, suggesting it can enrich the environment by promoting social cohesion and cultural exchange. A well-managed diversity can prepare all students for an increasingly globalized world. Both perspectives agree on the importance of targeted policy design to shape and control the impact of immigrant student composition on educational outcomes.

According to the United Nations, Europe has historically been a major migration-receiving country, and since 1990, it has maintained a gap of around 10 million compared to second-places Northern America (Figure 2.3). Today, more than 10% of the population of the European

³ Dustmann & Fabbri (2003) stress the importance of knowledge of host country language, focusing on its key role in the development of host country human capital and economy.

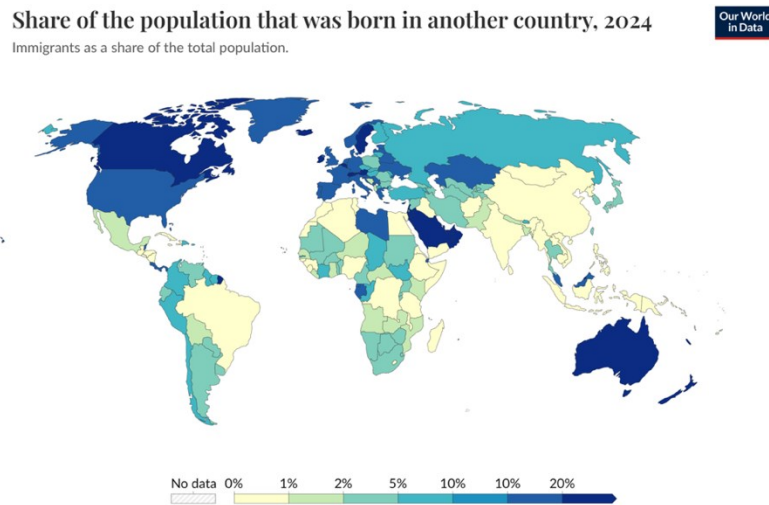
Union countries, United States, Canada and Australia is composed by *international migrants*⁴ (Figure 2.4).

Figure 2.3. Number of international migrants by region of destination (1990-2024)



Note: Regions ordered by number of international migrants in 2024. Data source: United Nations (2025).

Figure 2.4. Immigrants as share of total population. Data source: United Nations Department of Economic and Social Affairs

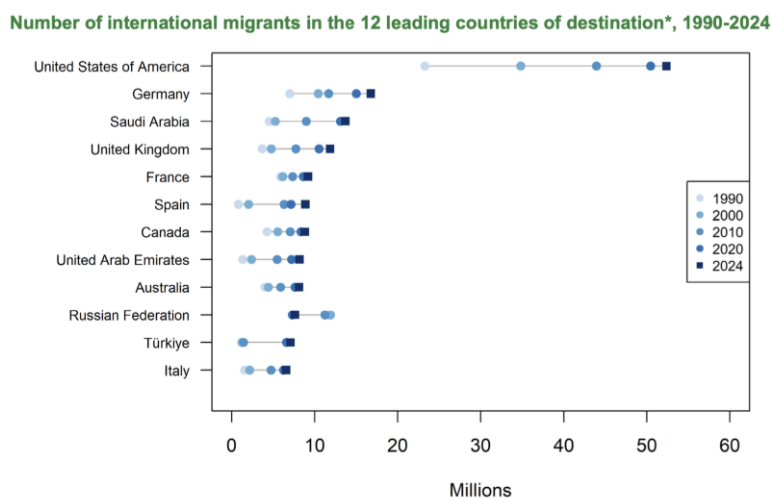


⁴ The United Nations Population Division defines *international migrants* as “any person who has changed his or her country of residence. This includes all migrants, regardless of their legal status, or the nature, or motive of their movement.”. Information available at <https://www.un.org/en/global-issues/migration#:~:text=For%20statistical%20purposes%2C%20the%20United,or%20motive%20of%20their%20movement> (accessed June 3, 2025).

Data source: Our World in Data⁵.

In order to stress one more time how much migration movements have hit Europe, it is enough to see that 5 out of the 12 leading countries of destination are European (**Errore. L'autoferimento non è valido per un segnalibro.**). These figures only emphasize the magnitude of the migration phenomena, highlighting how important it became to understand the impact of increased immigrant concentration, in order to exploit any benefit that comes from a more diverse, inter-cultural society. Italy is found among the countries that are receiving the largest share of migrants in **Errore. L'autoferimento non è valido per un segnalibro.**.

Figure 2.5. Number of international migrants in the 12 leading countries of destination.



Note: *Figure includes destination countries with 5 million or more international migrants in 2020. Destinations are ordered by number of international migrants in 2024 (United Nations, 2025).

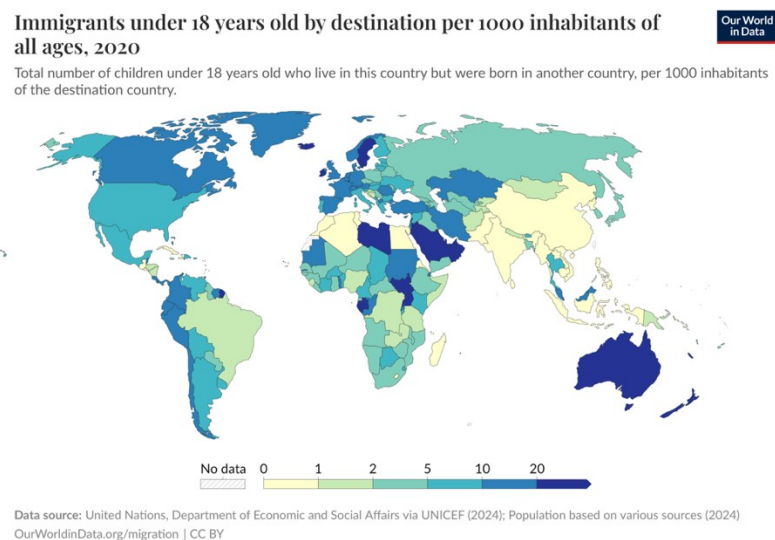
Clearly, the population of migrants in school age has been increasing as well in these areas (Figure 2.6). Most large-receiving countries have implemented through the years different policies for the school system to ensure support to non-native students, to foster their integration and limit any possible detrimental effect coming from the placement of possibly weaker students in schools. We propose a summary of some of these policies, to give an overview of how this issue has been addressed in the last years. There is extensive evidence that immigrant students tend to underperform in school compared to their native peers, most likely because the language of instruction in school is usually different than the one spoken at home. Non-native

⁵ Available at

<https://ourworldindata.org/explorers/migration?time=2024&Metric=International+immigrants&Period=Total+number&Sub-Metric=Per+capita+%2F+Share+of+population&Age=All+ages&country=CAN~FRA~DEU~GBR~USA~IND~CHN~SYR~YEM> (accessed June 3, 2025).

children usually need extra support to master the host-country language, in which native pupils have a clear advantage. If immigrant children do not receive adequate support for learning the language – which is a key to success in school – their integration and future success is hampered.

Figure 2.6. Share of immigrants under 18 years old by destination per 1000 inhabitants of all ages.



Data source: United Nations Department of Economic and Social Affairs via UNICEF (2024). Retrieved from: Our World in Data.

Hakuta (1999) identified five different types of language support used in local schools in the US at the time of their research to support students with limited English proficiency who had an immigrant background. These approaches – though adapted to modern needs – are still in use to date. The first minimal model is the basic *immersion* of children into schools, with no specific language support provided. Another common and more methodical strategy is namely the *structured immersion*, in which non-native students receive instruction in English with an adjustment made to their level of English in the mainstream classroom, and no additional specific support through their native language is offered. A third approach is the *immersion with a preparatory base* according to which non-native students participate in preparatory classes and programs to strengthen their proficiency in the language of instruction before transitioning to the mainstream class. In the *transitional bilingual education* approach, students are initially instructed in their native language, then teaching gradually shifts to the common language of instruction. Where they adopt the *maintenance bilingual education* approach, they offer instruction to non-native students in both their native language and the common language of instruction, with the objective of developing academic proficiency in both languages.

These methods have since then developed and have been applied in other countries of the US with sometimes slightly modifications. The *immersion with a preparatory base* is today a widespread strategy across different countries. Belgium established in 2019 the DASPA – *Dispositif d’Accueil et de Scolarisation des Primo-Arrivants* – schooling program for newcomers, with the objective of ensuring the reception, orientation and optimal integration of newly arrived pupils into the French Community’s education system and preparing them to be responsible citizens that contribute to the development of a democratic and pluralistic society (Eurydice - European Commission, 2024). DASPA program provides preparatory French classes to immigrant children to become familiar with the school system before continuing their education in mainstream classes. The time that non-native children spend in these preparatory classes is not fixed but depends on their progress, and it can go from as little as one week to a year and a half. Besides adapting to the main language of instruction to foster their integration and success in the school system, these classes offer to the children that participate to the program the possibility of encountering and adapting to the country’s socio-cultural system, facilitating their inclusion and their future entrance in the labor market. While assisting non-native children reinforcing their knowledge of the learning language separately, this system allows to avoid any possible detrimental effect on native children that derive from teachers diverting their attention to weaker students. In addition to this program, the French community in Belgium has established partnership agreements with some migration countries to organize courses in the language of origin during pre-secondary and secondary education, to which pupils with foreign origin can participate on a voluntary basis, as well as optional courses on cultural openness for native students, as a way to support the cultural integration with their non-native peers.

France as well engages with DASPO program, that replaced the previously existing *classes paralleles* in 2012, with the purpose of assisting children of foreign origin to learn language of instruction and helping them in reaching as fast as possible the same performance level of native children in corresponding mainstream classes (Infor Jeunes Laeken, n.d.). According to French law⁶, in order to have access to DASPO program children must be newly arrived students at least 2 and half years old and under 18 years of age, arrived in the national territory less than one year prior or, if they are not newly arrived, they must be minor of foreign origin with no sufficient knowledge in the language of instruction who have attended French school for less

⁶ Article 2 of Décret 18/05/2012 available at https://gallilex.cfwb.be/sites/default/files/imports/37785_003.pdf (accessed June 4, 2025).

than one full year. While DASPO program has been proved effective in fostering the integration of migrant children into French education system – both in France and in Belgium – it is not a system free of flaws. First, in order to have access to DASPO funds, these parallel classes must reach a certain threshold, which is sometimes not met. The consequence is that schools with smaller immigrant concentration are not able to fund integration for non-native children. Additionally, there has been raising concern over how to deal with students who, after the maximum period established by law, are yet not ready to move to the mainstream class (CODE - Coordination Des ONG pour les Droits de l'Enfant, 2024). Another aspect to consider is that immigrant children often come from extensively different cultures and speak different languages, which makes it difficult to think that a standardized class can offer each pupil the assistance individually needed.

Cyprus began offering tailored lessons to migrant children in primary schools to improve their capability to read, write and communicate in Greek, through the Geia Xara project in 2018. The idea is to offer non-native children the foundation to progress academically and the possibility to interact with their peers to foster integration. Lessons are offered after regular classes and in four different language levels so that each child is helped according to their own ability and initial knowledge; at the beginning of the project the possibility to participate was offered also to parents, which could then act as intermediaries in their native language with their children. The program was extended up to 2022 (European Commission - Migration and Home Affairs, n.d.).

Greece has committed since 2021 to the ACE – All Children in Education – program, an initiative led by UNICEF, that targets migrant children aged 4 to 17 and their families offering them classes in Greek, English, but also mathematics and science, to favor children entrance into primary school. There is also the possibility to participate to the ALP – Accelerated Learning Program – which support students already enrolled in high schools in biology, history, social sciences, mathematics, physics and chemistry.

Among other European countries who offer specific immersion classes to migrant children there are Germany, Sweden and Poland. In Italy the legislative procedure to establish similar transitional classes, called *classi ponte*, started in 2008, but was soon blocked by many fearing that it would lead to segregation of migrant children. For the past decades there haven't been a specific policy to address the introduction of migrant children into schools in Italy, until 2024, when it established that, starting from 2025/2026, classes with a share of immigrants higher than 20% are entitled to request additional language teacher to assist non-native children to

overcome their weakness in the language of instruction⁷. For the purpose of computing the share of migrant children, only pupils who are enrolled for the first time in the Italian education system are considered, who are expected not to possess the basic language competences necessary to follow the mainstream classes.

2.3 The risk of school segregation

The European Commission - European School Education Platform (2023) - defines school segregation as the “phenomenon in which pupils who live in the same city are separated in different school communities based on their ethnicity, socio-economic status, disability or other characteristics”. Segregation is a form of physical and social separation of a marginalized group from members of a non-marginalized group that causes unequal access to mainstream high quality services. Residential segregation based on ethnicity and socio-economic status is a significant component of school segregation. Catchment areas⁸, early tracking, school admission policies – that can pose financial or religious cultural to enrolment often interlinked with residential segregation⁹ – are factors that can determine school segregation, benefiting some pupils over others.

The risk of children being segregated in schools on the basis of their ethnicity or socio-economic status is a common concern when discussing policies addressed to migrant students. Despite the large consideration given to the integration of people of foreign origin, there is still evidence of high social segregation in European countries, mostly on the basis of the socio-economic background (Marciniczak et al., 2016). Where students are unevenly distributed across schools depending on their characteristics, the resources that contribute to their educational success are unequally distributed as well. The issue of school segregation is closely related to migrants, who also usually have a poorer socioeconomic status. Racial segregation in

⁷ See Art.1 of Law Decree 31 May 2024, n.71, available at <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legge:2024-05-31;71> (accessed June 4, 2025).

⁸ A catchment area is the geographical area surrounding a school from which it usually takes most of its pupils. Distance from your house to the school or whether you have siblings in the same school are, among other, factors that determine admission to schools with catchment areas.

⁹ For instance, residential proximity is often a discerning element in Italian school to decide which children to admit.

school is also often found guilty of the achievement gap of non-native children with respect to native ones, as highlighted by literature (Card & Rothstein, 2007), and today there are numerous efforts aimed at reducing racial segregation in the European landscape. The ECASS – *European Cities Against Social Segregation* – project tackles school segregation in Europe and foster educational inclusion, coordinating policies to ensure social mixing in residential areas and strategic location of schools (European Commission - European School Education Platform, 2023). Today it is implemented only in three cities, Barcelona, Milan and Oslo, in order to test policy innovations aimed at the elaboration of strategic guidelines for policymakers in Europe. The *multi-school catchment area initiative* that merges school catchment areas to rebalance their social intake and the *AFFELNET school choice reform* that changed the way in which children were assigned to public high schools are two other policies introduced in Paris to strengthen social diversity in school (European Commission - European School Education Platform, 2023). Nevertheless, before engaging in any discussion on segregation and the implementation of related policies, it's important to understand what the consequences on students' school achievement can be. Immigrants are often located in disadvantaged residential areas with lower housing costs, together with native families from a poorer socio-economic background. The observed relation between immigrant concentration and poorer school performance could be due to the clustering of children of disadvantaged origin in schools of lower quality. The idea that where there are more immigrants school performance is lower can drive native families to move their children to schools with less migrants, exacerbating existing segregation. The key issue is to understand segregation from an efficiency perspective. If migrant children negatively affect native children but can have a positive or null influence on other non-native students, and vice versa, then school segregation is efficient, as it maximizes students' performance. Conversely, in the case in which children of foreign origin exert negative or null effects on their migrant peers but positively influence natives' performance, then pursuing de-segregation is the most efficient strategy. If one comprehends the efficiency of de-segregation, it would be possible to understand potential gains coming from the re-allocation of children across schools and classes. Redistributing immigrant children to schools with lower shares of non-native students to fight segregation is optimal under efficiency considerations if the gains coming, for instance, from improved total performance are larger than the losses. If we assume peer effects to be linear, redistributing students does not result in an efficiency gain. The student's outcome depends only on the average of the characteristics of their peers when we model linear peer effects; reallocating students change only the distribution of these characteristics but has no effect on the average. In order to study if reallocating students can increase total performance, it is necessary to think about different formulations of the

functional form between immigrant concentration and school performance. Non-linear peer effects are especially interesting. If we assume that the relationship between immigrant concentration and students' performance is concave – that is to say, there are marginally increasing negative effects – reallocating students can be efficient. Formally it can be written as (Andersen & Thomsen, 2011):

$$f(tm_1 + (1 - t)m_2) > tf(m_1) + (1 - t)f(m_2)$$

Where m_1 and m_2 are the immigrant concentration in school 1 and school 2 respectively and $f(m_i)$ is the function that relates the share of immigrants to students' performance. When $f(m)$ is strictly concave, the inequality holds true. The right-hand side represents the weighted average achievement of the two schools with two different shares of immigrant. In other words, redistributing immigrant children so that both schools have the same share comprised between m_1 and m_2 increases total performance, that reaches its maximum with $f(\bar{m})$.

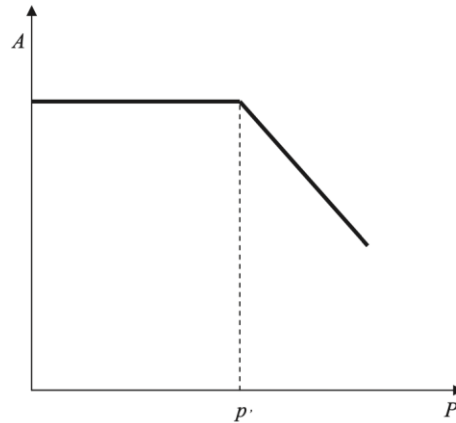
Conversely, if the effect of immigrant concentration is still negative but marginally decreasing, the relationship between immigrant concentration and students' performance is better represented by a convex function. In this case the negative effect of one additional immigrant student is smaller in schools with already high immigrant concentration, and with a strictly convex function it becomes:

$$f(tm_1 + (1 - t)m_2) < tf(m_1) + (1 - t)f(m_2)$$

In this case, the larger the gap between m_1 and m_2 the greater is students' achievement. According to this framework, segregation, in the sense of classes stratified by ethnicity, is technically efficient.

A particular case of non-linear relationship is the *tipping point*, in which the negative effect of immigrant concentration occurs only above a certain threshold. In this context, if one school is above the critical point, redistributing students can increase performance; otherwise, there is no efficiency gain coming from reallocation. Graphically it can be represented as in Figure 2.7.

Figure 2.7. Tipping point.



Note: Graphical illustration of a concave function with a tipping point.

Note: P in the horizontal axis is the share of immigrant students. A in the vertical axis is total achievement. Data source: Andersen & Thomsen (2011), p.32.

One final possibility is that peer effects are heterogeneous or asymmetric, meaning that they do not affect all students equally. It may be that immigrant students are more negatively affected by the presence of immigrant peers than native children or vice versa. It can be shown that even if peer effects within each group are linear, the aggregate effect is non-linear, and the same conclusions observed in the case of convex and concave functions hold true¹⁰.

In essence, in order to find out whether redistributing students lead to efficiency gains, the first fundamental step is to understand the functional form of the relationship between immigrant concentration and students' achievement, and in the case it is non-linear, whether it is concave or convex. Researching whether peer effects arising from immigrant concentration are asymmetric and heterogeneous across different groups of students can also help in understanding if the relationship is non-linear. In addition, Andersen & Thomsen (2011) emphasize that assessing if peer effects are asymmetric can also help to increase educational equity, if reallocating students can benefit specific groups of children.

¹⁰ Andersen & Thomsen (2011) offer an interesting and complete mathematical demonstration of this in pages 33-35. They are able to show that even if the effect of immigrant concentration is linear for both native and immigrant children separately, the effect on all students is still quadratic, formally: $A_T = \alpha_N + (\beta_N + \alpha_I + \alpha_N)p + (\beta_I - \beta_N)p^2$. Similarly, they show that if either or both peer effects on each separate group quadratic, the total relationship is still quadratic in any practical case.

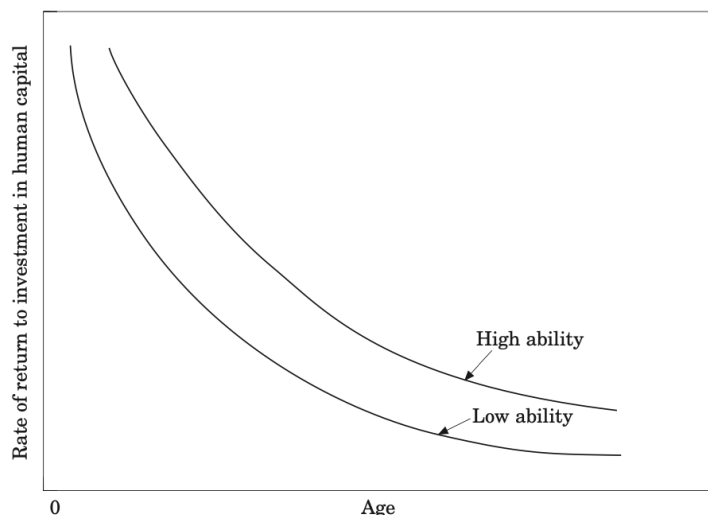
Nevertheless, mathematical theory collides with social norms; even if there was proof that complete segregation would be the most efficient option for all students in terms of efficiency, it could be ethically implemented as a school policy.

2.4 The role of primary education

The purpose of this study is to estimate how the concentration of non-native students can affect the long-term performance of their native peers. The idea is to verify whether any peer effect arising from immigrant concentration in primary school persist up to upper secondary school. The evidence will likely point in one of two directions: either peer effects stick over time, or they do not. If it was the first circumstance, policy makers should be interested in directly shaping school and classroom composition, in order to maximize any existing positive effect or minimize detrimental ones, boosting students' performance and human capital. If there is no detectable or sizeable immigrant peer effect that endures up to upper secondary school, then there should be no concern with respect to school composition. This analysis is based on the assumption that primary school experience can significantly shape children, more than what lower and upper secondary education can do. This premise is not without logical basis and scientific proof.

There is large evidence that what happens in early life deeply affect individuals. Effective and high-quality early learning is positively associated with school attainment and employment, as well as physical and mental health (Shuey & Kankaraš, 2018). Figure 2.8 illustrates the rate of return to investment in human capital over age and conveys one important information: for the same level of investment at each age, the return is much larger on young individuals than it is for older ones.

Figure 2.8. Rate of return to investment in human capital.



Source: Heckman (2000).

The long-term implications of early experiences are more evident if one thinks of the cognitive and non-cognitive skills that pupils learn during primary education. Throughout the years of primary schooling, children learn to read and to write and they found their basis in mathematics and sciences: there is strong evidence that early academic skills in language, literacy and numeracy are positively associated with later educational attainment¹¹. Sabates & Parsons (2012) find that low mathematical scores at age 10 are related with poor numeracy levels later in life. In addition to basic educational skills, children at this time start engaging with peers and adults that are out of the context of their family, most likely establishing the foundation of social interaction for the upcoming years¹². Any element that can compromise this learning-path can halter future school performance, even when education becomes more specific, and strongly related to the individual ability.

It is reasonable to assume that peer effects in higher education have a more subdued influence on other long-term outcomes. At the time of secondary education, especially at the upper level, it is plausible that most teenagers have a stronger sense of self and not influenced at a deeper

¹¹ Shuey & Kankaraš (2018) propose a useful and thorough resume of research that investigate the impact of early learning. In pp. 27-30 they focus on the effect on educational attainment. Interestingly, they recall some small-scale studies which suggest that even the styles of play in the earliest years of education may predict future educational outcomes.

¹² For a more detailed and scientific explanation of the role of early experiences on the memory system see also Nelson (2000).

level by their peers. Also, in Italy school tracking begins at the end of lower secondary school. Early tracking deeply affects peer composition, in a way that schoolmates tend to be similar in terms of ability, interests, and attitude towards schooling. From a certain point of view, peer effects in higher education are themselves an outcome of peer effects in earlier life.

3 Literature review

This chapter offers a comprehensive review of the existing literature on peer effects, with special attention to the role of non-native students. The objective of this review is to critically examine methodological approaches and empirical evidence that have shaped the current understanding of peer influences within educational settings.

3.1 Understanding peer effects and the challenges of identification

The term *peer effect* is used to describe the various mechanisms through which individuals' attitudes, behaviors and outcomes are influenced by their immediate social environment. Broadly speaking, peer effects refer to externalities in which behaviors, characteristics and achievements of one's peers affect individual outcomes. The significant role of peers in shaping the educational process was first acknowledged in the landmark report from Coleman et al. (1966), which recognizes fellow students as a fundamental component of a child's school environment and, consequently, an important influence on their educational experience. For the first time, experts began considering overall school and classroom composition as a determinant of schooling quality and educational outcomes.

As a matter of fact, the academic environment in which children learn is not solely determined by school-level resources and teachers' quality but is also shaped by the characteristics and behaviors of classmates. Manski (1993), tries to conceptualize these dynamics and identifies three types of effects that may explain the observed tendency of individuals within the same group to behave similarly, namely: endogenous effects, exogeneous (contextual) effects and correlated effects. In the first hypothesis, the propensity of an individual to behave in a certain way varies in response to the behaviors exhibited by the group. In the second, individual behavior is influenced by the exogeneous characteristics of the group, such as socio-economic background, parental education or language proficiency. According to the third hypothesis, similarities in a group emerge because individuals within a group share pre-determined characteristics or are exposed to similar environments. The fact that these three forces operate simultaneously within a given setting, makes isolating specific types of peer effects extremely challenging.

Today, it is widely recognized that peer composition constitutes a key determinant of student's educational outcomes, thereby fueling the interest in understanding both the nature and the direction of peer influences. In the context of education peer effects may arise through several

mechanisms. The most direct channel involves students providing instruction, guidance or assistance to one another. Higher-achieving peers can enhance classroom learning by raising academic standards, fostering competition, offering informal tutoring or facilitating knowledge spillovers. On the contrary, disruptive or worse-achieving classmates can lower classroom quality and divert teacher attention and school's resources. Furthermore, peer groups play a crucial role in transmitting aspirations, behavioral norms and attitudes toward schooling, influencing academic engagement and achievement. Peer effects that generate from classroom composition are of particular relevance to the present study. Classroom composition can shape the learning environment in various ways, affecting the collective educational experience. For instance, wealthier parents may contribute resources that have spillover effects within the whole classroom. Ethnic or linguistic balance within a class can influence students' outcomes and social development through their impact on the nature of interactions and the learning opportunities available to each student. Exposure to different culture, language, social norms, or to socio-economically diverse peers can enrich students' social capital and improve their communication and collaboration skills. Yet, such diversity can also introduce challenges. Non-native pupils may require language support, remedial instruction or additional attention from teachers, potentially reducing instructional time and resources available to the other students. In addition, processes of segregation and social clustering – whereby students gravitate towards ethnically or socially similar peers – may limit opportunities for meaningful peer learning and interaction. Hoxby (2006), when examining classroom racial composition, emphasizes the importance of the responses of teachers and school administrators to classroom diversity as another key aspect of peer influences. If teachers expect less from minority or immigrant children, they may lower academic standards in classroom with high proportion of non-native students. In such cases, native students might experience negative peer effects not directly attributable to the characteristics, behaviors and abilities of their immigrant peers, but rather to the altered expectations and instructional practices of teachers in response to class composition. Sacerdote (2011) provides a valuable synthesis of some of the mechanisms through which peer effects might operate in educational settings. The most common estimated model of peer effects in literature is the *linear-in-means* model, in which the outcome of interest is modeled as a function of the average background characteristics and average outcomes of peers, in addition to the individual-level background controls. This framework however presents notable limitations. Primarily, it assumes that peer effects operate exclusively through the mean of peers' characteristics and outcomes. As a consequence, it overlooks any effect that arises from other features of the distribution, such as mean preserving changes in the variance of peers'

characteristics. In practical terms, the model assumes no distinct effect from having either an exceptionally good peer or a particularly disruptive one, beyond what is captured by the average. This simplifying assumption allows researchers to isolate mean peer effects but precludes analysis of more nuanced, distributional or complex peer effects. According to Sacerdote (2011), this model generates two additional conceptual challenges. First, from a social welfare perspective, this model is not interesting nor useful, as it posits the net effect of reallocating peers across classrooms is zero. If an extraordinarily good student is exchanged for a particularly weak one between two classes of equal size, the positive effect exerted by the good peer is offset by the negative effect of the bad peer at the aggregate level. Second, recent empirical evidence suggests that peer effects are unlikely to be linear-in-means. Research like the one by Lavy, Silva, et al. (2012) show for instance that the strongest peer influences are exerted by the students in the bottom 5% of the ability distribution, whose presence is detrimental to pupils' learning, while they find no significant effect on educational outcomes coming from the top 5% of the ability distribution.

An alternative theoretical framework is the *bad apple* model, introduced by Lazear (2001), which posits that the most relevant peer effects are exerted by those classmates at the end of the tail of the ability distribution, meaning the less academically able and less disciplined students in the classroom. The "bad apple" pupil may engage in disruptive behavior, diverting teachers' and classmates' attention from respective instructional tasks, or demand disproportionate amounts of teaching support due to their low ability, thereby detracting resources from other students and reducing the quality of the learning environment. As opposed to that, the *shining light* model suggests that an exceptionally high-achieving student can provide a positive role model, raising expectations and performance for the entire class.

Without digressing into a comprehensive and exhaustive catalogue of peer effect models, it is worth noting that most frameworks focus on academic ability as the principal driver of peer effects. Such ability-centric perspective may prove inadequate when investigating peer effects exerted by immigrant students. The presence of immigrant peers can influence outcomes through multiple mechanisms beyond academic ability, including language proficiency, cultural integration, social dynamics and teacher-students' interactions. Consequently, frameworks that conceptualize peer effects as driven solely by academic ability may fail to capture the broader social, linguistic and contextual dimensions that are particularly relevant in heterogeneous educational settings.

Regardless of the specific peer effect of interest, the primary methodological step to take is the identification of the causal effect. There are at least three reasons that explain why simple OLS

estimation may be problematic. First, if peer effects do exist, an individual's outcome both affects and is affected by the outcomes of their peers, giving rise to endogeneity concerns. This simultaneity issue is known as the *reflection problem*, as it was first outlined by Manski (1993), which makes it difficult to disentangle whether a student's performance is the cause or the consequence of their peers' average performance. Without appropriate instruments or experimental variation standard OLS estimation would most likely be biased and inconsistent because of this reverse causality. Second, even when simultaneity is fully addressed, distinguishing between endogenous and exogenous peer effects, as defined earlier, is inherently difficult.

Third, the issue of self-selection further complicates identification. It is plausible that students self-select into peer groups based on unobserved characteristics such as ability, motivation and socio-economic status, leading to biased estimates of peer effects. In addition, the selection problem could extend to the allocation of students across classrooms within the same school. It may be that administrators cluster higher-ability students in the same class, or strategically assign less able pupils to classes with better teachers and resources. In such cases, the distribution of students across classrooms is unlikely to be random. At a broader level, residential sorting processes also shape the distribution of students across schools. Families do not randomly locate to neighborhoods, but rather select into residential areas based on housing costs, employment opportunities and socio-economic characteristics. As a result, students from lower socio-economic background often cluster in the same schools, exacerbating existing educational inequalities. This aspect is particularly important when studying peer effects associated with immigrant students, given the potential for both residential and educational segregation. Apart from residential proximity, families may choose school based on neighborhood characteristics, perceived peer quality and socio-economic composition, leading to a systematic sorting of students into schools.

In response to these challenges, economists and social scientists have developed multiple identification methods and credible measurement of peer effects while mitigating source of bias. These approaches can be broadly grouped into four categories: quasi-experimental designs, fixed effects models that exploit for example within-school or within-classroom variation, instrumental variables approaches and aggregation procedures. Each of these methodologies offers advantages and limitations, and their application has significantly advanced the understanding of peer effects in education.

3.2 Measuring the invisible: identifying peer effects in schools

The literature on peer effects has often produced mixed and sometimes contradictory results, largely due to the substantial challenges posed by identification. Empirical analysis of the causal effect of peers on individual-level outcomes requires disentangling a complex web of social influences while accounting for multiple confounding factors. Such task inevitably involves making assumptions about the nature and the structure of social interactions. Since peer effects started attracting systematic attention, most research has focused on socio-economic background, gender, and ethnic differences as key channels through which peer influences manifest in educational settings.

Quasi-experimental frameworks offer particularly valuable opportunities to study peer effects. Such natural experiments are rather rare and relatively scarce in literature. To our knowledge only Angrist & Lang (2004), Gould et al. (2009) and Geay et al. (2013) have successfully exploited quasi-experimental variation to estimate peer effects at the school level. While the research by Gould et al. (2009) will be discussed later in this chapter, it is worth at this stage to review the methodologies and findings of Angrist & Lang (2004) and Geay et al. (2013), which represent influential contributions. At the post-secondary level quasi-experimental evidence is provided by Zimmerman (2003) and Brunello et al. (2010). Zimmerman leverage the random assignment of college freshman to dormitories at Dartmouth college to determine peer effects among roommates in academic outcomes. Brunello, De Paola and Scoppa similarly exploit random assignment of roommates for first-year students in the University of Calabria to determine whether roommates peer effects exist and whether they vary with the field of studying.

Angrist & Lang (2004) examine peer effects in the United States, exploiting the Metco program, one of the largest and longest-running desegregation initiatives in the country. Established in 1966, Metco fosters desegregation of minority students, mostly African American in this case by facilitating their transfer from Boston's urban schools to other districts' schools, altering classroom composition and providing a valuable setting to assess peer influences. Participation to the program is entirely on a voluntary basis for both minority students and receiving school districts, generating an exogenous source of variation in classroom peer composition. Initial OLS estimates find no significant effect of the presence of Metco students on the achievement of native students, who are typically White. Recognizing the potential concerns in the OLS specification, in particular the possibility of a spurious negative correlation between the fraction of Metco students and native students' outcomes, the authors adopt an instrumental variable strategy. They leverage a modified version of the Maimonides' rule – previously used by

Angrist & Lavy (1999) – exploiting the fact that Metco students are assigned to schools partly on the basis of space constraints and that class size in Brookline is capped at 25 pupils. Their IV estimates are consistent with the OLS findings, showing no significant negative effect of increasing the fraction of minority students on the educational outcomes of their native peers, a result that is in contrast with most existing literature reporting negative peer effects associated with minority or immigrant students concentration.

Geay et al. (2013) similarly investigate the relationship between the percentage of non-native English speakers on scores in reading, writing and mathematics in English schools. Beside standard OLS models, they estimate an instrumental variable model that exploits the shock to the demand for catholic schooling following the entrance of ten Central and Eastern European countries in the European Union in 2004. The instrument leverages the observed preferences for Catholic education among Eastern European immigrant families, which generated an exogeneous variation in the proportion of non-native speakers in such schools. Given the nature of their instrument, the authors restrict their sample to exclude non-native speakers of non-white ethnic origin for whom the instrument is not valid. As a consequence of their strategy, their estimates capture the Local Average Treatment Effect, that is in this case the causal effect of immigrant peers on native English speakers that attend Catholic schools with the group of pupils that have been affected by the instrument. Their findings shows that the modest correlation observed in the baseline specifications is ruled out as soon as a very limited number of pupils' characteristics is introduced as control or school fixed effects are added. They conclude that much of the negative correlation is explained by previous sorting of non-native speakers into schools with less desirable characteristics.

Aggregation procedures and fixed effects are other common strategies that are frequently employed to address the identification problem of peer effects-related research, taking different forms depending on the level of aggregation, whether at classroom, school or even country level. A particularly common approach exploits variations in cohort compositions. It builds on the assumption that such year-to-year variations in peers' composition are idiosyncratic and beyond the management of families and schools administrators. In other words, differences between adjacent cohorts can be considered exogeneous and unexpected in a way that it is plausible that not parents nor school's principals can predict them and act accordingly in the distribution of pupils. Several studies adopt these strategies, including those by Hoxby (2000), E. Hanushek et al. (2003), E. A. Hanushek et al. (2009), and Brunello & Rocco (2013).

Hoxby (2000) exploits cohort-to-cohort variation within the same grade and school in Texas public schools to establish the existence and direction of peer effects while mitigating selection

bias stemming from parental and school-level behavior. By assuming that peer compositions fluctuations between adjacent cohorts are unpredictable and exogenous, Hoxby demonstrates that classroom composition indeed matter for academic achievements. In particular, their findings indicate that adding higher achieving peers benefit all students. An interesting point they make is there is evidence suggesting that peer effects are stronger within racial groups rather than between them. That is, Black students are more affected by the ability and behavior of their Black peers than by those of non-Black classmates.

E. A. Hanushek et al. (2009) conduct a similar analysis using panel data from Texas school to explore the effect of school racial composition on students' outcomes for native and minority students. By using a rich fixed effects specification – including school, grade, year effects and their interactions – and exploiting cohort differences in achievement and racial composition, the authors find that a higher concentration of Black students at the school level negatively affects the achievements of both Black and White pupils. Importantly, the negative effect is larger and higher in significance for Black students that it is for their White peers, in line with what indicated by Hoxby (2000). They also notice that the inclusion of the proportion of Hispanic students does not halter the estimated effects, implying that the observed outcomes are driven by the share of Black students rather than the broader concentration of minority pupils.

Brunello & Rocco (2013) take a novel approach by employing cross-country data from the OECD Program for International Student Assessment (PISA) covering European and Anglo-Saxon countries to determine whether rising shares of immigrants in schools and classrooms affect the educational outcomes of native pupils. They address self-selection biases resulting from the non-random distribution of students across schools by aggregating data at the country level. Nonetheless, it is plausible to think that immigrants self-select into countries as well. To avoid any bias that may arise from that, they include country and time effects as well. Besides that, they incorporate additional controls such as the growth rate and the level of real GDP per capita, public expenditure per pupil at secondary school level and total immigrant stock. Furthermore, they include the Duncan index or segregation index (see Duncan & Duncan (1995)), to account for the likely uneven distribution of immigrants across schools in the same country. Interestingly they observe a negative correlation between the share of immigrant pupils and the segregation index, suggesting that countries with a relatively higher share of immigrant students tend to display a more even distribution across schools. Their results show a small negative effect of the share of immigrant students on the school performance of 15 years old native pupils, that varies by gender and parental background. In particular, the estimated effect

is larger for females and for natives with poor parental background¹³. To address potential attenuation bias in their estimates, Brunello and Rocco employ a split-sample IV strategy. They randomly divide their sample in two, measuring the immigrant share in each sample, and use one subsample's measure as an instrument for the other's. This procedure is repeated 500 times and yields estimates that are only slightly larger than the baseline coefficient, suggesting that attenuation bias is a minor problem in this context. Lastly, they find that the negative peer effect is stronger in countries where immigrant segregation across school as measured by the Duncan index is higher.

Besides these important papers, McEwan (2003) uses within-school variation in peer characteristics to investigate peer effects on students' achievements in Chile. Ammermueller & Pischke (2009) similarly exploit within-school within-grade classroom variation, arguing that classes within schools are formed randomly with respect to family background characteristics. Card & Rothstein (2007) aggregate data at the metropolitan level and exploit variation across metropolitan areas to measure the degree to which racial segregation affects the black-white gap in test scores. Lavy, Silva, et al. (2012) use within-pupil variation and exploit variation in achievements across three compulsory subjects to estimate ability peer effects in England. Seah (2021) uses variation in the share of immigrant pupils across two different grade levels within the same school to address endogeneity issues.

The use of instrumental variables remains the most debated strategy for identifying peer effects, particularly because of the difficulty in finding credible and valid instruments. Among the contributions that tries to resolve this challenge, the study by Boado (2007) is especially noteworthy. Boado investigates peer pressures resulting from the concentration of non-native students in French schools at the secondary level school that can affect educational outcomes, measure by number of years spent in lower secondary school, marks in the 8th and 9th grade¹⁴, and track selection in upper-secondary school. The author employs two instruments, one for the school environment and one for the neighborhood. The first one is a dummy equal to 1 whether the school is located in a priority education area, to capture school quality. The second one is a subjective measure of neighborhood quality based on the degree of satisfaction reported in the

¹³ They identify children with good parental backgrounds as those with a number of books at home larger than the country average, and with poor parental background those who have fewer.

¹⁴ Lower secondary school in France lasts four years. The 8th and 9th grade, in France respectively named 4th and 3rd grade, are at the lower secondary level.

family questionnaire over the district in which the family resides. This latter instrument builds on the assumption that families who report higher dissatisfaction are more likely to live in deprived districts, which tend to have higher immigrant concentrations. Boado posits that this variable can explain why immigrants and natives from low socio-economic backgrounds cluster in the same areas. Using data from the Panel of Students in Secondary Education (1995-2001), a large dataset produced periodically by the French Ministry for National Education which includes explicit information about the family's migration history, they find negative correlation between the concentration of foreign students and all three proxies of educational success in the baseline OLS estimates. The standard OLS models however do not take into account the prior sorting mechanisms that distributes pupils across school divisions. When addressing the non-random sorting using the IV strategy, the coefficient of interest loses statistical significance. They conclude that the low level of attainment that is usually found in schools where there are more foreigners is more likely the result of the disadvantaged socio-economic background of the students rather than peer effects per se. Their inferences do not dismiss the existence of peer effects but rather questions the concentration of non-native students as a relevant mechanism of such effects. They conclude that simply interacting with foreign pupils rather than native ones does not affect school outcomes. Boado highlights how estimates of peer effects heavily depend on the assumptions and contextual information embedded in the chosen identification strategy, with more sophisticated techniques often challenging the prevailing narrative about how concentration of ethnic minorities and foreign students in schools is a harm to natives' educational attainment.

Other research who attempts to employ instrumental variables are from Evans et al. (1992), Gould et al. (2009), Hunt (2017).

3.3 Do immigrant peer effects last?

The existing literature on long-term effects of peer influences is notably limited, primarily due to the lack of appropriate longitudinal panel data. The ideal framework would provide information on peers merged with long-run outcomes of individuals, a combination that is rarely available. In addition, even when such data are available, it is challenging to isolate separately the effect of peers over extended periods of time in a credible way. To the best of our knowledge there are only two studies that attempted to infer something about the long-term impact of peer effects on educational outcomes, provided by Gould et al. (2009) and Black et al. (2010). While

the former focus specifically on immigrant peer effects, the latter examines peer group composition more broadly, considering individual characteristics such as gender and age.

The study that is the closest to our research is that of Gould et al. (2009), who investigate the long-term impact of immigrant peers in elementary school on high school outcomes of native students in Israel. Their study – as mentioned earlier in this chapter – exploits a quasi-experimental setting created by the large migration wave from the former Soviet Union to Israel in the early 1990s. The validity of their identification strategy relies on the substantial variation in immigrant allocation across schools generated by the size of the migration wave. Their empirical strategy to address the endogenous placement of immigrants across schools draws on random variation in the number of immigrants across grades within the same school. Their core assumption is that – conditional on the total number of immigrants in grade 4th and 6th in a given school – the number of immigrants in 5th grade can be as good as random, as it is defined solely by exogeneous factors such as the year of birth. Their findings suggest that immigrant concentration in 5th grade has a marginally significant negative effect on matriculation rates and a small and statistically insignificant effect on dropout rates once they control for immigrant concentration in adjacent grades. Interestingly, they find this effect to be comparable in magnitude to that of a modest reduction in the father's education level. To reinforce the credibility of their estimates, they conduct a placebo test examining whether the immigrant concentration in the 2nd grade of a given school affects outcomes of students in the 5th grade, finding no significant effect. This supports the idea that their strategy successfully captures the causal impact of peers rather than unobserved school-level characteristics.

An important result emerging from their analysis is that the negative impact of immigrant concentration is stronger for students from a lower socio-economic background. Moreover, the negative effect of immigrant presence appears to diminish as immigrant concentration grows, suggesting that sufficiently large presence of immigrants may facilitate their integration, mitigating peer-related disadvantages. The authors address an additional potential bias stemming from the possibly endogenous grade placement of immigrants. They fear that immigrant students, especially those recently arrived from the former Soviet Union, were more likely to be held back a grade. Their data support this idea as they find that a large number of non-native students are in a grade different to the one they should be placed into given their year of birth. To correct for this, they adopt an instrumental variable approach, instrumenting the number of immigrants in a grade with the *predicted* number of immigrants in said grade based on the exogenous year of birth. Their IV estimates are not different from the OLS coefficients. Lastly, they avoid the problem of endogenous classroom placement of immigrant

students by computing immigrant concentration and the grade-school level rather than at the class-grade-school level, a common strategy that will be employed in this work as well. While effective in mitigating selection issues, this method is arguably suboptimal, as it is likely that most peer effects actually happen at class level, where daily interactions occur, rather than at school level.

Despite effectively identifying the impact of immigrant peers at the elementary school level on drop-out rates and matriculation exams' results, the study does not provide evidence on how immigrant concentration affects native students' general performance, as it can be measured by standardized test scores.

Black et al. (2010) uses Norwegian administrative and schools' data to examine long-term effects of peer group composition – defined by gender, age and socio-economic status – on various outcomes, including IQ scores, teenage childbearing, educational track and labor market performance measured by earnings. While they do not investigate immigrant concentration, their study contributes to the literature on the persistence of peer effects. Using idiosyncratic variation in peer characteristics and an instrumental variable strategy they find that neither peers' age and mother's education used as a proxy for the socioeconomic status do influence long-term outcomes, whereas they find evidence that the proportion of females affect positively long-term outcomes of girls and negatively those of boys.

3.4 Evidence from Italy and research gaps

A limited number of studies have examined peer effects arising from immigrant students in the Italian context, including Contini (2013), Tonello (2016), Ballatore et al. (2018) and Frattini & Meschi (2019).

Contini (2013) uses INVALSI data for the 2nd, 5th and 6th grades in 2010, when the assessment was administered to the entire population of children. Assuming random distribution of pupils across classrooms and using within-school random variability across classes they find that the concentration of immigrant students is detrimental to the performance of other immigrant students and low social status natives. Better performing native students seem to slightly benefit of higher concentration of immigrant students. They draw two important policy implications. The first is that the concentration of immigrant pupils in Italian schools should not be a major concern, as the evidence on detrimental effects are rather small. The second is that, since there is a minor negative effect on the performance of other immigrant students and low social status

native students, schools and classes should host children with heterogeneous family backgrounds.

Tonello (2016) retrieves as well data from INVALSI but they employ data from the 8th grade and draw their strategy on within-school idiosyncratic variation in the non-native share between adjacent cohorts to test the impact on both language and math scores, finding that immigrant school share has a weak negative effect on the performance of native peers and that these effects are mostly identified for male students, whereas results for female students are usually not statistically significant. Moreover, they find that both low and high achieving pupils are affected by the concentration of immigrant students, though low achievers suffer more the negative externalities in comparison to the high achievers. The effects they estimate are nonlinear and increasing in the share of immigrant students in the school. They believe their results to be consistent with the *integration model*, according to which – for sufficiently low values of non-natives school share – non-native students' disruption does not halt the educational process.

Ballatore et al. (2018) adapt the empirical strategy of Angrist & Lang (2004) to their context. They use a Maimonides-type rule of class formation – making use of the enrolment thresholds in the class formation process in Italy - to solve issues of non-random distribution and estimate the effect of PEC – Pure Ethnic Composition – on mathematics and reading INVALSI test scores. They find statistically significant negative effects of PEC on natives' performance, less precise for math test scores and larger in size when considering only first-generation immigrants. By comparison of these results to those obtained exploiting variation between schools withing the same educational system – a strategy that builds on the one of Contini (2013) – they find evidence that principals tend indeed to allocate immigrant students to smaller classes or to classes with better performing native pupils.

Overall, results on peer influences exerted by non-native pupils remain mixed and do not consistently point in a single direction. Findings seem to be highly sensitive to the empirical strategy adopted to identify the causal effect of interest. More importantly, there is lack of evidence on the long-term implications of immigrant peer effects. Yet, understanding these long-term consequences is essential for at least two reasons. First, while contemporaneous school performance is important, it likely does not fully capture the cumulative and potentially persistent influence of peer composition on individuals' educational trajectories, future opportunities and labor market outcomes. If negative effects deriving from the concentration of non-native pupils in school were to last over time, it could lead, for instance, to delayed graduations, reduced access to higher education, or altered occupation choices. Standardized test scores at a given point in time as affected by synchronous immigrant peers share may fail

to detect enduring effects. If on the contrary there were lasting positive spillovers deriving by interactions with immigrant peers, it could justify effort towards more heterogeneous classroom composition.

Second, policy debates over school segregation and class composition are often motivated by concerns over social integration and long-term economic equity. If immigrant peer effects dissipate over time, then fears about school ethnic composition being harmful to natives' outcomes may be overstated. Conversely, if negative effects persist, it would justify policy interventions to address inequalities in early peer environments. In this way, understanding whether immigrant peer effects extend beyond school years and shape adult outcomes is not only a gap in literature but also a question of substantial policy relevance.

Frattini & Meschi (2019) try to identify the effect of immigrant peers on the performance of native students enrolled in vocational training institutions. They focus on data from Lombardia, the largest region of Italy and one experiencing sizeable increase in the population of immigrants. They exploit random variation across classroom and across adjacent cohorts within schools, and find that the presence of immigrant students negatively affects the performance of natives in mathematics, but they find no effect on literacy scores. They conclude that immigrant peers exert adverse effects on the performance of native students. Nonetheless, their results may be driven by the specific population they consider. They recognize that the vast majority of native students enrolled in vocational training institutions fall into the bottom end of the national ability distribution. Their results are more easily interpreted as the effect of immigrant presence on the performance of low-achievers.

To our knowledge, no research has yet investigated long-term outcomes of native students in Italy. This study aims to fill this gap by examining the impact of immigrant concentration at the school level in 5th grade on INVALSI standardized test scores in 8th and 13th grade.

4 Data

This chapter outlines the structure and characteristics of the data used to estimate the long-term effects of immigrant concentration on students in Italy. It presents the dataset employed for the analysis, describes the process of sample construction, introduces the key variables, and provides some descriptive statistics.

4.1 Data sources

This study relies on data provided by INVALSI, the *Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione* (National Institute for the Evaluation of the Education and Training System), which represents a priceless source of data for investigating peer effects in the Italian education system. INVALSI carries out systematic and periodic assessments of the knowledge and skills of Italian students, as well as the overall performance of the school system. Established in the mid 90s, the Institute has progressively expanded the scope and the frequency of its evaluations and, today, serves as the primary institution responsible for monitoring educational outcomes in Italy.

As declared in article 2 of its Statute¹⁵, the core purpose of INVALSI is to promote the improvement and development of the national education system, conceived as a driving force of the Italian economy. This mission underscores the central role that education plays in fostering individual opportunities and national economic performance, an aspect that positions INVALSI data as particularly valuable for economic studies and social research.

Since its birth, the Institute has been entrusted with both evaluating the school system and surveying learning outcomes of students. The key feature of INVALSI assessments is their capacity to ensure comparability across a decentralized and heterogeneous educational system, characterized by significant variations in school types and resources, regional economic conditions and demographic profiles. The use of standardized tests ensures comparability of results across different schools and regions, and it has become common practice in almost all European countries. A systematic evaluation of the educational system allows to intervene on its flaws and obstacles that limit students' opportunities. The preparation of the tests is entrusted

¹⁵ Available at

https://www.invalsi.it/amm_trasp/documenti/attigenerali/Statuto_Invalsi_Post_Decreto_218.pdf (accessed April 22, 2025).

to over two hundred authors that are selected among teachers, school principals, researchers from the Institute and national and international universities. Each question aims to measure a specific skill, and it must not favor or penalize for individual characteristics such as gender and type of school. The assessments are pre-tested multiple times on selected samples of students to ensure their precision and reliability¹⁶.

At present, standardized tests are administered yearly to students in the 2nd and 5th grades of primary school, the 8th grade of lower secondary school and the 10th and 13th grades of upper secondary school. INVALSI tests take the form of a census and are distributed to the whole population of students. Participation rates are typically high in most regions. Most recent available report from the Institute sets the coverage rate above 90% in almost all regions, with peaks at 98%, for the year 2021-2022 (INVALSI, 2022). The assessments evaluate core competencies in reading and mathematics, with an additional English proficiency test¹⁷ introduced only for the 8th and 13th grades. The tests are designed according to national standards and offer longitudinal comparability, thus facilitating panel analyses.

A key strength of INVALSI dataset is the rich set of background variables collected for each pupil. Beyond test scores, the Institute gathers comprehensive information on the socioeconomic and cultural background of students, including the place of birth of the parents, their educational attainment and current occupation, whether the students attended nursery school and kindergarten, as well as details on the possession of material goods, which are used as proxies for an economically and culturally supportive home environment. These data are aggregated into the Economic, Social and Cultural Status (ESCS) index, an important input that eventually allows to control for background heterogeneity when analyzing educational outcomes.

Another crucial feature of INVALSI data, and a particularly important one for the purpose of this research, is the availability of information on the students' origin. Children are identified

¹⁶ More on the specific process of creation of INVALSI assessment is available at https://www.invalsiopen.it/area-prove/invalsi-secondo-invalsi-per-saperne-di-piu/?_gl=1*_mgsalw*_ga*ODM0MTk3ODkxLjE3NDUzMTA3MTU.*_ga_5LBB0NMXR*cZ3NDcyMTUxOTckbzgkZzEkdDE3NDcyMTU0MDYkajAkbdAkaDA (accessed May 14, 2025) and at <https://www.invalsiopen.it/prove/come-nascono-le-prove/> (accessed May 14, 2025).

¹⁷ Detailed information on the competencies and skills measured for each subject are available at <https://www.invalsiopen.it/prove/cosa-misurano-le-prove/> (accessed May 14, 2025).

as native if they have at least one parent born in Italy, regardless of the students' own place of birth. All other students are either classified as first-generation immigrants – those born abroad to non-Italian parents – or second-generation immigrants – those born in Italy to foreign parents¹⁸. Such classification allows for more detailed examination of the potential differential effects exerted by distinct immigrant groups.

Information on the place of birth of both students and their parents is grouped into four broad categories: Italy, European Union countries, non-EU European countries, and non-European countries. While no more granular data on the specific country of origin is available, this classification still allows for more detailed analysis of outcomes by macro-geographical area. If there were more precise information on the specific origin of non-native students, it would be possible to explore whether peer effects vary according to migrants' origin or to their native language. It is reasonable to think that non-native children whose first language belongs to the Romance languages may face fewer difficulties to acquire Italian, which becomes especially important if we believe any potentially adverse effect arising from immigrant concentration to be due to language barriers.

Finally, it's worth noting that INVALSI data also include school identifiers, class identifiers, and municipality codes, making it possible to control for school and other levels of fixed effects, and eventually to investigate the phenomenon of interest at different levels of aggregation.

4.2 Building the dataset and defining the sample

As it was suggested in the previous chapter, one major obstacle to research on long-term peer effects is the lack of longitudinal data that track students over time. INVALSI data are available in separate year-grade-subject datasets and do not follow students over their school life, but rather just capture the situation at a given point in time. Since 2013 however it became mandatory to register each student using the SIDI code, a unique identifier assigned to each student in the Italian education system by the Ministry of Education, which makes it possible to build a longitudinal dataset aggregating information from different years.

¹⁸ Such information are available at <https://www.invalsiopen.it/studenti-immigrati-inclusione/#:~:text=immigrati%20di%20prima%20generazione%20-%20nati,in%20Italia%20da%20genitori%20stranieri> (accessed April 22, 2025).

We retrieved data for the 5th grade in 2014-2015, the 8th grade in 2017-2018 and the 13th grade in 2022-2023, that capture the same cohort of students over time. These are the most recent available data that carry for each student the unique identifier. A total of 277.821 children were tracked across the three years using the SIDI code. It is evident that this sample of students contains only pupils who remained in public school and did not repeat any grade after the 5th. Students who moved to a private school, failed a grade, missed from school on the days of the INVALSI assessments or even moved out of the country are lost in the process of merging the different datasets. This aspect will be discussed in more detail when talking about the limitations of this research and of its external validity. In addition, the SIDI code was not available for all children, most likely due to data entry error¹⁹. Observations for which such information was not available were dropped from the various datasets prior to the merging process. Further information on the number of observations for which SIDI code was not available or that were lost during the merging process in each individual dataset are available in the Appendix A.4.2.

Linking student records over multiple INVALSI survey waves made it possible to construct a longitudinal panel. Each student in our dataset is observed at the 5th grade in primary school, at the 8th grade in lower secondary school and at the 13th grade in upper secondary school. Each wave of INVALSI carries information on the characteristics of pupils, on their socio-economic background, school outcomes and INVALSI test scores.

One major issue of INVALSI assessments is that they are not mandatory; in other words, data from the Institute refer only to the *test-takers* rather than to the whole population of Italian students. This aspect questions the representativeness of the sample and hence the external validity of any conclusion drawn from these data. This becomes especially an issue if the population of students missing from the INVALSI assessments is not distributed at random. There are concerns that poor-performing students might be systematically absent from these tests because of poor attitude towards education or even because they are suggested to do so. While INVALSI tests are not high-stakes, school principals and teachers may take into account matters of public image and accountability, and they may encourage children expected to perform poorly to miss the assessment, in a try to boost the attainment level of the classroom

¹⁹ Up to 2018 the task of entering data was entrusted to teachers. Since then, assessments are carried out on computers so that the risk of human error in reporting data is minimized. This information is available in the form “Le prove su computer” at <https://www.invalsiopen.it/area-prove/invalsi-secondo-invalsi-per-saperne-di-piu/> (accessed May 14, 2025).

and of the school. If lower-achieving pupils are systematically more likely to be absent from these tests, it would lead to a bias in the measure in which higher-achieving students would be overrepresented. Any estimate on average achievement levels or group comparisons could be biased upwards. This issue of attrition will be addressed carefully in Section 7. In addition, if academically weaker students, who are more likely overrepresented among immigrant groups, are systematically more likely to skip the tests, it would lead to an underestimation of the actual share of immigrant students in the peer group. As a consequence, we would fail to effectively measure peer composition and ultimately to capture the true impact of immigrant concentration on natives' outcomes.

For the scope of this research, we need to compute the share of immigrant students in each school at the 5th grade. To understand to what extent the coverage rate may constitute an issue for this analysis, we retrieve information available in the INVALSI official reports on the participation rates of sample classes at the region level.

The Institute publishes periodically reports containing detailed aggregate information, including participation rates by region. These reports usually refer only to *classi campione*, or sample classes. In each wave of INVALSI, the Institute selects casually a set of classrooms across the country to best represent the national population of students and schools. Since this selection of classes is designed to be an appropriate sample of the broader population of interest and of its characteristics, we assume the reported participation rates to be a reliable proxy for overall participation.

Figure 4.1. Participation rates of sample classes at the 5th grade in Italian regions in 2015.



Data source: INVALSI (2015), Rilevazioni Nazionali degli Apprendimenti 2014-2015, p.10.

To minimize the threat of lack of representativeness in our data, it is suitable to drop all observations from the regions whose participation rates in 5th grade are below 75%, namely: Campania, Puglia, Sardegna and Sicilia. The participation rates for these regions are available in Table 4.1.

Table 4.1. Participation rates in 2015 for regions with high abstention rates.

	Campania	Puglia	Sardegna	Sicilia
Participation rate	65,8%	74%	29,8%	71,5%

Data source: Invalsi (2015), rilevazioni nazionali degli apprendimenti 2014-2015, p.12.

The lowest coverages are registered in the regions of Sardegna and Campania. Deleting observations from these regions is an imperfect solution to guarantee that our findings are representative of the whole population of students in Italy, and not only to *test-takers*.

After dropping observations from regions with low participation rates, we were left with 216.414 observations. Table 4.2 shows the distribution of pupils according to their origin. Non-native students are further disaggregated in first-generation and second-generation immigrants.

Table 4.2. Distribution of children by origin in the final sample.

	Natives	All immigrants	Immigrants of I generation	Immigrants of II generation
Number of pupils	201.007	15.407	4.027	11.380
Share over the population	92.88 %	7.12 %	1,86 %	5.26 %

The share of immigrants over the total population appears to be relatively small; however, it is in line with the number of non-native students found in other research on peer effects arising from immigrant concentration, such as those by Gould et al. (2009), Contini (2013) and Seah (2021).

On the 31st of December 2015, 8.3% of residents in Italy had foreign citizenship (Istat, 2016) and 10.3% of students enrolled in primary school in Italy in 2014-2015 were non-native (MIUR, 2017). The share of immigrants we observe is just below this figure from MIUR. Nevertheless, the distribution in Table 4.2 refers only to the students we observe in all the three grades. In the full dataset on 5th grade test-takers, 10% of students were non-native. We conclude that our measure of immigrant concentration in primary school – which we compute before the merging process – is effectively capturing the exposure of native children to non-native peers. We do

not find evidence of immigrant students being disproportionately absent – compared to native children – from the INVALSI assessments under exam.

Also, the final sample contains 11.185 unique schools²⁰.

Table 4.3 shows the share of immigrant students in each region in the final sample. The autonomous province of Bolzano presents an unusually large share of non-native pupils with respect to the other regions, especially when counting only second-generation immigrants. This is most likely due to its strong economy and job opportunities, especially in tourism and industry. The last column of the table also reports the Duncan index at the regional level, that measures the extent to which native and immigrant students are unevenly distributed across schools in our dataset. This index ranges between 0 and 1, where values closer to zero indicate low segregation while values closer to 1 indicate higher segregation. Interestingly, the regions with the highest concentration of immigrants are those where the non-natives are less segregated, whereas the regions with less immigrants show the highest values of the Duncan index. It may be that, in regions with more immigrants, schools adapt by becoming more inclusive and diverse, whereas in regions with less immigrants there are fewer incentives for dispersion so that non-native pupils concentrate in a limited number of schools. Besides that, social networks, housing patterns and local policies also play a crucial role in shaping the distribution of children across schools.

²⁰ These are identified as *plessi* in the Invalsi dataset. One school can be made of multiple *plessi*, which are basically the physical building in which lessons take place and where children interact daily.

Table 4.3. Share of immigrant students and Duncan index disaggregated by region.

	Share of immigrant students* (%)	Share of I generation immigrant students* (%)	Share of II generation immigrant students* (%)	Duncan index (0-1)
Prov. Aut. Bolzano (l. it.)	18.26	6.25	12.01	0.27
Umbria	9.52	2.14	7.37	0.36
Emilia-Romagna	8.78	2.36	6.42	0.38
Lombardia	8.19	2.06	6.13	0.44
Piemonte	8.05	1.97	6.07	0.49
Prov. Aut. Trento	7.85	2.10	5.75	0.49
Friuli-Venezia Giulia	7.69	2.06	5.63	0.46
Veneto	7.67	1.80	5.87	0.43
Marche	6.81	1.97	4.85	0.42
Toscana	6.66	1.87	4.78	0.45
Lazio	5.99	1.49	4.50	0.47
Liguria	5.80	1.87	3.93	0,53
Abruzzo	4.54	1.31	3.23	0.5
Valle d'Aosta	3.05	1.53	1.53	0.54
Molise	2.90	1.13	1.78	0.62
Calabria	1.95	0.90	1.05	0.69
Basilicata	1.49	0.59	0.90	0.72

Note: *all shares are computed over the final sample and are presented in percentage.

Table 4.4 and Table 4.5 show the distribution of children by place of birth, separately for native and first-generation immigrant students. Second-generation immigrant students are excluded because, by definition, they are born in Italy. The share of students for which the place of birth is not available is also shown in the last column. We observe that there are very few children who are born abroad by an in Italian parent. As for first-generation immigrants, most of them come from non-specified countries out of Europe.

Table 4.4. Shares of native students in the sample by place of birth.

	Natives				
	Italy (or Republic of San Marino)	UE	non-UE European country	Other	Not available
Share of pupils	98.93 %	0.31 %	0.23 %	0.51 %	0.01 %

Table 4.5. Share of first-generation immigrant students in the sample by place of birth.

	Immigrants of I generation				
	Italy (or Republic of San Marino)	UE	non-UE European country	Other	Not available
Share of pupils	-	33.80 %	24.16 %	42.02 %	0.02 %

4.3 Key variables and descriptive statistics

This section defines some of the key variables used in the analysis and descriptive statistics to provide an overview of the data distribution and more specific sample characteristics. The most important variables are those that identify students depending on their origin. The first one, *imm*, is a dummy variable that takes value 1 when the student is of foreign origin, irrespective of whether they are first- or second-generation immigrants. This is the key variable in this analysis, and it is employed to compute the share of non-native children. Two additional dummies, *imm_I* and *imm_II*, that flag respectively first- and second-generation immigrant students. These variables make it possible to further discern whether these two categories on non-native children have different impact on native students, investigating closely the potential heterogeneity in peer effects arising from early exposure to immigrant peers. If one expects the limited language proficiency of non-native children to be the main driver of adverse peer effects, it is likely that second-generation immigrants, who grew up in the host country and had time to master the mainstream language, produce little to no detrimental effect on native children. Table 4.6 below shows selected summary statistics for the final sample.

Table 4.6. Selected summary statistics for the sample.

	Full sample (0-1)	Natives (0-1)	Immigrants (0-1)		
			All immigrants	I generation	II generation
Who are males	0.467	0.471	0.414	0.412	0.415
With highly educated father	0.185	0.190	0.122	0.121	0.123
With highly educated mother	0.236	0.242	0.154	0.165	0.150
With employed father	0.887	0.894	0.788	0.733	0.807
With employed mother	0.659	0.680	0.381	0.323	0.401
Who are a grade behind	0.012	0.006	0.09	0.259	0.030
Who are a grade ahead	0.007	0.007	0.005	0.007	0.005
Who attended nursery	0.273	0.274	0.264	0.225	0.278
Who attended kindergarten	0.825	0.830	0.757	0.646	0.796

We consider a parent to be highly educated whenever they possess a degree, a PhD, or any other qualification above diploma (e.g. Academy of Fine Arts, conservatory). Immigrant children systematically have parents with lower education. The gap with respect to native children is especially large when consider the level of education of the mother, most likely because of structural, cultural and socio-economic barriers that limit women's access to education in certain countries²¹. Factors such as gender inequality in educational opportunities, lower household's investment in girls' schooling, weak educational infrastructure, political instability and poverty lead to a systematic educational gap for women which translates into lower average

²¹ Additional information on women's education around the world are available at <https://www.worldbank.org/en/topic/girlseducation> (accessed May 16, 2025) and at <https://www.unicef.org/education/girls-education> (accessed May 16,2025).

education level among immigrant mothers. A similar pattern is recognized for the employment status of the parents. In particular for women, their lower educational attainment limits their job opportunities along with the influence of cultural norms and gender roles. We excluded stay-at-home parents for the count of employed parents, as they do not earn income.

A large portion of first-generation immigrants, the 25,9%, is a grade behind with respect to the one they should be assigned to; it is often common practice to place immigrant students, especially those recently arrived in the country, in a grade prior to the one they should be assigned based on their year of birth to improve their chances of adapting to the Italian school system. A similar pattern was observed also by Gould et al. (2009). Such policy is frequently used to give them additional time to learn the mainstream language, adjust to new teaching methods and fill any possible gap in their education. Grade repetition is considered a way to support their integration and reach a more solid educational foundation before continuing their studies. The share of children who are ahead is relatively small in every subgroup of the sample.

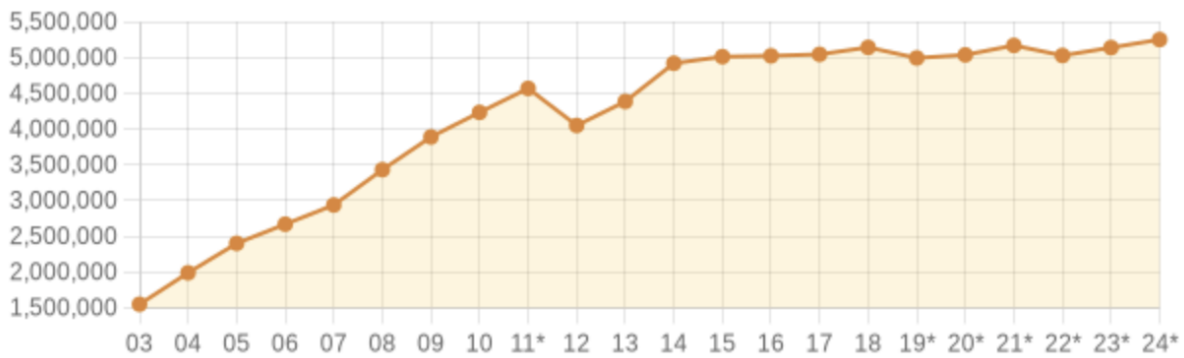
Interestingly, the share of children who attended nursery and kindergarten is similar in every segment of the sample that we are considering. Still, the lowest value is observed again for first-generation immigrant students, whose average age of arrival in Italy computed in this subgroup is 4.26 with a standard deviation of 0.05.

It is important to mention, for the sake of transparency, that there is a large portion of missing values in the variables related to nursery and kindergarten attendance; these figures shall be examined accordingly, in particular in the distribution of missing values itself is a driver of information.

4.4 Immigrant children in Italy

Over the past decades, developed countries such as Italy have witnessed a dramatic increase in the number of incoming migrants which lead to a rise in the enrollment shares of immigrant students and the subsequent change in the composition of classrooms and schools in the host countries. This is the reason of the growing interest towards understanding what the impact of a greater concentration of immigrant pupils on the performance of native students can be. Figure 4.2 shows how the population of Italian residents with foreign citizenship evolved over time. Starting from 2003, there is a constantly increasing trend up to 2011. Over the span of 10 years the population of migrants in Italy basically triplicated. After a marginal drop in 2012, the community of non-Italian residents remained overall stable.

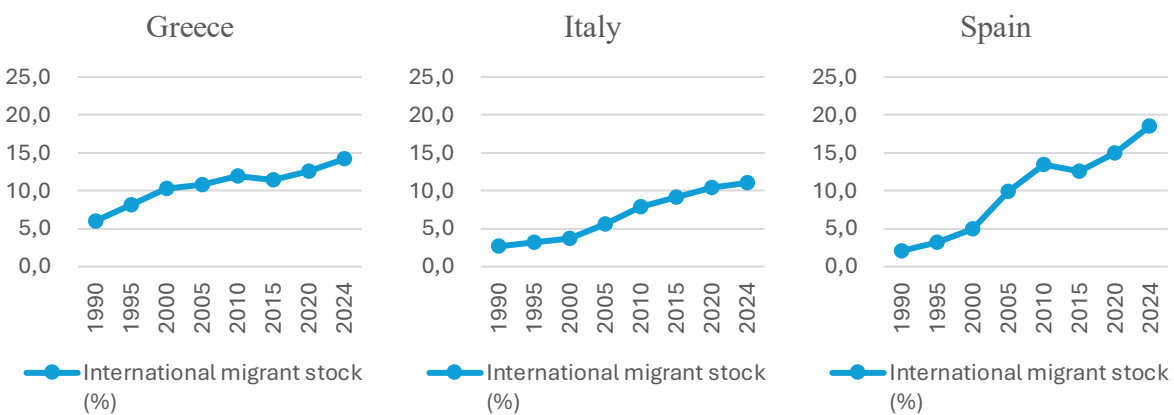
Figure 4.2. Population in Italy with foreign citizenship.



Data source: <https://www.tuttitalia.it/statistiche/cittadini-stranieri-2024/>.

What makes the Italian case particularly interesting is that – unlikely other major European countries – significant migration flows began relatively recently. Figure 4.3 and Figure 4.4 illustrate the evolution of the international migrant stock as a percentage of the total population starting from 1990 onwards in a selection of European countries. Nations such as France, Germany and the United Kingdom have historically experienced higher and more sustained migration flows, likely making them accustomed to managing ethnic diversity in schools, workplaces, and society at large. Conversely, southern European countries maintained relatively low migrant stocks during the 90s. In particular, Italy’s share stood just at 2,1% of total population, which quintuplicated over the span of 10 years up to 2024.

Figure 4.3. International migrant stock from 1990 to 2024 as percentage of total population in Greece, Italy and Spain.



Data source: United Nations

(<https://www.un.org/development/desa/pd/content/international-migrant-stock>).

Figure 4.4. International migrant stock from 1990 to 2024 as percentage of total population in France, Germany and United Kingdom.



Datasource: United Nations

(<https://www.un.org/development/desa/pd/content/international-migrant-stock>).

Consistent migration flows are a relatively new phenomenon for Italy. Today, there are 5.422.426 non-native resident in Italy²², around 9,2% of total population. 58,3% of total residents of foreign origin are located in north Italy, while south Italy welcomes only 12,4% of them. According to Istat, the largest incoming migration flows in 2024 came from Romania, Albania, Morocco and China²³.

²² This information are available in the ISTAT database at https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_FOREIGNIM/DCIS_POPSTRRES1/IT1,29_7_DF_DCIS_POPSTRRES1_1,1.0 (accessed May 21, 2025).

²³ Information retrieved from https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_FOREIGNIM/DCIS_POPSTRCIT1/IT1,29_317_DF_DCIS_POPSTRCIT1_1,1.0 (accessed May 21, 2025).

Figure 4.5 shows the distribution of non-native residents across regions in 2015 according to Istat. We can easily observe that regions in the northern area of Italy are characterized by a stronger presence of non-native residents, most likely due to their economic prosperity, job opportunities and even educational quality.

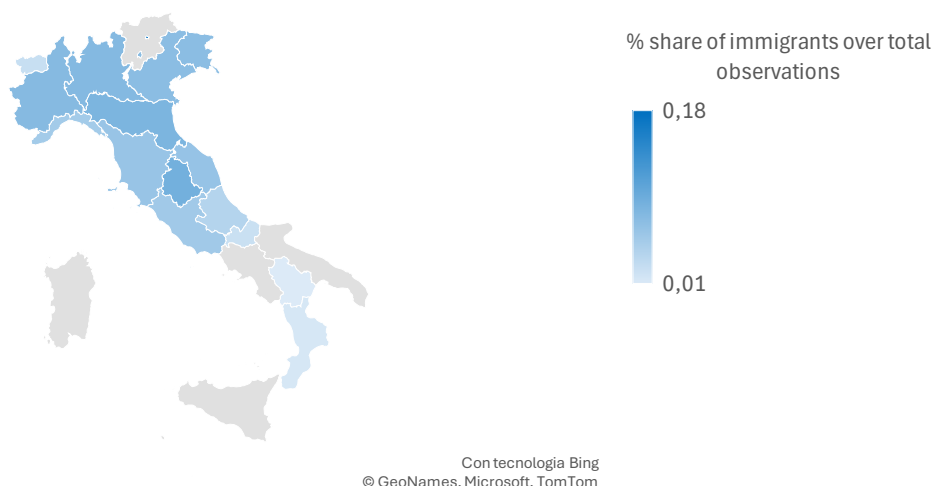
Figure 4.5. Distribution of foreign residents across Italian regions in 2015.



Note: white areas count from 9.082 to 50.063 migrant residents; light blue areas from 50.063 to 384.241; dark blue areas from 384.241 to 1.124.292. Data source: Istat database(https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,POP,1.0/POP_IN_TCENSPOP/DCIS_RICPOPRES2011/IT1,164_164_DF_DCIS_RICPOPRES2011_1,1.0)

These data strengthen again the idea that the dataset in use for this research correctly captures the distribution and concentration of immigrants across Italy. Figure 4.6 shows how non-native students in our sample are allocated across Italian regions in 2015. It is possible to observe the same pattern seen in Figure 4.5. The only difference is for the region of Lazio, that appears to have a strong immigrant presence in the data from Istat but not as much in our dataset; however according to Figure 4.1, this is the only region with coverage rate slightly inferior to the other areas (between 75% and 90%). The discrepancy observed between data by Istat and by INVALSI is most likely due to lower participation rate in this region in 2015.

Figure 4.6. Distribution of immigrant concentration across Italian regions in 2015 in the dataset retrieved from INVALSI data.



Note: We do not show information on the share of immigrants in the areas in grey. Campania, Puglia, Sardegna and Sicilia were purposefully left out of the final sample. In the region of Trentino Alto Adige we have information only in the provinces of Bolzano and Trento.

4.5 The performance of students in Italian schools

Our objective is to understand the impact that immigrant concentration has on native students' school performance. Those who argue that the concentration of immigrant students is detrimental to the performance of native students usually emphasize the challenges posed by increased classroom heterogeneity. Such perspective highlights that immigrant pupils, especially those with limited proficiency in the host country's language or who come from a more disadvantaged socioeconomic background, might struggle to keep pace with the rest of the classroom. Critics believe that this may complicate teachers' tasks and force them to devote additional time to these students, thus diverting attention from the rest of the class. On the contrary, those who claim that immigrant concentration can have positive effects believe that immigrant families are often positively selected, meaning that they tend to be more motivated or place higher value to education, contributing to a more achievement-oriented classroom and stimulating natives' curiosity and adaptability.

Before proceeding to the empirical analysis it is essential to briefly discuss students' achievement in the INVALSI assessments.

Table 4.7 shows average INVALSI test scores in mathematics and reading for the 5th grade, for both the entire sample and separately for native and immigrant students. On average, native pupils perform better than their non-native peers in both subjects. The performance gap is more

pronounced in reading, where the difference in average scores reached approximately 17 points, compared to about 12 points in mathematics. These findings reinforce the hypothesis that immigrant students may face additional challenges with the mainstream language, which likely affects their performance across all subjects taught in Italian. Interestingly, while native students achieve similar results in the two subjects, immigrant pupils perform slightly better in mathematics than in reading, most likely due to reduced language dependency in mathematical reasoning tasks.

Table 4.7. Average INVALSI assessment scores at the 5th grade.

5th grade scores in:	All children				Natives		Immigrants	
	Mean	S.D.	Max	Min	Mean	S.D.	Mean	S.D.
Mathematics	216.6	39.6	364.7	14.6	217.5	39.5	205.5	39.2
Reading	216.9	38.8	392.9	-23.4	218.0	38.6	201.0	38.8

Table 4.8 presents the same statistics for the 8th grade, where English proficiency assessments in reading and listening are also administered. It is interesting to notice that at this level immigrant children outperform their native peers in English, particularly in the listening component, where their average scores exceed that of natives by 6,3 points. This result is particularly relevant, as English represents a foreign language for both native and non-native students, placing them on a more equal starting point compared to Italian language, where Italian native students hold an obvious advantage. The stronger performance of non-native students in English may indicate that they are in fact school-oriented and academically motivated, contradicting the assumption that immigrant concentration necessarily lowers overall classroom achievement. In mathematics and reading, the same pattern observed in 5th grade persists, with native students achieving higher scores.

Table 4.8. INVALSI assessment scores at the 8th grade.

8th grade scores in:	All children				Natives		Immigrants	
	Mean	S.D.	Max	Min	Mean	S.D.	Mean	S.D.
Mathematics	216.3	37.4	374.2	71.7	217.0	37.4	207.6	35.9
Reading	216.0	34.8	353.6	22.7	217.1	34.6	201.6	33.9
English reading proficiency	216.2	34.8	316.0	53.4	216.1	35.0	216.7	34.5
English listening proficiency	214.2	35.8	323.6	75.8	213.8	35.8	220.1	35.7

Table 4.9 displays results for the 13th grade, confirming the trends observed in the earlier grades. Immigrant children outperform natives in English tests, while native students get better results in mathematics and reading. Across the board, reading test scores are the lowest for both grounds, with average performances declining more steeply in this subject than in others overtime. Overall, evidence suggests a consistent achievement gap between native and immigrant students in subjects taught in Italian, due to language barriers. Conversely, immigrant students perform better in English, where the initial advantage of native students is neutralized, supporting the idea that immigrant students are not necessarily low-achieving and that, when language related disadvantages are reduced, they can perform at or above the level of their native peers.

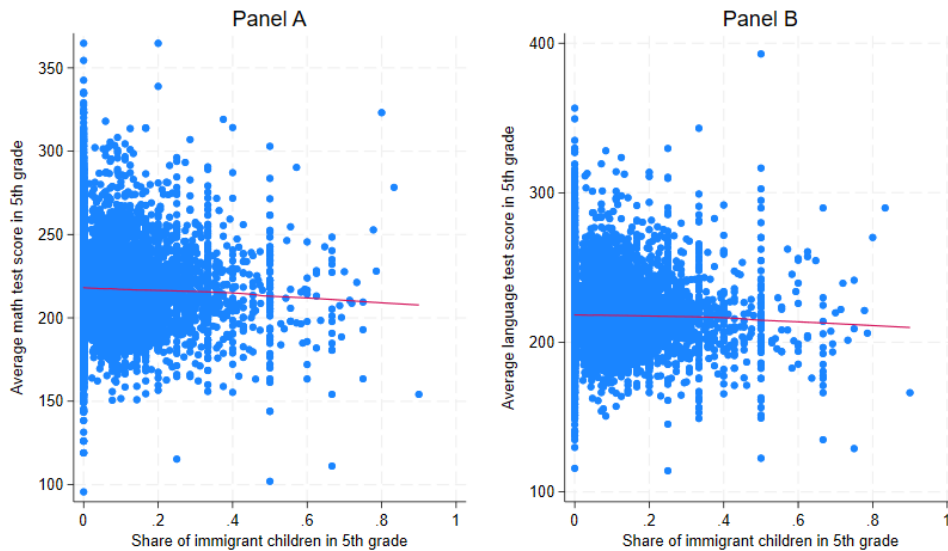
Table 4.9. INVALSI assessment scores at the 13th grade.

13th grade scores in:	All children				Natives		Immigrants	
	Mean	S.D.	Max	Min	Mean	S.D.	Mean	S.D.
Mathematics	203.7	37.7	330.2	72.5	204.3	37.8	196.2	35.9
Reading	197.4	38	329.9	2.2	198.3	38.0	185.7	35.5
English reading proficiency	212.6	38.4	286.8	95.8	212.5	38.5	214.0	37.2
English listening proficiency	219.1	38.6	303.0	108.0	218.8	38.7	223.0	37.1

To gain a preliminary understanding of how immigrant concentration relates to students' achievements, it is useful to visualize the association between native students' average test scores and the share of immigrant pupils in 5th grade computed at the school level. The figures below display results from non-parametric regressions using *lowess*, illustrating the relationship between school-level average scores of native students and the percentage of immigrant pupils in 5th grade. Using *lowess* non-parametric regression is preferable because it flexibly captures potential nonlinearities in the relationship between immigrant concentration and test scores without imposing a predetermined functional form, thereby providing a more accurate and data-driven representation. Figure 4.7 illustrates how native students' performance in mathematics and reading INVALSI assessment in 5th grade varies according to the share of immigrant peers enrolled in their school. Panel A depicts the association between the average mathematic test score of native students in 5th grade INVALSI assessment and the concentration of immigrant at school level in the same grade; Panel B uses school-level average scores in reading test of native pupils. Both graphs show a mostly flat relationship, suggesting that while native students

perform slightly worst in both subjects in schools with a higher share of non-native children, there is no sizeable visible relationship.

Figure 4.7. Relationship between school average test scores of natives in 5th grade and percentage of immigrant students in the school in 5th grade.



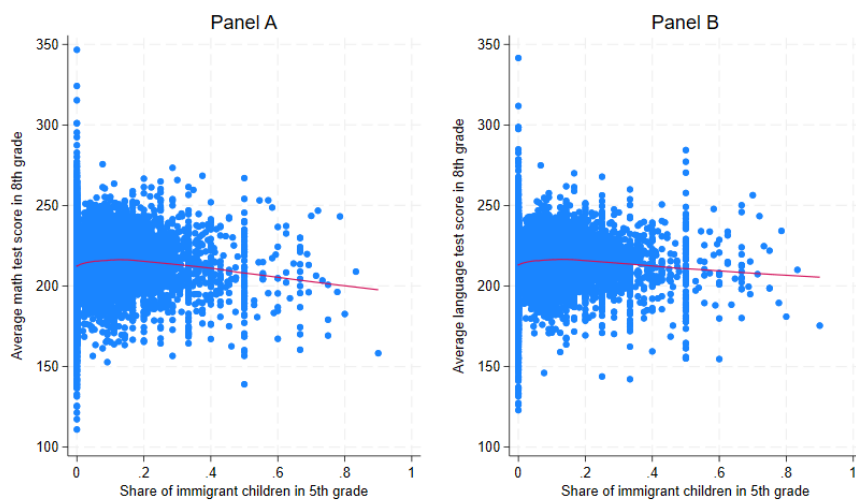
Note: figure A reports the association using average score in math test; Figure B using average score in reading (language) test.

Figure 4.8 and Figure 4.9 shows the relationship between test scores of native students in 8th grade and the concentration of immigrant children in 5th grade. Note that, at each grade, we will use as treatment for our analysis the share of non-native schoolmates computed in primary school, to capture whether any effect or early immigrant exposure lasts up to the end of upper secondary school.

The trend depicted in this graph is quite more interesting. The fitted line in panel A of Figure 4.8 suggests a non-linear relationship where schools with a moderate share of immigrant students perform better in mathematics assessment than both schools with very low and very high immigrant concentrations, with the latter showing noticeably lower average test scores. Such a pattern with inverse-U shape may imply that an intermediate level of immigrant presence may foster positive effects, while very high concentrations might be associated with the challenges that negatively impact performances. Panel B of Figure 4.8 display a similar trend for average reading test scores, though the gap in performance between schools with very low immigrant concentration and schools with a very high one is smaller than the one observed in panel A. Panel A and B of Figure 4.9 shows the relationship between average test scores in English reading and listening assessments and immigrant concentration. According to these graphs, native students who attended elementary school with very low or very high share of

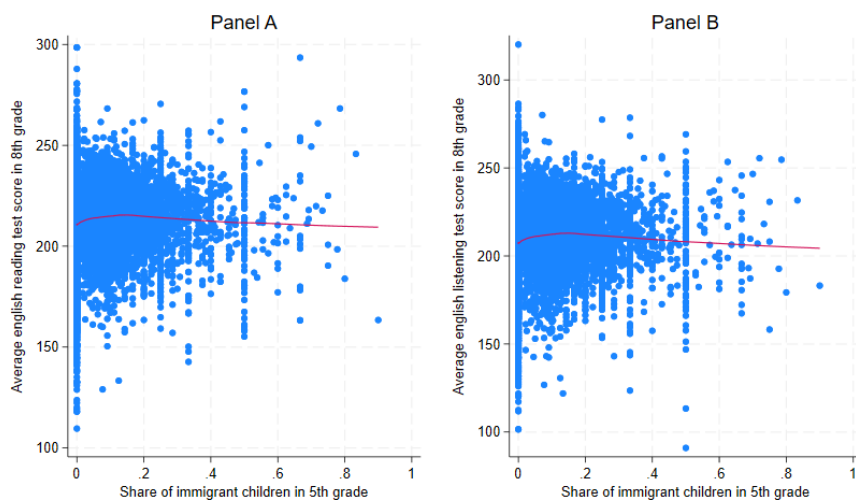
non-native children perform similarly in 8th grade, whereas those who shared 5th grade with a moderate fraction of immigrants perform slightly better. These figures suggest that native pupils who were enrolled in elementary school with a modest fraction of immigrant schoolmates appear to have benefited from the presence of non-native peers. Conversely, pupils who attended 5th grade in schools with a high concentration of immigrant children perform worse than the other students across all subjects, with the largest performance gap observed in mathematics and reading.

Figure 4.8. Relationship between school average test scores of natives in 8th grade and percentage of immigrant students in the school in 5th grade.



Note: figure A reports the association using average score in math test; Figure B using average score in reading (language) test.

Figure 4.9. Relationship between school average test scores of natives in 8th grade and percentage of immigrant students in the school in 5th grade.

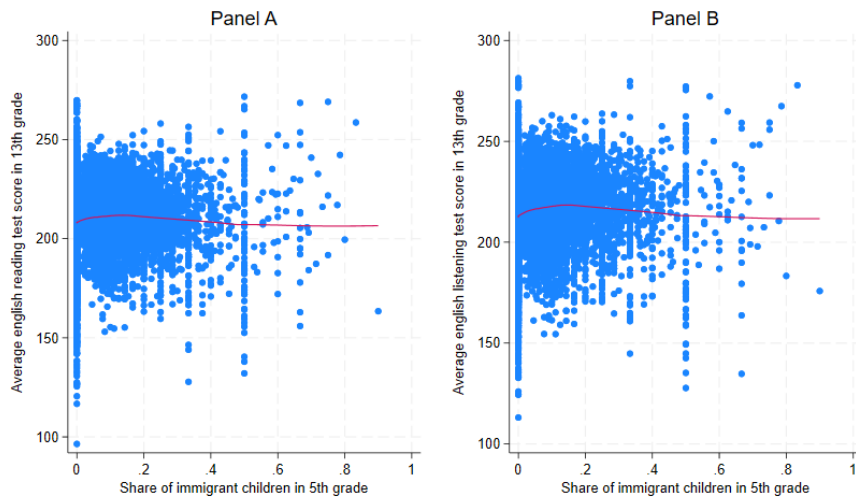


Note: figure A reports the association using average score in English reading test; Figure B using average score in English listening.

Figure 4.10 and Figure 4.11 depict the relationship between performance in 13th grade and the share of immigrant schoolmates in 5th grade. The same pattern observed for 8th grade results can be seen for upper secondary school results. Native pupils who attended 5th grade with a moderate immigrant concentration perform better than their native peers who attended schools with very low immigrant share, whereas native students who experienced higher immigrant concentration at elementary school perform slightly worse.

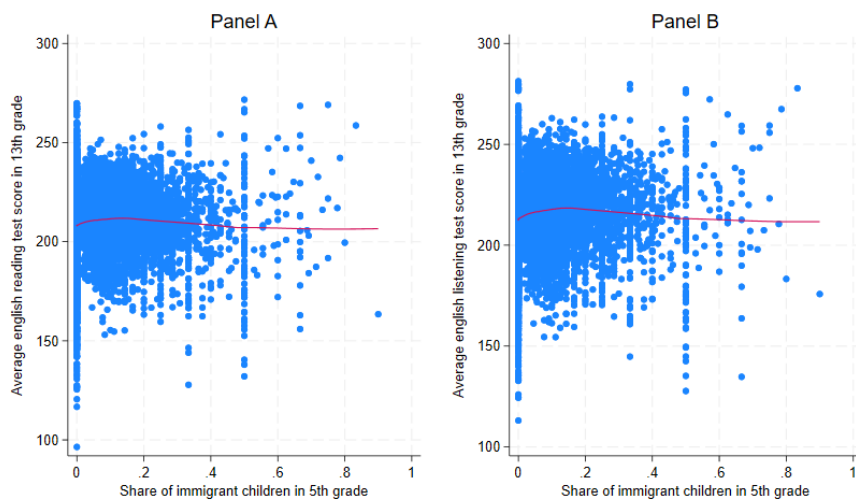
While these figures offer an interesting view of our data, it is not possible to draw any conclusion. Indeed, the non-parametric *lowess* regression only capture the raw relationship between two variables of interest, and it does not control for any potential confounding factor like socio-economic background, parental education, prior achievement or school characteristics. It is very likely that the pattern observed is driven by omitted variables. Above all, it is probable that this apparent association between 5th grade immigrant concentration and weakly worst performance in 8th and 13th grade is the result of unobserved socio-economic background characteristics. As outlined in the previous sections, children are not randomly distributed across schools. It is plausible that natives from a poorer socio-economic background and immigrant families cluster in the same areas because of economic constraints and job opportunities, hence we observe the negative relation between high immigrant share in primary school and test scores in later grades. Graphics that draw separately the correlation between test scores and shares of first- and second-generation immigrants are available in the appendix A.4.5. As expectable, when making the distinction between first-generation and second-generation immigrants, we find that children in schools with higher shares of first-generation immigrants perform slightly worse than their native peers in school that have higher shares of second-generation immigrants. Indeed, it is plausible that children born in Italy by foreign parents have better knowledge of the host-country language so that any possible negative effect arising from friction between native and non-native children are eventually cancelled out.

Figure 4.10. Relationship between school average test scores of natives in 13th grade and percentage of immigrant students in the school in 5th grade.



Note: figure A reports the association using average score in math test; Figure B using average score in reading (language) test.

Figure 4.11. Relationship between school average test scores of natives in 13th grade and percentage of immigrant students in the school in 5th grade.



Note: figure A reports the association using average score in English reading test; Figure B using average score in English listening test.

4.6 Data limitations

Despite INVALSI offering rich data and being an important source of information for social research, some limitations must be acknowledged for the sake of transparency and to determine the degree of external validity of any result that will be found.

First of all, as addressed earlier in this section, INVALSI data refer to *test-takers* rather than to the whole population of students enrolled in Italian schools, which can make it difficult to generalize findings. The consequences of this are especially amplified if we assume that the population of students who forgo the tests is not randomly distributed. That is to say, if lower-achieving students, immigrant students, or any other specific subgroup of children are systematically absent from INVALSI assessments, any result drawn from the empirical analysis of these data will be biased and not representative of the behavior of the whole population of students. Moreover, the nature of this data allows us to compute only the share of immigrant *test-takers* rather than the share of immigrants among enrolled students, which is again especially important if we expect immigrant students to be worse-performing or less-school oriented than native pupils and more likely to stay at home on the day that INVALSI tests are distributed. The latter, however, is not a significant concern in this analysis, as we observed that the share of immigrant *test-takers* available in the dataset is basically equal to the share of immigrant students enrolled in schools in Italy as measured by the MIUR.

To address the issue arising from potentially observing only a portion of the actual population of students in Italy, it was decided to drop any observation from regions who registered very low participation rates in the INVALSI tests in 5th grade in 2014-2015, as the exposure to immigrant concentration will be computed at this level. Further information on this procedure were provided earlier in this section. All other regions from which the observations are drawn have participation rates above 90%, except for Lazio that is between 75% and 90%, so that we minimize any risk coming from not observing students with specific characteristics or socio-economic backgrounds that make them more likely to miss the tests. By considering only regions with a sizeable coverage we hope to compute capture effectively the actual population of students in Italy.

Another major limit of this dataset is the so called *dispersione scolastica*²⁴, that refers to the fact that we are observing only students who stayed public school up to the end of upper secondary school, who never failed a grade and had to repeat a year or never missed schools on the days that INVALSI tests are administered. That is because of how the specific way in which the panel was prepared, as there were no longitudinal datasets available with the information needed for this analysis on peer effects. Only students that were present in each wave of INVALSI in the 5th grade in 2014-2015, 8th grade in 2017-2018 and 13th grade in 2022-2023 were connected through their SIDI code and kept in the final sample. All other observations were lost in the process. Because of this, we are potentially failing to observe the group of worse-performing students who have higher chances of failing a grade in school or dropping out after concluding compulsory schooling. Dropping-out rates or the share of students who are held back a grade are meaningful outcomes when it comes to understand the long-term school performances. Due to the nature of our data, we will not be able to draw any conclusion on the impact of being exposed to higher immigrant concentration in primary school on dropout rates and likelihood of grade repetition. This is especially concerning in the case of heterogeneous peer effects. If, for instance, immigrant concentration affects to a greater degree low-performing students, to the point that they are more likely to drop out of school or be held back a grade but has no significant effect on better-performing students, we would most likely underestimate any adverse effect exerted by immigrant students on their native peers. Ideally it would be useful to have a detailed well-built longitudinal dataset that tracks students over their whole school career – and not only at specific grades as INVALSI does – with eventually insights on dropping rates. While it appears that there is no effective solution to such limited longitudinal depth, this aspect must be taken into account when considering any result coming from the analysis. Hopefully this is just a marginal issue, as according to the statistics office of the *Ministero dell’Istruzione e del Merito*²⁵, dropout rates in upper education have been decreasing in the past few years. For instance, between school year 2020/2021 and 2021/2022 – right

²⁴ The Italian term for school dropouts. The term *dispersione* however make it clear that some students are “lost” over time, and we have no information on the reasons why this happens. It may be that they voluntarily drop out because of poor school attainment; children from an extremely poor background may be forced to leave after finishing compulsory school to find a job; less academically able pupils are likely to be held back a grade at some point during their schooling career.

²⁵ The Italian Ministry responsible for the education and instruction of the country’s youth, with a special reference to the public school system.

before our sample of students attended 13th grade – only 2,55% of the total students who had previously enrolled in school left with no specific reason (Ufficio di Statistica, 2023). In addition, they report that the higher dropout rates are historically registered in the southern part of Italy; observations from these regions were previously dropped from our sample. Nevertheless, to ensure that this attrition impacts as little as possible our estimates, we will perform an additional robustness check in Section 7, by weighting each student in the sample by the probability of missing from 8th grade and 13th grade observations.

A further important drawback of the dataset is the absence of information regarding the specific country of origin or mother tongue of immigrant students. Such details would be highly valuable for investigating whether the specific ethnic or cultural backgrounds of immigrant pupils play a role in the determination of peer effects. Indeed, two classrooms of identical size and with the same proportion of immigrants could display very different social dynamics depending on the internal composition of the non-native group. If all immigrant children in a class share the same native language and similar cultural background, they are more likely to form cohesive social networks within the classroom. This would reduce both their incentive and necessity to engage with their native peers or with the host country's language and culture. There is evidence that individuals tend to form personal networks with others similar in terms of sociodemographic and behavioral characteristics. McPherson et al. (2001) argue that race and ethnicity create some of the strongest social divisions, further reinforced by factors such as religion, language, and geographical proximity. Kao & Joyner (2004) using data from the United States show that friendships among adolescents occur predominantly within the same ethnic group. This propensity is often explained by greater social ease, shared cultural references, reduced language barriers and fewer cultural misunderstandings. Consequently, without detailed data on the ethnic and linguistic composition of immigrant groups within schools, it becomes challenging to fully capture heterogeneous peer effects in this framework. The social integration and peer influence dynamics are likely to differ substantially between a group of immigrants from diverse backgrounds and one composed of children from the same country and linguistic community. No specific solution can be implemented to address this issue with the available dataset, and it remains a limitation of the present analysis. Hopefully, future data collections will provide more detailed information on students' country of origin and linguistic background, allowing for a more nuanced and comprehensive investigation of this important dimension of peer effects.

Some additional issues of INVALSI data may be due to human-based mistakes, as up to 2018 teachers themselves were mandated to correct and input children answers to the tests and

personal information on their socio-economic background²⁶ on computers to submit them to the Institute responsible of INVALSI assessments. Moreover, there may be some bias due to these tests being not high stakes, meaning that students may not take them seriously. In this case, the score they obtain at the test would not be a true measure of their knowledge and skills.

²⁶ More information on computer-based INVALSI are available at https://www.invalsiopen.it/area-prove/invalsi-secondo-invalsi-per-saperne-di-piu/?_gl=1*_mgsalw*_ga*ODM0MTk3ODkxLjE3NDUzMTA3MTU.*_ga_5LBB0NMXR*czE3NDcyMTUxOTcjbzgzZzEkdDE3NDcyMTU0MDYkajAkbDAkaDA (accessed May 18, 2025).

5 Empirical strategy

5.1 Theoretical framework and challenges in identifying peer effects

The main purpose of this research is to investigate whether peer effects associated with immigrant concentration in elementary school persist over time. In other words, it aims to estimate the effect of immigrant presence among schoolmates in 5th grade on students' outcomes in 5th, 8th, and 13th grades, and to assess how the magnitude and direction of these effects eventually evolve over time. Any reasoning builds on an education production function of the form:

$$y_{ist} = \beta_0 + \beta_1 M_{cs} + \beta_2 X'_i + \varepsilon_{ti} \text{ with } t = 5, 8, 13$$

where y_{ist} represents the outcome of interest (e.g. test scores) of student i at time t who attended primary school s in 5th grade, X'_i is a vector of covariates that capture family background and other characteristics of individual i ²⁷, M_{cs} is the share of non-native students in class c in school s measured in 5th grade and ε_{it} is the standard error term. As highlighted earlier in this work a central challenge in estimating peer effects lies in the potential issues generated by omitted variables. If we fail to do comparison *ceteris paribus*, differences in the explanatory variable of interest will most likely capture other aspects that we are not observing. There exists a number of factors that we cannot control for, such as family size, parents' income, class size (this is indirectly and partially controlled for because we use the *share* of immigrants over school population), specific country of origin, that can affect students' performance and attainment in school, as well as their choice of studies at the beginning of tracking²⁸. By failing to control for similar aspects, our estimates of the coefficient of interest will pick up differences among students in our sample that are not due to the immigrant concentration to which they were exposed in 5th grade. Suppose, for instance, that family size FS_i affects the school performance of students in our sample; however, we have no information on the number of components in each household, so that the regression we estimate is:

$$y_{ist} = \alpha + \beta M_{cs} + \gamma X'_i + v_{it} \text{ with } v_{it} = \varepsilon_{it} + \delta FS_i$$

²⁷ These characteristics are assumed to be time-invariant, hence why X'_i is not time indexed.

²⁸ Tracking in Italian schools happen at the end of 8th grade. Children are asked to choose the type of school they want to attend at upper secondary level and the main subjects they want to study.

The OLS estimate of β is:

$$\hat{\beta} = \frac{cov(y_{ist}, \tilde{M}_{cs})}{V(\tilde{M}_{cs})}$$

Where \tilde{M}_{cs} is the residual of the regression of M_{cs} on X'_{cs} ²⁹. With some mathematics it can be shown that:

$$\hat{\beta} = \beta_R * \frac{cov(M_{cs}, \tilde{M}_{cs})}{V(\tilde{M}_{cs})} + \gamma_U * \frac{cov(FS_i, \tilde{M}_{cs})}{V(\tilde{M}_{cs})} = \beta_R + \gamma * \lambda_D$$

Where β_R is the true coefficient of interest, γ is the coefficient of FS_i and λ_D represents the relationship between FS_i and M_{cs} . The composite term $\gamma * \lambda_D$ represents the size of the omitted variable bias derived by failing to control for family size.

While it is important to adopt specific strategies to minimize the size of OVB and obtain more reliable measures of the coefficient of interest, we can justify that the issue of biased estimates in our framework is likely to be attenuated. The focus of this research is on comparing the effect of immigrant concentration in 5th grade on students' performance at different educational stages. Essentially, this involves estimating three separate regressions³⁰:

$$y_{5is} = \alpha_0 + \alpha_1 \tilde{M}_{cs} + \varepsilon_{5i}$$

$$y_{8is} = \pi_0 + \pi_1 \tilde{M}_{cs} + \varepsilon_{8i}$$

$$y_{13is} = \delta_0 + \delta_1 \tilde{M}_{cs} + \varepsilon_{13i}$$

In order to assess how peer effects evolve over time we compare the OLS estimates of the coefficient of interest in different regressions.

²⁹ This formula comes from the Frisch-Waugh-Lovell theorem, which shows that the coefficient of M_{cs} in a multivariate regression of y_{it} on M_{cs} and X'_i can be obtained by first regressing M_{cs} on X'_i , collecting the residual \tilde{M}_{cs} , and then regressing y_{it} on \tilde{M}_{cs} . This works because the effect of M_{cs} on y_{it} is identified from the variation in M_{cs} that is orthogonal to the controls – ensuring a ceteris paribus comparison. See Angrist & Pischke (2014) p.89.

³⁰ For simplicity we write these regressions in their partialled out form, where \tilde{M}_{cs} is the residual of the regression of M_{cs} on X'_i . This is possible as the set of covariates is assumed to be time-invariant.

Suppose we want to understand how the impact of immigrant concentration changes between 5th and 8th grades. We will compare $\hat{\alpha}_1$ and $\hat{\pi}_1$:

$$\hat{\pi}_1 - \hat{\alpha}_1 = \pi_1 - \alpha_1 + \frac{\text{cov}(e_{i8}, \tilde{M}_{CS})}{V(\tilde{M}_{CS})} - \frac{\text{cov}(e_{i5}, \tilde{M}_{CS})}{V(\tilde{M}_{CS})}$$

If $\frac{\text{cov}(e_{i8}, \tilde{M}_{CS})}{V(\tilde{M}_{CS})} \cong \frac{\text{cov}(e_{i5}, \tilde{M}_{CS})}{V(\tilde{M}_{CS})}$ the overall bias tends to zero and differences in the coefficients reflect genuine changes in peer effects. Even if the two terms are not identical, as long as the covariances decrease or remain similar, it is likely that the omitted variable bias is attenuated when comparing effects over time. This equality depends on how similar the two error terms across two grade-specifications are. Unobserved characteristics at the individual level, that would fall into the residual, are not likely to be an issue as they are – with high probability and on average – time invariant. For instance, individual ability is likely to be an intrinsic characteristic that students carry with them from primary to upper secondary education with little variation. Other unobserved factors such as specific school characteristics could represent an issue. Fundings to education, teacher quality or school location – in an urban or rural area – are some aspects specifically related to school that can affect students’ performance. However, families’ school choice is driven by many other characteristics that are with high probability related to their socio-economic background. Parents with higher levels of education, that attach large value to schooling, will choose “better schools” for their children, at any stage of their education. By these means, it is probable that specific school-characteristics are similar enough between grades, because schools themselves are similar (located in urban areas, prestigious or with positive reputation). Our approach mostly relies on the assumption that unobserved factors contributing to OVB are relatively stable over time.

The primary source of omitted variable bias in this context is the non-random sorting of children across residential areas and schools. Families indeed do not randomly allocate across towns, but choose where to live based on housing costs, employment opportunities and other factors, many of which are difficult to observe. It is reasonable to assume that families with similar socio-economic background and other unobserved characteristics cluster in the same areas and select in which school to enroll their children depending on residential proximity. Consequently, pupils with similar characteristics cluster in the same schools. In addition, non-random allocation of students into classes by school administrators or education authorities may reinforce this sorting. Essentially, this means that there may be multiple unobserved factors influencing both immigrant presence and school performance.

The key assumption is that these unobservable factors influencing both student outcomes and immigrant concentration in school over time are relatively stable over time. This assumption is plausible in contexts characterized by low residential and school mobility and where the mechanisms of students' allocation to schools and classes remain relatively unchanged across school levels. As a result, the potential omitted variable bias affecting the estimates is expected to be similar at different stages of schooling.

One potential limitation to this assumption is the issue of attrition over time, which arises from the loss of observations due to students repeating grades, dropping out of school or being absent in the 8th and 13th grades waves of the INVALSI assessments for any reason. As a result, the longitudinal sample may disproportionately consist of students who progress through school without interruption, who are – with high probability – from a more advantaged background. This selective attrition may violate the assumption that omitted variable bias remain stable, if unobserved characteristics – such as family motivation – or observed ones – such as exposure to immigrant peers – influence the likelihood of staying in the sample.

For these reasons, it is crucial to adopt appropriate econometric techniques to mitigate any residual endogeneity and enhance the credibility of our estimates. Also, it will be relevant to check to what extent our treatment affect attrition in 8th and 13th grade, as this would indicate that peer composition influences sample selection.

5.2 Identification strategy and limitations

If it is the case that attrition across grades introduce selection in the sample, our assumption that OVB are relatively stable over time trembles. It becomes central to control for the non-random distribution of children across classrooms and schools to properly identify peer effects, as failure to do so could exacerbate the bias in estimating the influence of immigrant peers over time.

A first and critical concern arises from the non-random distribution of children into classes within the same school. Allocation decisions may depend on observed and unobserved characteristics such as language proficiency. This endogeneity in classroom composition can lead to biased estimates of peer effects if, for example, immigrant students are disproportionately placed in certain classes due to unobserved factors also correlated with educational outcomes. One theoretically appealing approach to address this issue would be to instrument the class-level immigrant concentration with grade-level immigrant concentration within the school. The idea is that while class composition may be endogenous, the overall

immigrant share at the grade level is less likely to be influenced by within-school sorting and thus can serve as an exogenous source of variation.

However, as pointed out by Gould et al. (2009), when class size is fixed the IV estimates tends to be numerically identical to the OLS ones. Given that in the Italian school system class size within the same grade and school are fairly stable, this IV strategy would likely not add much value to the estimates.

To avoid the issue of nonrandom distribution across classrooms, the number of immigrant students in 5th grade is computed at the school level rather than at class level. While most peer effects likely operate within classrooms, class-level immigrant shares could be endogenous due to non-random allocation of children according to unobserved characteristics. This choice sacrifices some granularity but enhances the reliability of the measure by reducing endogeneity concerns. Additionally, heterogeneity in classroom resources and teacher quality may vary across classrooms within the same school thus confounding class-level effects. For example, higher-performing classes might be assigned more experienced teachers, while lower-achieving or immigrant-heavy classes might be allocated less experienced staff — or, conversely, schools may direct their best teachers to struggling groups to narrow achievement gaps. This issue is especially relevant in the Italian school systems, where students are typically placed in fixed classes that remain stable over multiple years, often with the same teachers throughout. This structure fosters long-term peer relationships and teacher-student dynamics, making classroom-specific factors particularly influential for educational outcomes. In contrast, systems like those in the United States, Canada, or the United Kingdom use a course-based structure, where students enroll in different classes with varied peers and teachers, reducing the persistence of peer and teacher effects over time. As we do not have further information on classroom composition or quality of teachers and other resources, we rely on the solution of computing immigrant concentration at school-level, sacrificing some precision in a try to obtain more consistent estimate of immigrant peer effects. Despite the loss of within-school granularity, the use of school-level immigrant share is still informative as it captures the broader social and educational environment influencing all students and thus can proxy for important peer effects that happen beyond classroom, while mitigating the bias arising from endogenous class sorting. The key explanatory variable is the *fraction* of immigrants in elementary school at the 5th grade so that we partially capture differences in school sizes as well, recognizing that the influence of one immigrant student in a school of thirty differs from their influence in a school of one hundred.

To obtain a valid measure of immigrant exposure, we compute the treatment before the merging process as:

$$\text{Share of immigrants} = \frac{\text{number of immigrant test takers in 5}^{\text{th}} \text{ grade}}{\text{number of total test takers in 5}^{\text{th}} \text{ grade}}$$

By doing so, we consider all observations available in 5th grade, including those eventually lost when merging the different grade-level datasets.

Beyond class-level challenges, the study addresses also the potential bias introduced by non-random spatial distribution of children across towns and schools. Immigrant families tend to cluster in specific geographical areas, and these areas often differ systematically in socioeconomic characteristics, local policies, school quality, and labor market conditions — all factors potentially correlated with student outcomes. Additionally, families with better socioeconomic background may purposefully move their children to schools with lower immigrant concentration, if they believe that high share of non-native pupils is detrimental to their children performance, further exacerbating the separation of possibly higher achieving pupils and lower achieving ones in different schools or even towns. To control for such unobserved heterogeneity, the main empirical model includes province-level fixed effects. These fixed effects absorb time-invariant differences across provinces, thereby reducing confounding bias due to area-level factors that influence both the immigrant share and student achievement. By controlling for unobserved, time-invariant characteristics of geographical areas, fixed effects should help reducing the size of the distortion induced by the non-random placement of children into schools. Fixed effects are a common solution that was largely employed in literature on peer effects as previously highlighted. Ideally, fixed effects at more granular geographical levels, such as municipalities or even schools, would better control for unobserved heterogeneity. Unfortunately, the current INVALSI dataset lacks detailed municipality identifiers, restricting the spatial fixed effects to the province level, and heterogeneity between schools within the same province remains partially uncontrolled. Moreover, school-level fixed effects are not feasible because the key independent variable is measured at the school level, and including fixed effects would absorb all variation in the explanatory variable. In addition, the identification strategy assumes that the time-invariant unobserved factors at the province level sufficiently capture the relevant confounders. However, any time-varying shocks or policy changes that affect both immigrant concentration and student outcomes within provinces may still bias the estimates. To mitigate this, future research could explore the inclusion of time-varying controls or dynamic fixed effects, if data availability permits.

It is important to acknowledge that some limitations remain. The reliance on school-level immigrant shares may underestimate the true magnitude of peer effects operating in class. Similarly, unobserved variable that vary within province over time could bias the results. Despite these limitations, the adopted identification strategy balances data constraints and econometric methods to produce consistent and interpretable estimates of immigrant peer effects on native students' performances in school.

6 Results

6.1 Baseline results

In this section, we present the results from the estimation of our model. The dependent variables are the logarithm of the scores in mathematics and reading in 5th, 8th and 13th grade separately. Log-transforming test scores facilitates the interpretation of the estimated coefficients and allows for more straightforward comparison with findings from related studies. Differently from some previous studies, we consider outcomes in both subjects. Prior research has often excluded language test scores, given the evident initial disadvantage faced by non-native students in the host-country language. To our eyes, however, this is precisely what makes it especially interesting to examine how the effects on reading test scores evolve over time. If immigrant students display weaker initial performance in this subject, this may intensify the potential negative consequences associated, for example, with teachers reallocating attention towards weaker-performing students, potentially diverting resources from native pupils, who are likely to possess a stronger command of the language regardless. By considering only the subgroup on native students, we hope to be able to capture the possible negative effect caused by higher concentration of students who do not master the mainstream language.

We first present the result obtained on the whole sample, which includes native and non-native pupils. Table 6.1 shows the estimates of the baseline model in which the dependent variable is the log test score in mathematics in 5th grade with errors clustered at the school level. The simplest model that regresses log math score in 5th grade on the share of immigrants results in a -0.0649 coefficient significant at the 1% level. Interestingly, when adding province fixed effects (column 2) the coefficient grows in absolute value – becoming more negative – and remains statistically significant. At the same time R^2 quintuplicates, suggesting that it is important to account for the non-random allocation of children across different areas and that province-level fixed effects successfully do so. Column 3 includes gender as control, and while the R^2 grows significantly, the coefficient does not change much. Gender is an important component that explains a portion of the variation in the variable of interest. Column 4 includes important variables that carry information on the socio-economic background of the student, such as whether they have at least one parents with higher education³¹ and at least one parent employed. The inclusion of a dummy equal to 1 if the student is in a lower grade than the one

³¹ To the scope of this research, we consider higher education a degree or a PhD, or anything above diploma.

they should be assigned to based on their year of birth (*delayed grade placement*) makes it possible to account for potentially lower achieving students³². Introducing these controls shrinks the coefficient of interest noticeably. The specification in column 5, that is the preferred one, accounts for all meaningful children's characteristics, including whether they attended nursery and kindergarten, and the cheating propensity score compute by the INVALSI Institute. Nursery and kindergarten attendance are important aspects that shall be included for two main reasons. Considering native students only, early introduction into the educational system can foster their development of social skills. A kid that was introduced since the earliest years of their life into an unfamiliar setting is likely to be different from one who spent those same years with family members only. In particular, children that were early exposed to and encouraged towards social interactions, may react better to increased classroom heterogeneity, maybe even positively leveraging it. With respect to immigrant students, this dummy partially captures for how long they have been in Italy and how early they were introduced into the Italian school system, as the INVALSI Institute records whether children attend nursery and kindergarten only in Italy. Attendance abroad is not positively registered because the purpose of such information is to monitor educational paths in the Italian system. Non-native children that were introduced since the first years of their lives in environments shared with native pupils are likely to have better knowledge of the host-country language and to be more similar in terms of behaviors and social skills. Unfortunately, however, these two variables were filled with missing values in the INVALSI dataset; for this reason, two additional dummies were included in the model, that take value 1 when there is no information on nursery and kindergarten attendance respectively for the pupil. In this way, we manage to not lose too many observations and also to pick up any possible existing pattern behind those missing values, that is to say if the lack of information is an information itself with respect to children's preexistent

³² At the elementary school it is unlikely that children are purposefully held back a grade, as it often happens in lower and upper secondary education. However, there are still important circumstances in which it can happen. Immigrant students are often placed in a grade prior to the one they should be assigned depending on their age, to facilitate their introduction into the school system and allowing them to reinforce their knowledge in the mainstream language. This is especially common in Italy as, differently from other European countries, there exist no parallel classes that allow non-native children to keep up with the school program at the exact grade in which they should be placed while strengthening their skills in the host country language. In addition, there are some exceptional case in which children loose a year because of health reason – in the event that they lose too many school days – or because of behavioral reason.

characteristics. The cheating propensity score is another valuable input that can explain variation in students' performance. This indicator is computed by INVALSI at the class level, and it is based on anomalous statistical indicators detected in the response patterns during the test, and it allows to take into account the risk that some classes may have engaged in opportunistic behaviors, boosting test grades. When adding these controls, the R^2 triplicates. With respect to the specification in column 4, the estimated coefficient indicates a more adverse effect of immigrant concentration on students' performance in 5th grade. In particular, an increase of one point in the share of immigrants, reduces children performance by -5,83%, significant at the 1% level.

Table 6.1. Baseline regression results for 5th grade mathematics test scores.

	Dependent variable: log math score in 5th grade				
	(1)	(2)	(3)	(4)	(5)
Share of immigrants	-0.0649*** (0.00377)	-0.0744*** (0.00852)	-0.0725*** (0.00854)	-0.0537*** (0.00841)	-0.0583*** (0.00660)
Female			Yes	Yes	Yes
At least one parent with higher education				Yes	Yes
At least one parent employed				Yes	Yes
Delayed grade placement				Yes	Yes
Cheating propensity					Yes
Nursery					Yes
Kindergarten					Yes
Missing values dummies					Yes
Fixed effects		Yes	Yes	Yes	Yes
R-squared	0.00137	0.00708	0.0241	0.0439	0.144
N	216414	216414	216414	216414	216414

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the full sample. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6.2 and Table 6.3 show the same specifications used in Table 6.1, repeated using log test scores in Mathematic in 8th and in 13th grade³³. Similar patterns to those highlighted in Table 6.1 emerge. Initially, with no fixed effects at province level (column 1), no significant effect of immigrant concentration is detected neither on 8th grade performance nor on 13th grade performance. The estimated coefficient becomes significant at the 1% percent level as soon as province fixed effects are added to the model (column 2). Interestingly, the estimated coefficient in column 2 of Table 6.2 is larger in absolute value by 1.7% points than the one estimated in Table 6.1. Adding gender as a control (column 3) does not change largely the estimated coefficient again. According to the preferred specification (column 5), one additional unit in immigrant concentration in 5th grade decreases students' performance in mathematics in 8th grade by -6.69%, larger than the estimated coefficient in column 5 of Table 6.1 by 0.0086 percentage points. Apparently, the effect of immigrant concentration measured in 5th grade has a more negative impact on results in 8th grade compared to that in 5th grade. This can be explained by the cumulative nature of peer effects or by contextual influences over time. It looks like the immediate impact is more limited, while in the medium term, a higher concentration of immigrant peers is likely to have influences learning dynamics and social integration experiences, in such a way that manifest clearly after several years. In addition, this result can be explained by the possible increasing difficulty in compensating for the initial disadvantage created by non-native share over time, causing the negative effects to become more pronounced as students progress through the educational system.

³³ The minor difference in sample size is due to the fact that the test score in a specific subject is not available for some students. It may be that students who attended school on the day of the reading test were missing for any reason on the day of math test. We assume the distribution of these missing values to be random, as the number is extremely limited. As so, it should not influence the main results.

Table 6.2. Baseline regression results for 8th grade mathematics test scores.

	Dependent variable: log math score in 8 th grade				
	(1)	(2)	(3)	(4)	(5)
Share of immigrants	0.00536 (0.00350)	-0.0914*** (0.00611)	-0.0899*** (0.00613)	-0.0661*** (0.00566)	-0.0669*** (0.00573)
Female			Yes	Yes	Yes
At least one parent with higher education				Yes	Yes
At least one parent employed				Yes	Yes
Delayed grade placement				Yes	Yes
Cheating propensity					Yes
Nursery					Yes
Kindergarten					Yes
Missing values dummies					Yes
Fixed effects		Yes	Yes	Yes	Yes
R-squared	0.0000109	0.0634	0.0756	0.114	0.115
N	216403	216403	216403	216403	216403

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the full sample. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6.3 displays the results obtained using math scores in 13th grade. Column 5 suggests that the negative effect exerted by non-native share at the elementary school level lasts up to 13th grade; in particular, a one percentage point increase in the share of immigrants in 5th grade, reduces performance in mathematics in 13th grade by -6.61%, an effect similar to the one identified in 8th grade and larger in absolute value than the one identified in 5th grade.

Table 6.4, Table 6.5 and Table 6.6 show the same specifications using reading test scores in 5th, 8th and 13th grade, with errors clustered at the school level. Again, including province fixed effects always increase the magnitude of the negative effect – with the coefficient of interest being more negative – and the R², suggesting these models to be more effectively capturing the variation on test scores.

Table 6.3. Baseline regression results for 13th grade mathematics test scores.

	Dependent variable: log math score in 13th grade				
	(1)	(2)	(3)	(4)	(5)
Share of immigrants	0.0250*** (0.00380)	-0.0793*** (0.00536)	-0.0771*** (0.00539)	-0.0523*** (0.00496)	-0.0528*** (0.00500)
Female			Yes	Yes	Yes
At least one parent with higher education				Yes	Yes
At least one parent employed				Yes	Yes
Delayed grade placement				Yes	Yes
Cheating propensity					Yes
Nursery					Yes
Kindergarten					Yes
Missing values dummies					Yes
Fixed effects		Yes	Yes	Yes	Yes
R-squared	0.000200	0.0587	0.0814	0.115	0.115
N	216414	216414	216414	216414	216414

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the full sample. Significance levels: * p < 0.10, ** p < 0.05, *** p < 0.01.

As it was observed for the previous models, the inclusion of specific covariates reduces the magnitude of the estimated coefficient. The specification in column 5 of Table 6.4, that considers all meaningful individual characteristics and province-level fixed effects, suggests that a one unit increase in immigrant share reduces students' performance in reading in 5th grade by -7.71% significant at the 1% level. As expected, it emerges here that immigrant concentration has a more detrimental effect on performance in reading than it has on mathematics, for which the estimated coefficient in 5th grade was -5.83%, almost two percentage points of difference.

Table 6.4. Baseline regression results for 5th grade language test scores.

	Dependent variable: log reading test score in 5th grade				
	(1)	(2)	(3)	(4)	(5)
Share of immigrants	-0.0840*** (0.00371)	-0.0966*** (0.00850)	-0.0967*** (0.00850)	-0.0738*** (0.00827)	-0.0771*** (0.00753)
Female			Yes	Yes	Yes
At least one parent with higher education				Yes	Yes
At least one parent employed				Yes	Yes
Delayed grade placement				Yes	Yes
Cheating propensity					Yes
Nursery					Yes
Kindergarten					Yes
Missing values dummies					Yes
Fixed effects		Yes	Yes	Yes	Yes
R-squared	0.00236	0.00758	0.00770	0.0362	0.0848
N	216391	216391	216391	216391	216391

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the full sample. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

At the 8th grade (Table 6.5), differently from what was observed for the specification using Mathematic scores, the effect of immigrant concentration remains significant at the 1% level even when no province-level fixed effects are applied. According to column 5, there is just -0.16 percentage points difference between the estimated effects in 5th grade and 8th grade: one unit increase in the immigrant share in 5th grade decreases performance in reading in 8th grade by -7.55%, significant at the 1% level.

Table 6.5. Baseline regression results for 8th grade language test scores.

	Dependent variable: log language test score in 8th grade				
	(1)	(2)	(3)	(4)	(5)
Share of immigrants	-0.0285*** (0.00331)	-0.0995*** (0.00555)	-0.100*** (0.00551)	-0.0748*** (0.00500)	-0.0755*** (0.00508)
Female			Yes	Yes	Yes
At least one parent with higher education				Yes	Yes
At least one parent employed				Yes	Yes
Delayed grade placement				Yes	Yes
Cheating propensity					Yes
Nursery					Yes
Kindergarten					Yes
Missing values dummies					Yes
Fixed effects		Yes	Yes	Yes	Yes
R-squared	0.000342	0.0400	0.0457	0.0901	0.0908
N	216408	216408	216408	216408	216408

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the full sample. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

While the estimated coefficient of interest remains statistically significant up to 13th grade (Table 6.6), one can observe that the size of the effect become smaller over time, differently from what recorder for the specifications using math test scores. However, up to 8th grade, the effect on reading performance seems more detrimental than the one on mathematics scores. That is most likely because of the additional care non-native children require in Italian language as they do not possess the innate advantage of Italian-born children. Conversely, mathematics is a relatively new subject for both native and immigrant children, and the smaller negative effect is probably due just to the difficulty coming from the subject being taught in Italian also to non-native children. The estimated effects in 13th grade on both subjects are relatively similar, with a gap of 0.22 percentage points only.

Table 6.6. Baseline regression results for 13th grade language test scores.

	Dependent variable: log language test score in 13th grade				
	(1)	(2)	(3)	(4)	(5)
Share of immigrants	-0.00107 (0.00414)	-0.0908*** (0.00611)	-0.0929*** (0.00601)	-0.0634*** (0.00545)	-0.0639*** (0.00551)
Female			Yes	Yes	Yes
At least one parent with higher education				Yes	Yes
At least one parent employed				Yes	Yes
Delayed grade placement				Yes	Yes
Cheating propensity					Yes
Nursery					Yes
Kindergarten					Yes
Missing values dummies					Yes
Fixed effects		Yes	Yes	Yes	Yes
R-squared	0.0000003	0.0396	0.0581	0.0968	0.0970
N	216414	216414	216414	216414	216414

Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the full sample. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

These results indicates that there may be some negative effects originating from immigrant concentration that last up to upper secondary school. While the gap between the estimated coefficient for the two subjects in 5th grade is quite large – 1.88% difference – this asymmetry is lost over time, and the estimated effects of immigrant share in 5th grade on performance in both subjects in 13th grade are similar.

Nevertheless, the main concern of this study is to understand how immigrant concentration affects native pupils specifically. Table 6.7 exhibits the result of the preferred specification (as the one in column 5 of Table 6.1) applied to the subsample of native children only, using robust standard errors clustered at the school level. Individual characteristics comprises all controls included above, that is to say: gender, dummies for occupation and education of the parents, nursery and kindergarten attendance (together with dummies to account for missing values), whether the child is a grade behind and the class level cheating propensity score. At first glance, it can be observed that the coefficient of interest remains statistically significant at the 1% level up to 13th grade for both subjects, pointing to the persistence of the effects arising from early immigrant exposure. For what concern mathematics test scores (columns 1 to 3) the estimated

effect is negative but small in size, suggesting that immigrant concentration reduces performance in 5th grade by 2.65%, in 8th grade by -4.21% and in 13th grade by 3.53%, all coefficients being significant at the 1% level. Apparently, the effect of immigrant share computed at the elementary school level grows in absolute value by 1.56 percentage points between 5th and 8th grade, suggesting that not only immigrant peer effects last over time but they also become more detrimental. The estimates coefficients on reading test scores appear to be more consistent over time, showing a stable similar in size effect of immigrant share across different grades.

However, to assess whether the trajectory highlighted by these estimates is realistic, it is essential to formally test whether the estimated coefficients across different specifications are statistically identical. Comparing the effect on 5th grade outcomes with that on 8th and 13th grade outcomes allows to evaluate whether the influence of early exposure to immigrant peers remains stable, diminishes or amplifies as students progress through the educational system. To test so, we perform a Wald test following a *suest* procedure on Stata³⁴. The null hypothesis is that the coefficients of interest are equal across the two models under comparison. In order to determine the actual trajectory of peer effects, we begin by testing whether the coefficient of interest is statistically equal across grades at the same subject-level. When testing the coefficients of interest across the models in column 1 and 2 of Table 6.7 – on 5th and 8th grade mathematics test scores respectively – we reject the null hypothesis ($\chi^2 = 5.54, p = 0.0186$), which indicates that the estimates are statistically different across the two specifications. In other words, there is a statistical difference between these two coefficients, meaning that the effect of immigrant exposure in 5th grade are slightly more detrimental to 8th grade outcomes than they are to contemporaneous 5th grade outcomes. One possible explanation for the increasing effect of immigrant concentration on mathematics performance over time is that mathematics skills are more cumulative and sequential in their nature. Early gaps in the learning environment – such as high exposure to poorer performing peers, lower classroom resources or disruptive peer dynamics – can compound as students progress through higher grades. As a result, we observe that the negative peer effect of higher immigrant concentration in 5th grade

³⁴ Stata's *suest* (Seemingly Unrelated Estimation) procedure is a method that combines separately estimated regression models and recovers a joint variance-covariance matrix of the coefficients, allowing for valid cross specifications comparisons. The Wald test, performed using the *test* command, assess whether the coefficient of immigrant concentration is statistically equivalent across models.

seems to amplify by 8th and 13th grade. In contrast, language skills are less sequentially dependent and are more influenced by exposure to language both in and out of school. Native students might compensate for early classroom disadvantages through family resources or informal social environments, which helps to maintain more stable outcomes in language tests over time, despite initial classroom composition. Table 6.8 reports the results of the equality test conducted on different pairs of specifications of Table 6.7. These tests reveal consistent patterns in the estimated effect of 5th grade immigrant concentration across different educational stages. In mathematics, as already highlighted, the effect differs significantly between 5th and 8th grade, so that we observe that the negative effect on mathematics increases significantly in lower secondary school, pointing to the possibility that peer effects arising from non-native shares accumulate over time or have a delayed effect. However, no statistical difference is observed between 8th and 13th grade or between 5th and 13th grade. While the effect on natives' performance in mathematics intensifies during the earlier years, apparently it stabilizes or potentially weakens at the end of upper secondary school, resulting in no detectable change in the long-term. Comparing 5th grade and 13th grade estimates can be however trickier, as the net change is moderate and less precisely estimated over a large span of years, and the test lacks power to detect a statistically significant difference. In contrast, the coefficients of interest in the reading specifications remain statistically equivalent across all grades comparisons, suggesting that – unlike in mathematics – the impact of immigrant school share on language outcomes is either stable or changes minimally throughout students' academic trajectories. In summary, we find that the share of non-native pupils in 5th grade negatively affects native students' performance in both mathematics and language up to the end of upper secondary school, with an effect that ranges between -2.49% and -3.53% in 13th grade depending on the subjects. The impact on test scores in language assessments appears stable over time and statistically significant at the 1% level. On the other hand, the effect on mathematics performance amplifies in 8th grade, suggesting that the peer effects may cumulate over time. As pointed out earlier in this section, this is most likely due to the sequential nature of mathematic knowledge and the greater difficulty in filling any preexistent gap in this subject.

When looking at Table 6.7 it may seem that the effect of early exposure to immigrant peers is more detrimental or inconsistent over time on performance in mathematics tests rather than in language ones. Similarly to what we did above, we test for the equivalence of the coefficients of interest across different subjects at the same grade-level (Table 6.9). According to these results, the effects of immigrant concentration on mathematics and reading in 5th grade are not statistically different ($\chi^2 = 0.02$, $p = 0.8867$). This suggests that early immigrant exposure does

not affect contemporaneous outcomes at a subject-specific level, but rather is likely to impact their overall school performance. However, by 8th grade, the difference between the coefficients of interest becomes highly significant, suggesting that the immigrant peer effect diverges substantially between the two subjects in lower secondary school, with a more detrimental effect on mathematics. In 13th grade, the difference is still statistically significant but less pronounced than in 8th grade, suggesting that the differential impact of immigrant concentration on mathematics and language skills continues up to upper secondary education. Taken together, there is evidence that while the effect of immigrant school share is similar across subjects in primary school, it becomes increasingly differentiated as students' progress through the education system. Interestingly, our finding that the immigrant peer effects are more detrimental to mathematics skills than they are to language skills, similarly to what highlighted by Contini (2013). In theory, we would expect to observe negative effects that are stronger on reading performance than on mathematics'. Non-native children usually have limited language proficiency and may require additional attention from teachers in this subject. Our results however suggest a different story. Contemporaneous effects – that is the degree to which 5th grade exposure to immigrant children affect 5th grade performance – are similar in magnitude and not statistically different across subjects. One percentage point increase in the immigrant share in primary school reduces performance by -2.65% in mathematics and by -2.56% in reading. Non-native children in primary school may lack proficiency in the mainstream language, potentially diverting teachers' attention. However, 5.26% of the immigrant children in our sample (7.12% of total observations) are second-generation immigrants, meaning that they were born in Italy. By the time they attend 5th grade, it is likely that they had time to strengthen their knowledge of the Italian language, thereby explaining why we do not observe particularly stronger effects in this subject. The distinction between first-generation and second-generation immigrants is important and will be addressed in detail later on. The effects on the two subjects statistically diverge in 8th and 13th grades. Apparently, early immigrant exposure has a larger in absolute value effect on mathematics up to the end of upper secondary school. These findings open up to two interpretations. First of all, mathematics skills are cumulative in their nature, making it more difficult to fill any potential initial gap. Second, language proficiency of native students is not influenced only by their school environment but is largely affected by their family and many other interactions that happen out of school.

Table 6.7. Main results for native students' math and reading scores in 5th, 8th, and 13th grades.

Dependent variable: log test score of native children only						
	Mathematics			Reading		
	5 th grade (1)	8 th grade (2)	13 th grade (3)	5 th grade (4)	8 th grade (5)	13 th grade (6)
Share of immigrants	-0.0265*** (0.00684)	-0.0421*** (0.00566)	-0.0353*** (0.00515)	-0.0256*** (0.00720)	-0.0268*** (0.00471)	-0.0249*** (0.00563)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.144	0.119	0.118	0.0828	0.0927	0.0994
N	201007	200999	201007	200987	201001	201007

Note: Robust standard errors in parenthesis are clustered at the school level. All regressions include a constant term. Significance levels; * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 6.8. Tests of equality of coefficients of immigrant share across grade-level specifications.

Subject	Comparison	χ^2	<i>p-value</i>	Interpretation
Mathematics	5 th grade vs 8 th grade	5.54	0.0186	Significantly different
	8 th grade vs 13 th grade	2.85	0.0914	Not significantly different
	5 th grade vs 13 th grade	1.75	0.1854	Not significantly different
Reading	5 th grade vs 8 th grade	0.03	0.8643	Not significantly different
	8 th grade vs 13 th grade	0.22	0.6360	Not significantly different
	5 th grade vs 13 th grade	0.01	0.9238	Not significantly different

Table 6.9. Tests of equality of coefficients of immigrant share across subject-level specifications, at the same grade.

Grade	Comparison	χ^2	<i>p-value</i>	Interpretation
5th grade	Mathematics vs Reading	0.02	0.8867	Not significantly different
8th grade	Mathematics vs Reading	15.79	0.0001	Significantly different
13th grade	Mathematics vs Reading	6.37	0.0116	Significantly different

Up to now, we used the INVALSI definition of native students, that is any children, born or not in Italy, who have at least one parent born in Italy. This definition is rather large and include children that may have recently arrived in the country from abroad, who might as well have limited language proficiency and require additional attention at the time of introduction into a foreign educational system. We restrict our sample to native children born in Italy only and estimate again all six specifications. To do so, we temporarily drop from our sample native children whose age of arrival is defined. INVALSI records for all children the age at which they arrived in the country. Pupils born in Italy report *mancante di sistema* (that can be translated as structurally missing) in that variable. We exploit this information to isolate native children born in the country. If we believe that born-abroad native children are similar in terms of educational experiences to immigrant students, then we would not be measuring exactly the causal effect of interest. Table 6.10 illustrates the results of the estimations on the reduced sample. The effects are similar in magnitude, direction and statistical significance to those found on the broader group of native children. There are two possible explanations to this. First, students born abroad by Italian parents are fully linguistically and culturally integrated by the time they enroll in school. Second, it may be that native children born in the country and those born abroad are affected similarly by early immigrant exposure, so that there is overall homogeneity within the native students only sample. As we do not find evidence that immigrant peers affect in different ways children born in Italy and children born abroad from Italian parents, all following estimations are computed on the sample of native pupils as identified by INVALSI data. Last but not least, these results add preliminary robustness to our initial analysis, showing that the core results are not sensitive to the way we define native students.

Table 6.10. Main results for native student’s math and reading scores in 5th, 8th, and 13th grades.

Dependent variable: log test scores of native children only (born in Italy)						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Share of immigrants	-0.0262*** (0.00686)	-0.0410*** (0.00566)	-0.0342*** (0.00513)	-0.0242*** (0.00720)	-0.0252*** (0.00467)	-0.0234*** (0.00560)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.143	0.118	0.118	0.0822	0.0916	0.0987
N	198892	198884	198892	198873	198886	198892

Note: Robust standard errors in parenthesis are clustered at the school level. All regressions include a constant. The regressions are estimated on the sample of native students born in Italy. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

There is a concrete possibility that not all immigrant children affect their native peers in the same way. Multiple research address scenario in which first-generation and second-generation immigrants exert different effects on their schoolmates. The former are more likely to be weaker in the host country language and thus require additional attention by teachers in order to fill the gap they have with respect to the other children who have a better knowledge of said language. Moreover, they had less time to adapt to the new culture, the potentially new schooling system and education dynamics. It is also important to consider that, since we are computing the share of immigrants at the elementary school level, it may be for most of immigrants of first generation that they have recently arrived in Italy, thus exacerbating one more time the educational gap with respect to their peers. Appendix A.6.1 shows the distribution of first-generation immigrant children by their age of arrival in Italy. Table 6.11 shows the results obtained from regressing native children’s outcomes on the share of first-generation and second-generation immigrants separately. Both first-generation and second-generation immigrant share seem to affect native students’ performance significantly up to 13th grade in both subjects. We do not observe substantial differences in the coefficients of first- and second-generation immigrants, pointing to the possibility that these two groups do not exert different effects on the performance of native children. We test the hypothesis that the two coefficients are statistically equal within each specification, and in all cases, we fail to reject the null, meaning that we have no evidence that the two peer effects differ significantly.

In order to infer something about the trajectory of these effects over time, we performed one more time comparison across different grade-level specification to test whether the coefficients

of interest are statistically significant. According to the results obtained, reported in Table 6.12, there is no statistically significant difference across coefficients of first-generation immigrant share over time. In other words, the moderately negative effect that we identify is stable over time and remains significant at the 1% level up to 13th grade in both subjects. Similar conclusions can be made with respect to the coefficient on second-generation immigrant share (Table 6.13), which appear to be mostly stable over time. The coefficients of interest are only statistically different between 5th and 8th grade when considering performance in mathematics assessments.

Following that, we observe again that early exposure to immigrant peers – whether they are first-generation or second-generation ones – negatively and consistently affects school performance of native students, up to the end of upper secondary school.

These findings are in line with what observed by Contini (2013), who studies elementary schools in the Italian panorama as well. When distinguishing between percentage of first-generation immigrants and of second-generation immigrants in 5th grade, they find very similar negative and statistically significant effect on 5th grade reading and mathematics test scores. In their initial within-school regressions, they observe slightly stronger negative effects on performance in reading, as observed earlier in this work as well. Nevertheless, when they point their attention towards immigrant peer effects on the subgroup of native students only, they find that first-generation immigrant share has a larger effect in absolute value on mathematics performance, irrespective of the socio-cultural background³⁵ of native children. The effect of second-generation immigrant share is less stable across subject-level categories and shows an inconclusive pattern over different socio-cultural categories.

³⁵ They use the number of books possessed at home as a proxy for the socio-cultural background of children. The idea is that pupils with more books at home are more likely to have parents with higher education or to live in an environment in which education is looked up to. They identify five categories of children depending on the number of books, ranging from 0 books at home to 200 books at home. They perform separate regression on each subsample of native children as identified by three of the specified categories (the minimum, the middle and the maximum category) and they estimate equal coefficients over all three specifications.

Table 6.11. Effects of first-generation and second-generation immigrant shares on native students' outcomes.

Dependent variable: log test score of native children						
	Mathematics			Reading		
	5 th grade (1)	8 th grade (2)	13 th grade (3)	5 th grade (4)	8 th grade (5)	13 th grade (6)
Share of I gen. immigrants	-0.0375** (0.0166)	-0.0427*** (0.0126)	-0.0360*** (0.0121)	-0.00864 (0.0179)	-0.0321*** (0.0110)	-0.0325** (0.0131)
Share of II gen. immigrants	-0.0213** (0.00925)	-0.0453*** (0.00891)	-0.0371*** (0.00779)	-0.0329*** (0.00940)	-0.0281*** (0.00700)	-0.0239*** (0.00816)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.144	0.119	0.118	0.0828	0.0927	0.0994
N	201007	200999	201007	200987	201001	201007

Note: Robust standard errors in parenthesis are clustered at the school level. All models include a constant. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6.12. Tests for equality of coefficients of first-generation immigrant share across grade-level specifications.

Subjects	Comparison	χ^2	<i>p-value</i>	Interpretation
Mathematics	5th grade vs 8th grade	0.11	0.7383	Not significantly different
	8th grade vs 13th grade	0.47	0.4911	Not significantly different
	5th grade vs 13th grade	0.01	0.9251	Not significantly different
Reading	5th grade vs 8th grade	1.91	0.1668	Not significantly different
	8th grade vs 13th grade	0.00	0.9685	Not significantly different
	5th grade vs 13th grade	1.69	0.1932	Not significantly different

Table 6.13. Test results for equality of coefficients of second-generation immigrant share across grade-level specifications.

Subjects	Comparison	χ^2	<i>p-value</i>	Interpretation
Mathematics	5th grade vs 8th grade	6.36	0.0117	Significantly different
	8th grade vs 13th grade	2.11	0.1467	Not significantly different
	5th grade vs 13th grade	2.94	0.0864	Not significantly different
Reading	5th grade vs 8th grade	0.24	0.6250	Not significantly different
	8th grade vs 13th grade	0.58	0.4477	Not significantly different
	5th grade vs 13th grade	0.74	0.3907	Not significantly different

There is another interesting hypothesis that is worth testing, that is whether male and female immigrant children exert different effects on their native peers in school. To do so, we compute separately the share of male immigrant students and female immigrant students in 5th grade and estimate the same specification including both as treatments. To avoid issues due to collinearity we excluded the variable *gender* from individual characteristics. Table 6.14 reports the results obtained. The estimated coefficients for *male immigrant share* are more negative and consistent across time and remain always significant at the 1% level. The share of female immigrants shows more inconsistent effects across grades and subjects. These findings suggest that the share of male non-native students steadily exert negative influences on the performance of native students. There exists a portion of literature on peer effects which confirms that male peers tend to have larger negative effects on achievement of other students. Our findings align with these theories, as we fail to detect consistent peer effect arising from the concentration of female immigrant students.

Table 6.14. Effects of male and female immigrant shares on native students' outcomes.

Dependent variable: log test score of native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Share of male immigrants	-0.0388*** (0.0126)	-0.0482*** (0.00944)	-0.0380*** (0.00896)	-0.0265** (0.0134)	-0.0339*** (0.00814)	-0.0332*** (0.00965)
Share of female immigrants	-0.0143 (0.0129)	-0.0361*** (0.0107)	-0.0331*** (0.00973)	-0.0246* (0.0137)	-0.0191** (0.00913)	-0.0154 (0.0108)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.129	0.108	0.0972	0.0824	0.0853	0.0791
N	201007	200999	201007	200987	201001	201007

Note: Robust standard errors in parenthesis are clustered at the school level. All models include a constant term. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.2 Heterogeneous peer effects

Research on peer effects often tackles the possibility that peer effects are not homogeneous across the population of students under analysis. There is extensive evidence that confirms the pivotal role of peers, while suggesting that effects arising from peer composition are most often heterogeneous.

E. Hanushek et al. (2003) find that peer average achievement affects all students, but those in the top quartile of the school achievement distribution seem to be less responsive to peer achievement. Zimmerman (2003) determines that only students in the middle 70% of the SAT distribution are affected by their roommates' performance, whereas the impact on weaker and stronger students is not statistically significant. Brunello et al. (2010) observe that peer ability has a statistically significant effect only in the hard sciences field; the same does not hold for social sciences and humanities, suggesting that the field of study may matter as well in determining the impact of peer effects. There is research that focuses also on whether different genders and ethnicities are affected differently by their peers. Gould et al. (2009) examines whether low-status children and high-status ones are affected differently by immigrant concentration in school on long term outcomes such as dropout rates. They define native children with *low socio-economic status* as those whose parents have less than 10 years of schooling or come from Africa-Asia ethnicity and find that immigrant presence does not affect

dropout rates not for *high-status children* nor for *low-status* ones. Conversely, they find that immigrant peer effects are strongly negative on matriculation rates only for *low-status children*, while null for *high-status* ones, supporting the idea that there might be heterogeneity in peer effects that depends on predetermined individual characteristics. Contini (2013) classifies native children according to the number of books at home used as a proxy of their socio-economic background and they find peer effects arising from immigrant concentration to be highly heterogeneous. In particular, students with low socio-economic background (with less than 10 books at home) whereas students with higher socio-economic background (at least 200 books at home) seem to slightly benefit from the presence of second-generation immigrants. Tonello (2016) tests whether male and female native children are affected differently by non-native school share but find no significant effect for neither of the two subgroups.

Following the steps drew from past research, we first try to assess whether male and female native children are affected differently in the long-term. Table 6.15 and Table 6.16 presents the result of male and female native students separately. For both these groups, the effect of early immigrant exposure appears to remain stable and significant up to 13th grade. Besides that, it doesn't seem like the trajectory of peer effects differ depending on the gender. Females seems to undergo on average larger negative effects of immigrant concentration both in their mathematics performance and language one. When testing for the equality of the coefficients across the two groups at the same grade and subject level, we do not find any statistical difference. In other words, we find evidence that male and female native students do not differ in the way they are affected by exposure to immigrant peers in primary school.

As for the trajectory of these effects, we test the equality of the coefficient of interest across different grade-specifications and for the same subjects. In the female sample, we find that the coefficient of the share of immigrants is statistically different between the specifications in column 1 to 3, whereas we do not find any statistical difference between the columns 4 to 6. The takeaway can be that the negative effect of immigrant concentration on reading is more consistent over time – and does not fade away – whereas that on mathematics performance follows a different path, while still remaining evident and significant over the whole school life of native children. Something similar is observed for the subsample of male native children.

Table 6.15. Main results for female native students' math and reading scores in 5th, 8th, and 13th grades.

Dependent variable: log test scores of female native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Share of immigrants	-0.0303*** (0.00809)	-0.0455*** (0.00704)	-0.0344*** (0.00687)	-0.0297*** (0.00802)	-0.0254*** (0.00611)	-0.0263*** (0.00694)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.142	0.109	0.0959	0.0864	0.0846	0.0828
N	106288	106281	106288	106279	106285	106288

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample of female native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6.16. Main results for male native students' math and reading scores in 5th, 8th, and 13th grades.

Dependent variable: log test score of male native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Share of immigrants	-0.0222*** (0.00803)	-0.0378*** (0.00677)	-0.0360*** (0.00698)	-0.0209** (0.00885)	-0.0283*** (0.00612)	-0.0233*** (0.00817)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.120	0.112	0.106	0.0794	0.0887	0.0809
N	94719	94718	94719	94708	94716	94719

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample of male native children only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Angrist & Lang (2004) suggest that within-gender effects are more important. In other words, females are more influenced by their female peers while males are highly affected by their male peers. To test for this occurrence, we estimate our main regressions on the samples of female and male native students separately, using both the share of male immigrants and the share of female immigrants as treatments. Table 6.17 and Table 6.18 show the results of these estimations. We observe that the coefficient of male immigrant share is always negative and significant at the 1% level in all specifications except for one, for both male and female native students. On the other hand, the effect of female immigrant share is not consistently estimated for any of the two subsamples. When considering only female native students, we observe that

they are weakly influenced by the presence of female immigrant peers, especially on 8th and 13th grade mathematics outcomes. Male native students seem to not suffer any detrimental effect arising from the concentration of female immigrant peers. These findings are in line with most existing literature on gender peer effects. Male non-native students seem to generate stronger negative spillovers on their native peers of both genders, that remain highly significant and similar in size up to 13th grade. We find evidence as well of within-gender effects; we observe weakly significant coefficients of the female immigrant share in the sample of female native students, but not as much in the sample of male native students. On average, male immigrant students appear to exert a large detrimental effect – that does not fade over time – both on female and male native students.

Table 6.17. Effects of male and female immigrant shares on female native students' math and reading scores in 5th, 8th, and 13th grade.

Dependent variable: log test score of female native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Share of male immigrants	-0.0389*** (0.0150)	-0.0497*** (0.0118)	-0.0375*** (0.0116)	-0.0239 (0.0153)	-0.0315*** (0.0106)	-0.0280** (0.0118)
Share of female immigrants	-0.0214 (0.0155)	-0.0411*** (0.0132)	-0.0311** (0.0129)	-0.0356** (0.0157)	-0.0192* (0.0116)	-0.0245* (0.0131)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.142	0.109	0.0959	0.0864	0.0846	0.0828
N	106288	106281	106288	106279	106285	106288

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample of female native children only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6.18. Effects of male and female immigrant shares on male native students' math and reading scores in 5th, 8th, and 13th grade.

Dependent variable: log test score of male native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Share of male immigrants	-0.0418*** (0.0150)	-0.0499*** (0.0123)	-0.0432*** (0.0129)	-0.0287* (0.0160)	-0.0345*** (0.0113)	-0.0346** (0.0148)
Share of female immigrants	-0.00151 (0.0152)	-0.0250* (0.0132)	-0.0283** (0.0132)	-0.0126 (0.0163)	-0.0217* (0.0119)	-0.0114 (0.0153)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.120	0.112	0.106	0.0794	0.0887	0.0809
N	94719	94718	94719	94708	94716	94719

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample of male children only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Another important dimension along which children differ deeply is their socio-economic background. We proxy it using parental education and estimate our regressions on separate samples of native children, to see to what extent they are affected by immigrant concentration in 5th grade depending on their socio-economic predetermined environment. We identify students with high background as those with at least one parent with higher education, that is a degree or anything above. Children with poor socio-economic background have at least one parent who completed, at maximum, lower secondary education. All other cases are flagged as average socio-economic background. Table 6.19, Table 6.20 and Table 6.21 below show the estimated coefficients for three samples of native students separately, depending on their socio-economic background. The estimated effects on children whose parents have low education are the strongest; also, they are always statistically significant at the 1% in all grade-level and subject-level specifications. The estimated effect on children with average socio-economic background is still significant and consistently estimated, but less negative. Finally, we find that children of parents with high education are not affected by early immigrant exposure. For these students, even contemporaneous effects are not strongly defined. These findings are in line with most existing research on education and peer effects, that find in general stronger negative effects on students with poorer background. Contini (2013) as well finds that the adverse results on matriculation rates are exclusively driven by the strong negative effects of immigrant concentration on natives from a low socio-economic background. One potential concern over these results is that such large heterogeneity in the effects may be partly driven

by greater exposure to immigrant peers for specific group of native children. However, that is not the case here, as the average immigrant concentration across groups is very similar: 0.104 (std.dev. = 0.9) for students from high socio-economic background, 0.116 (std.dev. = 0.1) for those from poor background, 0.109 (std.dev. = 0.9) for the native students from medium background. These small differences suggest limited variation across the three groups, hence we conclude that there is strong evidence that native children from a poor socio-economic background are those more adversely affected by early exposure to immigrant peers. A portion of the next section is dedicated to estimate our model on different subsamples of native students, identified by geographical areas and by the number of books at home, to test further heterogeneity in peer effects as well as the sensitivity of our specifications.

Table 6.19. Effect of 5th grade immigrant concentration on native children with high socio-economic background.

Dependent variable: log test scores of native students						
	Mathematics			Reading		
	5 th grade (1)	8 th grade (2)	13 th grade (3)	5 th grade (4)	8 th grade (5)	13 th grade (6)
Share of immigrants	0.0178* (0.0106)	-0.0137 (0.00874)	-0.00421 (0.00903)	0.0200* (0.0108)	0.00187 (0.00752)	0.0137 (0.00941)
R-squared	0.114	0.0636	0.0797	0.0502	0.0434	0.0565
N	52783	52783	52783	52781	52781	52783

Table 6.20. Effect of 5th grade immigrant concentration on native children with average socio-economic background.

Dependent variable: log test scores of native students						
	Mathematics			Reading		
	5 th grade (1)	8 th grade (2)	13 th grade (3)	5 th grade (4)	8 th grade (5)	13 th grade (6)
Share of immigrants	-0.0231*** (0.00897)	-0.0347*** (0.00867)	-0.0266*** (0.00783)	-0.0194** (0.00898)	-0.0191*** (0.00700)	-0.0191** (0.00853)
R-squared	0.131	0.0775	0.0877	0.0639	0.0490	0.0656
N	79140	79136	79140	79130	79139	79140

Table 6.21. Effect of 5th grade immigrant concentration on native children with poor socio-economic background.

Dependent variable: log test scores of native students						
	Mathematics			Reading		
	5 th grade (1)	8 th grade (2)	13 th grade (3)	5 th grade (4)	8 th grade (5)	13 th grade (6)
Share of immigrants	-0.0446*** (0.00945)	-0.0455*** (0.00744)	-0.0423*** (0.00755)	-0.0451*** (0.00999)	-0.0328*** (0.00699)	-0.0319*** (0.00813)
R-squared	0.131	0.0775	0.0877	0.0639	0.0490	0.0656
N	79140	79136	79140	79130	79139	79140

6.3 Testing nonlinear effects

Assessing the potential nonlinearity in the relationship between immigrant concentration and school performance is crucial, as it offers important evidence on the nature and dynamics of the underlying mechanisms governing social interactions in educational settings. To do so we introduce the quadratic term of the non-native share and we regress the following model:

$$y_{ist} = \alpha + \beta M_{is} + \gamma M_{is}^2 + \delta X'_{is} + \varepsilon_{its} \text{ with } t = 5, 8, 13$$

Where y_{ist} are the test scores in mathematics and language in 5th, 8th and 13th grades and the set of covariates X'_{is} contains the same controls as in the regressions above. Table 6.22 displays the results obtained by regressing the above equation on the subsample of native students only.

In mathematics, the linear effect of *immigrant share* is small and not statistically significant in 5th grade but becomes significantly negative in 8th and 13th grade. Specifically, a one percentage point increase in immigrant concentration is associated with a -4.4% decrease in 8th grade and -3.1% decrease in 13th grade in mathematics scores. In reading, the linear effect is also negative but statistically significant only at the 5% in 8th and 13th grade, remaining non-significant in 5th grade. According to these observations, immigrant peer effects tend to emerge or intensify during secondary education, particularly in mathematics. On the other side, none of the squared terms are statistically significant across grades or subjects, indicating that there is no strong evidence in favor of non-linear relationship between immigrant concentration and native students' educational outcomes. The estimated effect appears to be approximately linear within the observed range of the data.

However, we observe that the statistical significance of the linear coefficient weakens once the squared term is added to the specification. Such reduction in significance in the quadratic model – as opposed to the simpler linear specification – may be due to multicollinearity between the linear and squared terms, which increase standard errors of both making it harder to detect individual effects. Nevertheless, the coefficient remains somewhat similar in magnitude and sign, suggesting that the underlying pattern remains consistent. These results suggest that there is no strong evidence supporting non-linearities in the relationship of interest, indicating that the linear specification can adequately capture the observed patterns.

These results align broadly with the patterns previously observed in the non-parametric *lowess* plots. In particular, the negative and statistically significant effects of immigrant concentration emerging in 8th and 13th grade, especially in mathematics, are consistent with the downward trends detected in the graphical analysis. The lack of statistical significance for the quadratic

coefficients supports the relatively modest curvatures observed in the *lowess* trends. In general, both the descriptive figures and these regression models indicate that native students tend to perform slightly worst in school with higher share of immigrant peers, and that this association becomes more pronounced as students progress through the school system.

Table 6.22. Non-linear specification that includes the quadratic term of immigrant share.

Dependent variable: log test score of native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Share of immigrants	-0.00710 (0.0151)	-0.0437*** (0.0155)	-0.0311** (0.0136)	-0.0168 (0.0146)	-0.0256** (0.0127)	-0.0270** (0.0136)
(Share of immigrants)²	-0.0491 (0.0331)	0.00417 (0.0422)	-0.0105 (0.0361)	-0.0222 (0.0292)	-0.00310 (0.0340)	0.00531 (0.0348)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.144	0.119	0.118	0.0828	0.0927	0.0994
N	201007	200999	201007	200987	201001	201007

Note: Robust standard errors are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.4 Exploring potential tipping points

After testing for general non-linearities in the relationship between immigrant concentration and native students' performance and not finding robust evidence of it, it becomes relevant to examine whether a *tipping point* may exist. Tipping points imply the existence of a threshold in the share of immigrants after which the effect on the educational outcomes of native pupils becomes more significantly pronounced. Identifying whether a there exists a threshold is crucial for understanding whether the influence of school immigrant composition is stable across different exposure levels or if there is a critical volume of non-native peers beyond which detrimental effects begin or are more strongly identified.

In order to find whether the adverse effects become stronger or more consistently estimated after a certain threshold, we construct five bins of the variable *immigrant share* and estimate separately the effects on test scores of native students. The use of bins allows for a more intuitive test of whether the relationship between immigrant concentration and natives' school

performance intensifies at particular levels of exposure. Table 6.23 shows the estimates of the bin specification. The first bin is used as the baseline, hence why there is no estimated coefficient for it. There is no significant evidence of any a specific tipping point. The estimated coefficients are, on average, always significant at the 1% level. The negative effect of 5th grade immigrant concentration enlarges gradually as we move towards higher bins, but we cannot identify any specific threshold after which the negative effects become significantly larger in absolute value. These figures do not make us to reject the earlier hypothesis that the relationship between early exposure to immigrant peers and natives' performance in school is linear, still they leave space for more detailed research to investigate more nuanced tipping points, that a simple bin specification cannot determine.

Table 6.23. Bin specification.

Dependent variable: log test scores of native students						
	Mathematics			Reading		
	5th grade	8th grade	13th grade	5th grade	8th grade	13th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Effects of the share of immigrants in bins						
Bin [0 – 0.1)
Bin [0.1-0.2)	-0.00397** (0.00163)	-0.0038*** (0.00119)	-0.0049*** (0.00112)	-0.0049*** (0.00169)	-0.0019* (0.00103)	-0.0042*** (0.00123)
Bin [0.2-0.3)	-0.00391* (0.00214)	-0.0077*** (0.00165)	-0.0058*** (0.00151)	-0.00247 (0.00235)	-0.0041*** (0.00139)	-0.0042*** (0.00162)
Bin [0.3-0.4)	-0.00918** (0.00367)	-0.0142*** (0.00245)	-0.0103*** (0.00233)	-0.00531 (0.00377)	-0.0098*** (0.00199)	-0.0087*** (0.00261)
Bin [0.4-0.5)	-0.0190*** (0.00592)	-0.0286*** (0.00418)	-0.0232*** (0.00401)	-0.0217*** (0.00590)	-0.0224*** (0.00404)	-0.0164*** (0.00447)
Bin [0.5-1]	-0.0187** (0.00797)	-0.0249*** (0.00809)	-0.0284*** (0.00762)	-0.0174* (0.00930)	-0.0177*** (0.00666)	-0.0204** (0.00857)
R-squared	0.144	0.119	0.118	0.0829	0.0928	0.0995
N	201007	200999	201007	200987	201001	201007

7 Robustness and sensitivity check

In order to assess the credibility of the main findings of this research, this section presents a series of robustness and sensitivity checks. This section aims to verify whether the estimated effects arising from immigrant concentration are driven by specific modeling choices or variable definitions.

7.1 Placebo regressions

First, any assessment made up to now is based on the idea that we are capturing the effect of sharing classroom and school spaces with non-native students, that most likely have limited proficiency in the mainstream language and that may have different teaching needs. Following this assumption, we estimate the effect of *immigrant share* in 5th grade on school outcomes of one single cohort of students observed contemporaneously in 5th grade, and subsequently in 8th and 13th grade. To test our identification strategy, we perform some placebo regressions using as a treatment variable the immigrant concentration computed in 2nd grade at school level in the same year we observe pupils in 5th grade. Similar placebo regressions were carried out by Gould et al. (2009) as well. Under the assumption that educational outcomes of native students are influenced by their immediate peers, we should find little to no evidence of significative effects of 2nd grade peer composition on 5th grade students' performance. Information on peer composition at school level in 2nd grade were retrieved from INVALSI datasets. As stated in the earliest sections of this work, the Institute tests reading and mathematics skills both in 2nd and 5th grade at the elementary school level. These two datasets, once recovered, were merged together³⁶, and the share of non-native students was computed at school level as it was previously done on the 5th grade dataset. After that, the dataset on 2nd grade and our main dataset were merged with respect to the variable *codice_plesso*, which identifies schools uniquely within the same year. By doing so, each child in the main dataset was assigned with the fraction of non-native students – again only *test-takers* – enrolled in the 2nd grade of their own school. All observations to which it was not possible to assign the *immigrant share* computed in 2nd grade were dropped. Table 7.1 reports the results from the placebo regressions, that we estimate using the sample models as our standard regressions. We observe mostly weakly significative

³⁶ We obtained a sample of 498.107 2nd grade students. Of them, 1.92% were immigrants of first-generation and 9.13% were of second-generation.

effects, which support the idea that our original identification strategy successfully captures the effect of immigrant peer composition on educational outcomes. In addition to that, the estimated coefficients of these placebo regressions are notably small in size, reinforcing the effectiveness of our original estimates. The fact that the estimates of the effect of 2nd grade immigrant concentration are consistently smaller than those obtained using 5th grade share reinforce the idea that our original regressions are capturing the genuine effect of exposure to immigrant peers specifically in 5th grade. The small significance of the placebo coefficients, however, suggests that overall school composition – regardless of grade – may exert a weak influence on school performance. Nevertheless, the larger magnitude and stronger consistency of the 5th grade estimates indicate that contemporaneous peer exposure at the grade level plays a distinct and more substantial role in shaping educational trajectories. In any case, the weakly significant negative estimated coefficients may reflect that there is a residual component of the school context that we fail to observe. As existing literature pointed out, there are other school-level and individual-level that play – such as the teacher per pupil ratio – that shape the extent to which peer effect influence individual behavior and outcomes.

Table 7.1. Placebo regressions.

	Dependent variable: log test score of native students					
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share in 2nd grade	-0.0114* (0.00629)	-0.0279*** (0.00429)	-0.0192*** (0.00404)	-0.0137** (0.00695)	-0.0177*** (0.00358)	-0.00812* (0.00443)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.144	0.119	0.118	0.0826	0.0925	0.0993
N	197489	197481	197489	197470	197483	197489

Note: The treatment is the share of immigrant children computed in 2nd grade at school level. Each native student in the original dataset is assigned to the immigrant share computed in their school as identified by *codice_plesso*. Robust standard errors in parenthesis are clustered at the school level. All specifications include the constant term. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The main interest at this stage of the research is to make sure that our identification strategy and regression models are actually able to measure the causal effect of interest, that is the influence of immigrant students on the performance of their native peers. Placebo regressions are useful instruments to this scope. We decided to estimate another placebo regression,

assigning as treatment to each student the immigrant concentration computed for the child five rows ahead in the dataset, and the results are reported in Table 7.2. We observe no statistically significant effect of the placebo treatment on school outcomes of native students, as a further support to the effectiveness of our model. The significant effect of the placebo treatment in 8th grade on math test scores is most likely due to the remaining correlation existing between the placebo treatment and the original *immigrant share*. Because of the way in which the placebo treatment was assigned to the students in our dataset, there are a number of cases in which children were assigned their exact original treatment (think of the circumstance in which two children enrolled in the same school are few rows away in the dataset). In any case, we believe that these results point towards the effectiveness of our original estimates.

Table 7.2. Placebo regressions.

Dependent variable: log test score of native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo treatment	-0.00515 (0.00462)	-0.0143*** (0.00408)	0.000507 (0.00387)	-0.00303 (0.00478)	-0.00586 (0.00363)	0.00408 (0.00405)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.144	0.119	0.178	0.0827	0.0925	0.208
N	201002	200994	201002	200982	200996	201002

Note: the treatment is the share of immigrants computed for the student five rows ahead in the dataset. Robust standard errors in parenthesis are clustered at the school level. All specifications include the constant term. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.2 Non-random sorting across schools

One great concern of research on peer effects is the non-random sorting of students across schools. In most cases, school choice is driven by residential proximity, so that the primary sort of bias is induced by residential choices of households. As it was extensively addressed earlier in this work, families choose where to live depending on housing costs and job availability, so that they endogenously sort depending on their socio-economic background. As a consequence, we are likely to observe that families with similar cultural and socio-economic status cluster in the same areas, and their children go to the same schools. This issue was addressed in our identification strategy with fixed effects at province level. Some concern however remains that it might not be enough to fully address the issue. In particular, there are other unobserved and

usually unobservable factors that guide school choice. The so-called *native-flight* and overall self-selection represent a substantial challenge. Native family may purposefully enroll their children in schools with low immigrant share, if they expect their presence to be detrimental to the academic success of their own children. This, in addition, can exacerbate school segregation. Such phenomena are obviously more likely to happen in big geographical areas, where the number of different schools is sufficient to allow a certain degree of choice to families. On top of that, large, urbanized areas usually public transportation that facilitates the commuting to other schools. Tonello (2016) suggests the use of an interaction term between the share of immigrants in junior high school and a dummy for small municipalities to test if the identification strategy is susceptible to the endogenous placement of students in schools. If there was unaddressed school-sorting, we should observe a downward biased estimation of the coefficient of interest. as it would induce a spurious correlation between natives' test scores and non-native shares. We adapt this strategy to our framework, and we sort provinces into two groups, *small* and *large*. We count the number of unique primary schools³⁷ in each province and find the median. Unlike Tonello (2016), we build our two categories on the basis of elementary school because that's the level at which we computed our treatment. The smallest province in our dataset is Aosta (Valle d'Aosta) with only 8 different schools. The largest one is Rome (Lazio) with 464 unique schools. We use the median as the threshold that distinguish small and large provinces, so that we obtain two balanced groups. We then created a dummy equal to 1 for schools located in small provinces, identified as those with less than 37 unique primary schools. We then estimate the following model:

$$y_{its} = \beta_0 + \beta_1 M_{is} + \beta_2(\text{small province} * M_{is}) + \beta_3 X'_{is} + \varepsilon_{its} \text{ with } t = 5, 8, 13$$

We use fixed effects at region level rather than at province level, to ensure variation within the geographical unit. The interaction terms in Table 7.3 are consistently statistically insignificant in most specifications, with the exception of column 4 where a weakly significant negative interaction ($p < 0.1$) is detected for reading scores in 5th grade. These results suggest that peer effects to not differ systematically between small and large provinces. If school-sorting was

³⁷ We count the number of schools and not the number of *plessi*. The term *plesso* refers to the building in which teaching take place for a certain student. One school can have multiple *plessi*. Any potential sorting would take place at the school level rather than the *plesso* level, as parents can choose in which school to enroll their children but not in which building they should be placed.

taking place in larger urban areas – where families have larger degree of school choice – we would expect to find significantly smaller and negative effects in smaller provinces. The absence of statistically significant effects implies that the sorting dynamics that may take place at province level are unlikely to be a major source of bias in the main estimates. In other words, our original estimates of peer effects are not likely to be driven by differential sorting patterns across areas with different degrees of school choice. The robustness of the results across different provincial contexts adds credibility to our interpretation that early exposure to immigrant concentration has a lasting impact on native students’ performance, in particular in 8th grade. We observe that the estimated coefficient of *immigrant share* here is larger than the those we initially found. However, this only suggests that peer effects may be more pronounced in these areas, while the absence of significant interaction terms strongly confirm that the difference across provinces is not statistically robust.

Table 7.3. Robustness check.

Dependent variable: log test score of native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share	-0.0249*** (0.00715)	-0.0329*** (0.00622)	-0.0267*** (0.00949)	-0.0268*** (0.00759)	-0.0191*** (0.00504)	-0.0162* (0.00893)
Interaction term	-0.0148 (0.0167)	0.00583 (0.0121)	0.000973 (0.0114)	-0.0299* (0.0173)	-0.000598 (0.0104)	-0.0117 (0.0120)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.144	0.119	0.178	0.0829	0.0927	0.208
N	201007	200999	201007	200987	201001	201007

Note: The models include the interaction term between the share of non-native students and a dummy for small provinces. By these means, we test whether there is unaddressed non-random school sorting, under the assumption that it is most likely to happen in larger provinces. Robust standard errors in parenthesis are clustered at the school level. All specifications include the constant term. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Using fixed effects at the region level cause loss of granularity of information and may still not be a reliable source of information. To prove the validity of our findings, we regress our specifications separately on children attending primary school in small provinces and those attending school in large provinces (Table 7.4 and Table 7.5). We do not observe any substantial difference between the effect of early immigrant exposure in small and large provinces, supporting the idea that sorting at the primary school level is not biasing our main estimates.

Table 7.4. Robustness check.

Dependent variable: log test score of native students (small provinces)						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share	-0.00944 (0.0141)	0.0434*** (0.0128)	0.0432*** (0.0124)	-0.0124 (0.0149)	0.0330*** (0.0104)	0.0355*** (0.0125)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.136	0.0574	0.0594	0.0828	0.0526	0.0636
N	47475	47471	47475	47469	47475	47475

Note: Results from regressions on the sample of native children attending primary school in small provinces. Robust standard errors are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7.5. Robustness check.

Dependent variable: log test score of native students (large provinces)						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share	-0.0236*** (0.00769)	-0.0434*** (0.00660)	-0.0365*** (0.00589)	-0.0194** (0.00817)	-0.0267*** (0.00544)	-0.0237*** (0.00645)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.142	0.118	0.123	0.0789	0.0932	0.100
N	153532	153528	153532	153518	153526	153532

Note: Results from regressions on the sample of native children attending primary school in large provinces. Robust standard errors are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

However, the largest sorting is most likely to happen at higher level of education, starting from lower secondary school. While primary school choice is more likely to be exclusively driven by residential proximity – as children need to be taken to school – it may be that selection happens mostly in lower and surely in upper secondary school. Focusing on lower secondary school, we still expect that most sorting is still induced by parents' choice. We repeat the procedure above, but this time we identify small and large provinces on the basis of the number of unique lower secondary schools. This time, the smallest province, with only 1 upper secondary school, is Oristano (Sardegna). The largest one is again Rome (Lazio), with 385 schools. The threshold for large province is set at 75 schools. Fixed effects are at the region

level to ensure variability within the unit. As it can be observed in Table 7.6, the interaction terms are again not statistically significant in most specifications (exception made for column 6), thereby giving additional validity to our initial results. We do not find evidence that our original estimates are driven by sorting across schools at primary and lower secondary level.

Table 7.6. Robustness check.

Dependent variable: log test score of native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share	-0.0277*** (0.00819)	-0.0238*** (0.00646)	-0.0298*** (0.00615)	-0.0332*** (0.00884)	-0.0125** (0.00544)	-0.0157** (0.00657)
Interaction term	0.00920 (0.0132)	0.00981 (0.0110)	-0.00263 (0.00922)	-0.0110 (0.0139)	0.00296 (0.00921)	-0.0244** (0.00965)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.144	0.119	0.178	0.0828	0.0927	0.208
N	201007	200999	201007	200987	201001	201007

Note: These models include the interaction term between the share of non-native students and a dummy for small provinces (identified depending on the number of lower secondary schools). Robust standard errors in parenthesis are clustered at the school level. All models include a constant term. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We do not test in the same way the potential non-random sorting of students across upper secondary schools. Tracking in Italy becomes at the end of high school, so that we know for sure that there is a certain level of self-selection bias. Students with higher individual ability, better socio-economic and parental background or greater attitude towards school and education are most likely to attend *liceo*, which offers general education aimed primarily at preparing students for university education. These schools emphasize theoretical knowledge across different disciplines such as sciences or humanities. Otherwise, students can decide to attend technical institutes (*istituto tecnico*), that combine general education with technical training in a try to equip students with practical skills, or vocational institutes (*istituto professionale*), that – as the name suggests – offer vocational and job-specific training. While we cannot specifically control for track selection in the original models – because the choice of school track could itself be an outcome of early immigrant exposure – here we estimate our regressions on the stratified sample by type of upper secondary school attended (Table 7.7). The results obtained from this stratified analysis by type of upper secondary school attended reveal substantial heterogeneity in the estimated effect of immigrant concentration. The magnitude, direction and statistical significance of the coefficients vary considerably across tracks.

For children who eventually enrolled in a *liceo* with a scientific track, the estimates are generally positive and statistically significant in both mathematics and language. These results are in sharp contrast with the negative and significant coefficients found in the overall sample, suggesting that students in this academic track may experience different dynamics of peer interaction. One possible explanation is that students in *liceo scientifico* – who are positively selected in terms of prior achievement, motivation and often socio-economic status – might be less susceptible to any potential adverse peer externalities or may even benefit from a diverse peer environment. A closely related possible interpretation, is that students with high individual ability are not extensively affected by their peers, as cognitive ability plays the major role in shaping their educational outcome. This is not uncommon evidence in research on peer effects. Lavy, Paserman, et al. (2012) find that the proportion of low-achieving students has no effect on high-ability students³⁸. Sacerdote (2001) suggests that high-ability students in college are less sensitive to the academic background of their peers. We test again the null hypothesis that the coefficient of *immigrant share* is equal across models, to state the trajectory of these effects over time for this subgroup of students. We find that the coefficients in columns 1 and 2 are significantly different ($\chi^2 = 4.37, p = 0.0366$), as well as those in columns 2 and 3 ($\chi^2 = 5.10, p = 0.0239$). As for reading, we find the coefficients in columns 4 and 5 not to be statistically different ($\chi^2 = 0.36, p = 0.5469$), while those in columns 5 and 6 are ($\chi^2 = 4.93, p = 0.0264$). We derive that high-ability students are slightly positively influenced by the presence of immigrant peers, and that this influence can hold up to the end of upper secondary school.

Among students enrolled in other types of *liceo* in 13th grade, the estimated coefficients are mostly negative but still smaller in magnitude than the original ones and often not statistically significant. Exception is made on mathematics test scores in 8th and 13th grade, on which the effect is negative and highly significant. This pattern reflects a more mixed population of students in these schools in terms of individual ability, that leads to a more ambiguous dynamics of peer effects.

³⁸ They consider repeaters to be students with low cognitive ability on average. They regress the proportion of repeaters in a grade on the average score and other educational outcomes of students divided into three groups, depending on the number of advanced courses they are enrolled into. They find that the proportion of repeaters only negative affects regular students (considered to be those that are not enrolled in any advanced course), while there is no statistically significant effect on students enrolled in at least 2 advanced courses or in more than 3 advanced courses.

As for those attending technical institutes, the coefficients are consistently negative and mostly statistically weakly statistically significant, especially in mathematics. This suggests that early exposure to a higher share of non-native peers may have a modest adverse effect on achievement of children that ultimately decide to follow this track, potentially due to the cumulative disadvantage in prior academic preparation.

Finally, the strongest and most significant effects are observed among students enrolled in vocational institutes, where the immigrant share is significantly and negatively associated with both mathematics and reading scores in all grades (exception made for the specification in column 6). These results are not unexpected, as students in this type of school tend to come from more disadvantaged backgrounds and may attend schools with lower overall academic performance, amplifying any negative peer externalities. These estimates suggest that students who eventually enroll in professional institutes are negatively affected by early exposure to immigrant concentration up to the end of upper secondary school. When testing the null hypothesis that the coefficients across models are statistically identical, we do not reject it with respect to coefficients in columns 1 and 2 ($\chi^2 = 1.56, p = 0.2113$) and to those in columns 2 and 3 ($\chi^2 = 0.44, p = 0.5084$). The same holds for coefficients in columns 4 and 5 ($\chi^2 = 3.37, p = 0.0663$) and those in columns 5 and 6 ($\chi^2 = 2.67, p = 0.1025$). According to these results, we conclude that students with lower individual ability or lower socio-economic and parental background are strongly and negatively affected across their whole education by the presence of non-native peers.

Taken together, these findings provide compelling evidence that the effect of immigrant concentration is not homogeneous across educational tracks. These results underscore the importance of accounting for educational trajectories and selection mechanisms when evaluating the long-term impact of peer composition. Also, they align with most existing literature that finds lower-achieving pupils to be those strongly and usually negatively affected by the presence of other possibly low-achieving students. From a robustness perspective, stratifying the analysis by type of upper secondary school serves as a valuable sensitivity check. By disaggregating the sample, we observe that the effects vary substantially across more internally homogeneous groups of students. The fact that estimates differ markedly, and even reverse in sign in some cases, suggests that selection into educational tracks plays a non-trivial role in shaping long-term outcomes. This stratified analysis highlights important heterogeneity that remained masked in the baseline model, adding nuance to the interpretation of peer effects and exposing the underlying structure of selection and its consequences.

Table 7.7. Stratified analysis by school type in upper secondary education.

Dependent variable: log test scores of native students						
Liceo (scientific track)						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share	0.0230** (0.00921)	0.00505 (0.00725)	0.0205*** (0.00759)	0.0218** (0.0104)	0.0161** (0.00661)	0.0334*** (0.00941)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.134	0.131	0.146	0.0702	0.0792	0.0927
N	58163	58163	58163	58161	58160	58163
Liceo (other)						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share	-0.00903 (0.00916)	-0.0248*** (0.00731)	-0.0207*** (0.00709)	-0.00810 (0.00932)	-0.00957 (0.00669)	-0.00240 (0.00771)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.143	0.104	0.0992	0.0732	0.0720	0.0733
N	65262	65257	65262	65256	65261	65262
Technical institutes						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share	-0.0159* (0.00921)	-0.0237*** (0.00741)	-0.0159** (0.00746)	-0.0128 (0.00982)	-0.0118* (0.00694)	-0.0136 (0.00894)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.154	0.142	0.158	0.0771	0.0877	0.106
N	58336	58334	58336	58330	58334	58336

Professional institutes

	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share	-0.0700*** (0.0145)	-0.0529*** (0.0109)	-0.0454*** (0.0118)	-0.0595*** (0.0155)	-0.0319*** (0.0111)	-0.0122 (0.0132)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.171	0.110	0.114	0.0926	0.0792	0.102
N	19246	19245	19246	19240	19246	19246

Note: Stratified analysis depending on the type of school attended at upper secondary level. Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample of native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.3 The issue of attrition

During the initial merging process, we lost important observations on children that for any reason are not present in the 8th grade and 13th grade waves of 2017-2018 and 2022-2023 respectively. It could be students who were held back a grade at some point during their educational career, students that dropped out of school, or that moved to a private school in which INVALSI assessments are not conducted. Trivially, it could also simply be students who were sick on the day of INVALSI so that they stayed at home and missed the tests. One additional case is students who intentionally missed the tests, because of poor attitude towards education or even because they were suggested to do so. Attrition represents a great challenge to this work, especially if we expect it to be driven by the exposure to immigrant peers early in the school career.

The first aspect we can check is if students who are missing in 8th grade and 13th grade dataset are systematically different from those who remain in it. Table 7.8 shows some important information over baseline characteristics of students, separately for the sample used in our analysis, the group of students lost after 8th grade and that of students lost after 5th grade. As it can be observed, students in these groups differ substantially, especially in terms of origin and parental education. While only 7.11% of students in our final sample are first-generation immigrants, this share grows up to 12.55% when considering those students that we fail to observe after the 5th grade INVALSI wave. Also, the share of students with at least one parent with higher education drops among both groups of students who fall out of the sample at some point in time. Ultimately, we find that these students differ substantially in important

characteristics that affect school performance. Table 7.9 shows the average of the *immigrant share* value across the three groups. It is possible to see that these groups differ in the way that they were exposed early in their educational life to immigrant peers. When tested, we found that the mean of *immigrant share* is statistically different across the three groups, and that about 1.85% of variation in the share of immigrants in 5th grade is explained by group membership. While it is not a large value, it is still a figure we must take into account.

Table 7.8. Characteristics of sampled students, of students lost after 8th grade and of students lost after 5th grade.

	Students who remain the sample		Students lost after 8 th grade		Students lost after 5 th grade	
	Stock	%	Stock	%	Stock	%
Number			97.285		34.505	
Who are immigrants	15.407	7.11%	16.889	17.36%	11.162	32.34%
Of which of first generation	4.027	1.86%	5.563	5.72%	4.329	12.55%
Who are female	115.309	53.28%	39.985	41.10%	13.585	39.37%
Who have at least one parents with higher education	65.356	30.20%	12.595	12.9%	4.391	12.72%
Who have at least one parent employed	198.694	91.81%	74.298	76.37%	23.709	68.71%
Who are a grade behind	2.615	1.21%	4.617	4.28%	3.662	10.61%

Table 7.9. Average of immigrant share exposure across the three groups.

	Students who remain the sample		Students lost after 8 th grade		Students lost after 5 th grade	
	Mean	Std.err.	Mean	Std.err.	Mean	Std.err.
Average immigrant share	0.0988	0.0001	0.1276	0.0004	0.1436	0.0007

To understand to what degree attrition can affect our main results – that is to say to understand whether the immigrant share in 5th grade is itself a determinant of attrition – we perform a logit model to estimate the probability of not being in the sample in 8th grade or in 13th grade separately (Table 7.10), given the exposure to immigrant peers in primary schools. In doing so,

we hold on to our main identification strategy, so we include fixed effects at province level and cluster errors at the school level. We create two dummies, *attrition8*, equal to 1 if we lose the observations in 8th grade, and *attrition13*, equal to 1 if we lose it in 13th grade. When estimating the logit model of the effect of *immigrant share* on the two dummies alone, we find that early immigrant exposure substantially predicts the likelihood of dropping out of sample. However, as soon as we control for important individual characteristics such as gender or parental background, this effect shrinks. The main takeaway is that, while early immigrant exposure is indeed an important determinant of our sample, there other factors that play an important role in shaping the behavior of students in school.

Table 7.10. Logit estimates of the effect of immigrant share on the likelihood of not being observed after 5th grade or after 8th grade given the share of immigrants in 5th grade.

	Probability of not being observed after 5 th grade		Probability of not being observed after 8 th grade	
	(1)	(2)	(3)	(4)
Share of immigrants	2.576*** (0.0876)	1.985*** (0.0961)	1.290*** (0.0503)	1.051*** (0.0479)
Female		Yes		Yes
At least one parent with high education		Yes		Yes
At least one parent employed		Yes		Yes
Delayed grade placement		Yes		Yes
Fixed effects	Yes	Yes	Yes	Yes
N	442312	442312	442312	442312

Note: We estimate separately the likelihood of not being observed after 5th grade and not being observed after 8th grade. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The fact is that, while this is an important aspect to consider when discussing our results, there is not much that can be done for a few reasons. The first reason is that we have no information on the actual reason why we do not observe some students in the 8th grade and 13th grade INVALSI waves we employ. Of course, a student absent on that day for health reason is different from a student that was held back a grade at some point and is different from a student that, for example, moved abroad.

In order to assess whether our estimates remain valid, we adopt an IPW procedure. By these means, we assign to each observations the likelihood that they stay within the sample in 8th grade and 13th grade based on their observed individual characteristics. We then estimate again our main regression, but this time weighting each observation by the inverse of the probability

that they stay in the sample. By doing so, we observe in Table 7.11 results that are quite similar in magnitude, direction and significance to our original estimates of the effect of immigrant concentration. This reinforces the validity of our results, suggesting that they were not extensively driven by sample selection or unobserved attrition-related factors. The main concern that attrition may distort our findings is not supported by evidence.

Table 7.11. Results of weighted regressions.

Dependent variable: log test score of native children						
	Mathematics			Language		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Share of immigrants	-0.0239*** (0.00878)	-0.0363*** (0.00562)	-0.0301*** (0.00544)	-0.0237*** (0.00798)	-0.0171*** (0.00478)	-0.0205*** (0.00573)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.0426	0.115	0.118	0.0329	0.0914	0.104
N	201007	200999	201007	200987	201001	201007

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample of native children only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.4 Testing different samples

Another common approach to robustness checks involves testing the model across different subsamples. By estimating the main specification separately for distinct groups – such as geographic regions or socio-economic strata – it is possible to verify whether the initial results are consistent across varying contexts. This method helps to rule out the possibility that findings are driven by specific segments of the sample. In addition to that, it can also reveal meaningful and initially undetected heterogeneity in the strength or direction of the estimated effects.

First, we decided to estimate the model separately for each main geographic area. We considered the four areas identified in the INVALSI dataset according to Istat distinction³⁹: north-west, north-east, center and south. North-west area includes Valle d’Aosta, Piemonte, Liguria and Lombardia. North-east includes Veneto, Friuli-Venezia-Giulia, Emilia-Romagna, Bolzano and Trento. Regions in the center and in the south are respectively Toscana, Umbria, Marche, Lazio and Abruzzo, Molise, Basilicata, Calabria. Table 7.12 presents the estimated

³⁹ We group south and islands together because of the reduced number of observations from these areas.

effects of the share of immigrant peers in 5th grade on natives' school performance disaggregated at macro-region level. These results indicate that negative peer effects that derive from early immigrant exposure are not uniform across the country, but are rather geographically concentrated, particularly in the northern regions. In the north-west, immigrant share in 5th grade is associated with significantly lower test scores in both subjects and across all grades. The observed pattern is similar to the one highlighted in the main regressions. In the north-east, the estimated peer effect is smaller and only marginally significant in both mathematics and language. The coefficient in mathematics remains consistently significant through 13th grade, whereas the effect on language achievement fades entirely by 8th grade. In the center we observe the largest effect in absolute value in 8th grade on mathematics test scores, suggesting an amplification of peer effects during lower secondary school, as identified earlier. As observed for north-west regions, the effects of early immigrant exposure persist up to 13th grade. In contrast, no statistically significant effect is detected in the southern regions. Here, coefficients are small and imprecisely estimates, with larger standard errors and no consistent pattern overtime. However, it's important to remember that we are missing a consistent portion of observations from this area⁴⁰. Taken together, these findings suggest that the impact of early exposure to immigrant peers is highly context dependent. The negative effects appear to be driven by the northern area of Italy. The fact that significant peer effects of immigrant concentration are found in North and Center, while the south shows no robust effects, may reflect important structural differences in the distribution and integration of immigrant students across Italian regions. In the beginning of this work, we emphasized that immigrant population in Italy is more densely concentrated in the northern regions, where there are usually better economic opportunities and migration inflows have historically been higher. This great variation in immigrant shares across Italy suggest that there may be large differences in the degree to which native students are exposed to immigrant concentration. The low share of immigrant population in the southern regions of Italy can reduce the visibility of any peer dynamics (Table 7.13).

Moreover, initial calculations indicated that school segregations levels were higher in the south. Following that there are two important implications. First of all, the fact that immigrant students are apparently concentrated in a small number of schools in southern regions implies that the

⁴⁰ To deal with the issue posed by having information only on *test-takers*, rather than the actual population of students enrolled in Italian schools, we decided to drop observations from the regions of Campania, Puglia, Sardegna and Sicilia, who had very low participation rates.

majority of native students there have little to no contact with immigrant peers. In this case, estimating peer effects becomes statistically difficult as the relevant treatment is not sufficiently distributed. Second, higher segregation rates may generate larger selection bias in the composition of schools. In other words, students who attend schools with larger immigrant share in the south may differ systematically from those who are not exposed at all, making it harder to identify the peer effects of interest. Therefore, the lack of significant effects in these regions is likely to be constrained by the specific mechanisms of exposure and integration, suggesting that south Italy might not serve as a good proxy for everyday peer interaction because of the greater school segregation.

Table 7.12. Testing the model on geographical areas.

Dependent variable: log test score of native students only						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
North-west	-0.0307*** (0.0102)	-0.0548*** (0.00743)	-0.0386*** (0.00718)	-0.0268** (0.0104)	-0.0332*** (0.00637)	-0.0314*** (0.00832)
N	72388	72386	72388	72386	72386	72388
North-east	-0.0230** (0.0113)	-0.0235** (0.0111)	-0.0240** (0.00973)	-0.0216* (0.0118)	-0.0116 (0.00874)	-0.00848 (0.0103)
N	53038	53038	53038	53036	53037	53038
Center	-0.0309** (0.0156)	-0.0605*** (0.0107)	-0.0471*** (0.0108)	-0.0359** (0.0171)	-0.0433*** (0.00976)	-0.0332*** (0.0113)
N	52801	52800	52801	52789	52798	52801
South	-0.0167 (0.0310)	0.0191 (0.0232)	-0.0192 (0.0226)	-0.00156 (0.0377)	0.00154 (0.0238)	-0.0301 (0.0272)
N	22780	22775	22780	22776	22780	22780
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on the subsample on native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7.13. Mean value of immigrant share by geographical area.

Area	Mean (share of immigrants)	Std. dev.
North-west	0.129	0.112
North-east	0.138	0.109
Center	0.109	0.967
South	0.543	0.000

As suggested in the previous section, another interesting distinction can be made on the basis of socio-cultural background. While the main variable that define students' background are already included in our main model, we follow the idea of Contini (2013) and repeat separate regressions on subsamples of native children identified on the basis of the average number of books they have at home. The number of books at home is a common proxy usually employed in research on peer effects as a measure of the socio-economic and parental background. This variable theoretically captures both material and non-material aspects of the family environment and the household's characteristics. The underlying idea is that the number of books is a valuable measure of access to cultural capital, and it can signal the extent to which home environment supports learning and academic engagement. Brunello & Rocco (2013) include the number of books at home – together with the share of students with at least one parent having obtained a degree or anything above – as proxies of the average parental background. Brunello et al. (2017) find strong association between books-at-home and school outcomes, proving the efficiency of this variable in embodying cultural capital. Table 7.14 illustrates the results obtained on the different subsample of native children. We do not observe significant effect on the educational outcomes of children with very few books at home, as for those with a number of books at home sufficient to fill more than two shelves. The largest and most significant effects are identified for students with few books or some books at home. The estimated coefficients follow a trajectory over time similar to the one observed in our initial estimates. If we consider the number of books at home to be a consistent proxy of parental background and cultural capital, and if we believe that these are factors that play a pivotal role in the educational outcomes of pupils, then we find evidence that, as already confirmed in literature, peer effects tend to be more consistently estimated for average-achieving students. High-achieving students or those who come from particularly benevolent socio-economic backgrounds are usually more immune to the influence of their peers.

Table 7.14. Sensitivity check. Testing the model on samples of native students identified by the number of books at home.

Dependent variable: log test scores of native students						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Effect of share of immigrant on children with:						
Very few books (0-10)	-0.00747 (0.0189)	-0.0217 (0.0182)	-0.0198 (0.0182)	-0.00510 (0.0199)	-0.0296* (0.0159)	-0.00276 (0.0209)
Few books (11-25)	-0.0185* (0.0111)	-0.0442*** (0.00909)	-0.0302*** (0.00916)	-0.0309*** (0.0113)	-0.0246*** (0.00842)	-0.0302*** (0.0101)
Some books (26-100)	-0.0272*** (0.00849)	-0.0414*** (0.00746)	-0.0397*** (0.00783)	-0.0287*** (0.00899)	-0.0208*** (0.00672)	-0.0197** (0.00824)
Enough books to fill two shelves (101-200)	-0.00895 (0.0104)	-0.0295*** (0.00914)	-0.0222** (0.00931)	-0.0125 (0.0108)	-0.0166** (0.00830)	-0.0155 (0.0104)
Enough books to fill three shelves or more (more than 200)	0.00219 (0.0118)	-0.0133 (0.00987)	-0.00350 (0.0108)	0.0231* (0.0122)	-0.00736 (0.00924)	0.00910 (0.0118)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	201007	201007	201007	201007	201007	201007

Note: Robust standard errors in parenthesis are clustered at the school level. All specifications include a constant term. The regressions are estimated on native students only. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Another important robustness check consists of excluding potential outliers from the analysis. The matter with outliers is that they may exert disproportionate influence on the estimated coefficients. In particular, schools with exceptionally high or low immigrant concentration are likely to have specific unobserved characteristics that explain their unusual share of non-native students. For instance, schools that present immigrant concentration equal to zero may be located in exceptionally wealthy area, making them inaccessible to immigrant children. In addition to that, native pupils that attend these schools may have an uncommon and surprisingly high socio-economic and cultural background that variables related only to whether parents

have a job or whether they have a degree may not be able to capture. Another aspect to consider is related to that of the *test-takers issue*. As explained in the earliest sections of this work, the larger limitation of our dataset is that it conveys information only on children that attended school on the days of the INVALSI tests. We do not know anything about students that for any reason were absent from school on these days. The main concern is on the phenomenon of *strategic absenteeism*, which has been documented in standardized testing contexts, including in Italy. Strategic absenteeism refers to the intentional or informal practice by which teachers may suggest certain students – typically those expected to perform poorly – to stay at home on the days of the standardized assessments. Even though INVALSI are not formally high-stakes tests – meaning that no penalization is foreseen in the case of performance below average – there is still some anecdotal and empirical evidence that some schools engage in such practices to artificially inflate their average test scores. Considering that, there are a number of cases that must be considered. It may be that the immigrant concentration computed as zero in our dataset is not measuring the actual share of non-native students, as immigrant students may be initially less academically oriented and exhibit strategic absenteeism, either due to greater performance pressure, resource constraints or administrative discretion, thus creating bias in the computation of our treatment.

Schools with extremely large immigrant shares may as well be problematic. For these reasons, we repeated our estimation excluding from the sample the top and bottom 1% of the distribution of *immigrant share*. This led to deleting 2,117 observations from the dataset. Table 7.15 shows the estimated coefficients of interest in the sample cleaned of outlier values. To allow for straightforward comparison between Table 7.15 estimates and the results of the original baseline model, Table 7.16 reports together the original estimates of the baseline model and the trimmed estimates obtained on the reduced sample. The large similarity between the different estimates suggests that our results are not driven by few extreme values, supporting the external validity and generalizability of our findings.

Table 7.15. Robustness check.

Dependent variable: log test scores of native children						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Immigrant share	-0.0240*** (0.00770)	-0.0439*** (0.00546)	-0.0337*** (0.00516)	-0.0236*** (0.00804)	-0.0288*** (0.00467)	-0.0258*** (0.00569)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.144	0.119	0.118	0.0827	0.0927	0.0994
N	199768	199760	199768	199748	199762	199768

Note: The regressions are estimated on the sample trimmed of outliers. Robust standard errors in parenthesis are clustered at the school level. All models include a constant term. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7.16. Comparison of original estimates of baseline models and trimmed estimates on the reduced sample.

Dependent variable: log test score of native children only						
	Mathematics			Reading		
	5 th grade	8 th grade	13 th grade	5 th grade	8 th grade	13 th grade
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Baseline model</i>						
Share of immigrants	-0.0265*** (0.00684)	-0.0421*** (0.00566)	-0.0193*** (0.00471)	-0.0256*** (0.00720)	-0.0268*** (0.00471)	-0.00143 (0.00496)
<i>Reduced sample</i>						
Immigrant share	-0.0255*** (0.00691)	-0.0418*** (0.00572)	-0.0198*** (0.00477)	-0.0255*** (0.00724)	-0.0266*** (0.00476)	-0.00175 (0.00503)

7.5 Other educational outcomes

As much as the focus of this research was focused over estimation of long-term effects of early immigrant exposure on native students' performance, there are other interesting educational outcomes that are often considered in literature. In particular, apart from the direct effect on test scores, it is interesting to understand to what extent the immigrant share in 5th grade affects tracking decisions. We estimate the effect of immigrant concentration in 5th grade on the probability that the student is enrolled in a lyceum in 13th grade. To do so, we employ all our controls and keep adopting fixed effects at the province level. Table 7.17 suggests that early immigrant exposure negatively affects the likelihood of attending a lyceum the upper secondary

school. In other words, children that experience larger immigrant concentration in primary school are less likely to engage with general education, as opposed to vocational education offered in technical and professional institutes, when controlling for important individual characteristics. While we do not expect this to be the fundamental evidence of adverse effects arising from 5th grade immigrant concentration on educational trajectories of native children, it still adds validity to our findings that early immigrant exposure has weakly negative effects on the educational performance of native children, that does not fade over time but rather last consistently up to the end of 13th grade.

Table 7.17. Estimate effect of immigrant share in 5th grade on the likelihood of being enrolled in a lyceum in 13th grade.

	Probability of being enrolled in a lyceum in 13th grade
	(1)
Share of immigrants	-0.778*** (0.0860)
All controls	Yes
Fixed effects	Yes
N	201007

Note: Robust standard errors in parenthesis are clustered at the school level. The specification includes a constant term. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

8 Discussion of results and conclusions

During the development of this work, our main interest was directed at understanding how peer effects that arise from immigrant concentration in early education evolve over time. In the last years, the interest in measuring and understanding peer effects – especially those exerted by non-native peers – increased substantially, but there is still an undeniable lack of information on the stickiness of such effect, where for stickiness we intend their capability to persist over time. The fact is that, if we claim to be concerned about the influences of peers in school, we must take into account how these evolve across the educational life of students. When dealing with non-native concentration, it becomes fundamental to grasp fully the specific trajectories of peer effects, especially if any inference on them is intended at shaping potential school policies that address peer composition. Existing literature on immigrant peer effects observe at times contrasting results, showing that in most cases findings are highly dependent on identification strategy adopted, as well as country-specific immigrant pattern, ethnicity of non-natives and other factors – some even difficult to measure – that make it difficult to fully capture the entire story.

In addition to that, research on peer effects must address specific issues of non-random distribution of non-natives. Immigrants are unlikely to settle in a geographic area at random, and their choice is rather influenced by a multitude of aspects. For example, they might move to a specific region to reunite with their family, or because they know that there exists a large community of their own culture and ethnicity in that area. Their decision is likely also to be constrained by costs of life, housing costs, job availability, and in some particular cases even by proximity to reception centers. At the same time, we do not expect children to enroll into schools and be placed into specific classes randomly. School choice, in particular at primary and lower secondary level, is mostly driven by residential proximity. The consequence of that is that we tend to observe individuals similar in terms of socio-economic and cultural background clustered in the same areas, and hence their children being in the same schools. The risk that arises from this non-random distribution is that of overestimating any effect deriving from immigrant concentration. When native and non-native pupils from a poorer socio-economic background – which is usually a good predictor of their school performance – cluster in the same schools, we will tend to mechanically observe that students exposed to higher immigrant concentration suffer adverse effects to their educational performance, whereas students in areas with less immigrant exposure will be sort of immune of these effects. The non-random distribution of children across schools represents the greatest challenge in all research on peer effects. In this work we adopt fixed effects at province level, to account for the fact that

immigrants do not randomly decide where to reside. We acknowledge that provinces are still large geographical areas, and there likely is some residual heterogeneity within them; however, that is the smallest geographical unit reported in the datasets offered by INVALSI. In any case we believe economic situation and industries presence to be homogeneous enough within province, and our first results show that province-level fixed effects are enough to consistently estimate peer effects. A second concern was related to the potential endogenous placement of non-native children in classrooms within a school. In Italy, there are no specific rules related to classroom composition. The Italian government only in recent years published guidelines – that as such are not mandatory – according to which no more than 30% of a class should be on non-Italian citizenship, unless they are proficient in Italian. Still, it may be that administrators purposefully place children in specific classrooms considering teacher quality and available resources. When it comes to immigrant children who enroll in school after first grade – when classes have already been made – it may even be that their decision is driven by the classroom predetermined peer composition itself. If they expect non-native children to be disruptive and want to limit any potential damage on high-performing students, it is probable that such pupils will be placed in already poor performing classes. Or, in an opposite instance, it may be that they are placed into classes with exceptionally good teachers, that can offer them tailored support to facilitate entrance into the Italian school system. Since we cannot properly control for these aspects, we decided to compute the immigrant share at school level, rather than at class level. By doing so, we renounced to some granularity of information in a hope to effectively measure the exposure to immigrant peers of native children in primary school. Nevertheless, the concentration of immigrant share at the school level still offers important insights, as it can encapsulate the shape of the overall learning environment in which children begin to develop their educational skills. Despite most peer interactions happening at classroom level, it is clear that one school with higher concentration of immigrants overall differs from one with just a few.

We used 5th grade (2014-2015), 8th grade (2017-2018), and 13th grade (2022-2023) INVALSI waves to assess natives' performance in school. Our outcomes of interest are test scores in both mathematics and reading subjects. In order to test long-term effects, we needed to be able to follow the same students over time. The INVALSI Institute offers the possibility to do so since the introduction, in 2013, of the SIDI code which uniquely identifies students across years.

Our initial estimates tell that there exists an ongoing negative effect of early immigrant exposure that lasts up to the end of upper secondary school. When testing if the coefficients across grade-level specifications were statistically equal one another, we observe that they are not

significantly different across different grade specifications, meaning that these effects are persistent over time and do not tend to fade. Nevertheless, the effects on math performance differ from those on reading performance of native students. Apparently, immigrant concentration in 5th grade has a larger – in absolute value – negative effect on math test score, though the gap with respect to the effects on reading is contained. The best explanation for these findings is that math skills and knowledge are cumulative in their nature, and gaps in early concepts can make it difficult to catch up with new knowledge in the subsequent years of education. It is likely that in math, where precise terminology is essential, peers with limited language proficiency can provide less productive spillovers. Moreover, language skills can be easily contextual influenced, so that native students can improve their knowledge through daily social interactions that take place also out of school – whereas mathematics is usually the result of exclusively formal teaching. What we observe suggest that math performance is more sensitive to school peer composition for native students.

Any further assessment on the long-term impact of peer effects arising from early immigrant exposure assumes that non-native children are likely to have limited language proficiency in at the start of their education and, in some cases, may be less academically oriented. The other side of this assumption is that native students are proficient in Italian language. However, INVALSI defines native students as all children who have at least one parent born in Italy, irrespective of their own place of birth. In other words, there is a portion of native pupils who have migrated to Italy and began school in the country. While it is easy to assume that they are similar to native children who are born in Italy in terms of cultural background, it may be that they are not as proficient in Italian language. To ensure that this was not affecting our estimates, we re-estimate the regressions using a sample of native pupils exclusively born in Italy, excluding all children of Italian parents who may have recently moved to Italy. Even with this restriction, we found similar results and conclude that our findings are not extensively sensitive to the way in which we define native students.

If we were to stop here, we would conclude that all immigrant children are similar in terms of language proficiency and influence their native peers in the same way. Even without specific evidence it is trivial that this cannot be true. We made a distinction between non-native children on the basis of whether they were first- or second-generation immigrants and of their gender. We did not find significant evidence that first- and second-generation immigrants affect native children in a different way. The same holds true when we consider the share of male immigrants and that of female immigrants separately. However, when considering gender, evidence suggest that within-gender groups effects are stronger and more consistently defined. Indeed, when

estimating the same regressions on the two samples of male and female native students one at the time, a different pattern emerges. Male immigrants negatively affect both male and female native students, consistently up to 13th grade. On the other hand, there is no proof that female non-native students exert adverse effects on the performance of their native peers. We find weakly significant results on female native students, despite not consistently estimated across the different grades. The effects on male native children are even weaker and less precise. This specific finding aligns with most existing literature on peer effects that supports the existence of heterogeneous effects on the basis of gender that mostly take place within gender groups. Male students, irrespective of their origin, are more likely to be the disruptive elements in a classroom. Also, there exist some evidence pointing to the possibility that they are more influenced by group dynamics, making them the ideal subject of peer effects.

Native children as well differ between each other in a way that affects how they endure influences of their immigrant peers. One of the main drivers of differential effects of early immigrant exposure can be the socio-economic and cultural background. We distinguish native children into three categories: native pupils with parents with a degree or anything above that; students whose parents completed at maximum lower secondary education; native children left out of the two main categories. Parental education serves as a good proxy for the socio-economic and cultural background of children. We found evidence negative effects arising from immigrant concentration are mostly driven by the adverse influences on children from a poor or “medium” socioeconomic background, for which negative effects are strong, statistically significant and consistently estimated up to 13th grade. By contrast, there is no indication that native pupils from a high socio-economic background suffer negative effects from early immigrant exposure. When using the number of books at home as a proxy of the cultural background of native students, we observe similar results. The strongest adverse effects are concentrated among the kids with few books (11-25 books) at home and those with some books (26-100). Native students with more than 100 books at home do not suffer any effect from early immigrant exposure, not positive nor negative. The same holds true for students with very few books at home (less than 10).

Our analysis also yields proof that the largest negative immigrant peer effects are consistently estimated in the northern-central regions, whereas we fail to find such proof in the southern part of Italy, where segregation reaches higher values, as measured by the Duncan index.

The results of this research are proved to be robust to the potential sorting of children across schools in large provinces, where the greater number of schools allows families a higher degree of choice as to where to enroll their children.

Overall, there is strong evidence that early immigrant exposure in school affects negatively the performance of native children, and that such detrimental effects stick along their whole primary and secondary education. While these effects are can be considered weak in their size, they are still not negligible. What is particularly interesting is the trajectory we observed for these effects. Once we control properly for other important factors that shape performance in school, we would expect the share of immigrants in 5th grade – which is a transitional factor – to influence only contemporaneous outcomes. The fact that we are able to identify significant effects up to 13th grade, indicates that early immigrant exposure shape permanently the educational outcomes of native children.

As mentioned multiple times in this work, the fact that we observe only *test-takers* and the issue of attrition constitute a great challenge to the validity of this work. As for the former, INVALSI make it possible to observe only children who sat the tests. If the population of students missing from INVALSI observations is not randomly distributed, we would introduce bias in our estimates. Similarly, if the students we lose through the merging process due to attrition are systematically similar among each other and different from those who remain in the final sample, the resulting set of students may be non-representative of the true population of pupils in Italy. More importantly, it can be that our sample is made mostly of students that are on average high-achievers or have positive attitude towards school, so that they are unlikely to be purposefully absent form INVALSI tests or to be held back a grade at some point during schooling years. If that was the case, our estimates would detect an attenuated effect of immigrant concentration. Despite that, the fact that we still find significant results is an indicator that early interactions with immigrant peers strongly affect the performance of native children in schools, and that such influences stick in the long-term. Nevertheless, our weighted regressions suggest that attrition should not be a main issue as for the validity of our estimates.

The main findings of this paper are in line with existing literature on peer effects in Italy. As Contini (2013) and Frattini & Meschi (2019), we found at times more adverse effects on mathematics test scores, than on reading ones. We tend to observe long-term strong detrimental effect on mathematics performance because, as school proceeds, it becomes more difficult to fill preexistent gaps. Within the Italian framework, Ballatore et al. (2018) as well find evidence of large and statistically significant effects of the number of non-native students on both math and language performance. Among the restricted plethora of research on long-term peer effects, Gould et al. (2009) conclude that early exposure to immigrant children can negatively affect long-term educational outcomes of native students, even though there is a certain degree of heterogeneity depending on which outcomes are under analysis.

This paper opens up to proper research on the persistence of peer effects, an aspect that requires careful and specific analysis. Further investigation hopefully will have the chance to make use of observations from multiple cohorts and more detailed information on the specific country of origin and language spoken at home.

References

- Ammermueller, A., & Pischke, J. (2009). Peer effects in European primary schools: evidence from PIRLS. *Journal of Labor Economics*, 27(3), 315–348.
- Andersen, S. C., & Thomsen, M. K. (2011). Policy Implications of Limiting Immigrant Concentration in Danish Public Schools. *Scandinavian Political Studies*, 34(1), 1–99.
- Angrist, J. D., & Lang, K. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *The American Economic Review*, 94(5), 1613–1634.
- Angrist, J. D., & Lavy, V. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *The Quarterly Journal of Economics*, 114(2), 533–575.
- Angrist, J. D., & Pischke, J.-S. (2014). *Masterin 'Metrics - The Path from Causa to Effect* (Princeton University Press, Ed.).
- Ballatore, R. M., Fort, M., & Ichino, A. (2018). Tower of Babel in the Classroom: Immigrants and Natives in Italian Schools. *Journal of Labor Economics*, 36(4), 841–1163.
- Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*, 106(2), 407–443.
- Barro, R. J. (1996). *Determinants of Economic Growth: A Cross-Country Empirical Study* (Working Paper No. 5698). National Bureau of Economic Research.
- Barro, R. J. (2001). Human Capital and Growth. *American Economic Review*, 91(2), 12–17.
- Barro, R. J., & Lee, J.-W. (1993). International comparisons of educational attainment. *Journal of Monetary Economics*, 32(3), 363–394.
- Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70(5), 9–49.
- Black, S. E., Devereux, P. J., & Salvanes, K. G. (2010). *Under pressure? The effect of peers on outcomes of young adults* (Working Paper No. 16004). National Bureau of Economic Research.
- Boado, H. C. (2007). Immigrant Concentration in Schools: Peer Pressures in Place? *European Sociological Review*, 23(3), 341–356.
- Brunello, G., De Paola, M., & Scoppa, V. (2010). Peer Effects in Higher Education: Does the Field of Study Matter? *Economic Inquiry*, 48(3), 621–634.

- Brunello, G., & Rocco, L. (2013). The effect of immigration on the school performance of natives: Cross country evidence using PISA test scores. *Economics of Education Review*, 32, 234–246.
- Brunello, G., Weber, G., & Weiss, C. T. (2017). Books are Forever: Early Life Conditions, Education and Lifetime Earnings in Europe. *The Economic Journal*, 127(600), 271–296.
- Card, D., & Rothstein, J. (2007). Racial segregation and the black-white test score gap. *Journal of Public Economics*, 91(11–12), 2158–2184.
- CODE - Coordination Des ONG pour les Droits de l’Enfant. (2024). *L’Intégration Scolaire des Enfants Migrants infra-scholarisés et leur Accès aux Dispositifs Alternatifs*. Available at: https://lacode.be/wp-content/uploads/2024/09/lacode_L-integration-scolaire-des-enfants-migrants-infra-scolarises-et-leur-acces-aux-dispositifs-alternatifs_sept24.pdf (accessed June 4, 2025).
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*.
- Contini, D. (2013). Immigrant background peer effects in Italian schools. *Social Science Research*, 42(4), 1122–1142.
- Darling-Hammond, L. (2000). Teacher Quality and Student Achievement: A Review of State Policy Evidence. *Education Policy Analysis Archives*, 8(1), 1–44.
- Del Boca, D., Monfardini, C., & See, S. G. (2018). *Government Education Expenditures, Pre-Primary Education and School Performance: A Cross-Country Analysis* (Discussion Paper No. 11375). IZA Institute of Labor Economics.
- Duncan, O. D., & Duncan, B. (1995). A Methodological Analysis of Segregation Indexes. *American Sociological Review*, 20(2), 210–217.
- Dustmann, C., & Fabbri, F. (2003). Language Proficiency and Labour Market Performance of Immigrants in the UK. *The Economic Journal*, 113(489), 695–717.
- Etim, J. S., Etim, A. S., & Blizard, Z. D. (2020). Class Size and School Performance: An Analysis of Elementary and Middle Schools. *International Journal on Studies in Education*, 2(2), 66–77.
- European Commission - European School Education Platform. (2023, July). *Tackling school segregation in European cities*. Available at: [128](https://school-</p>
</div>
<div data-bbox=)

education.ec.europa.eu/en/discover/practices/tackling-school-segregation-european-cities
(accessed June 3, 2025).

European Commission - Migration and Home Affairs. (n.d.). *Geia Xara - Greek lessons and mediation services to support children's integration*. Available at: https://home-affairs.ec.europa.eu/projects/geia-xara-greek-lessons-and-mediation-services-support-childrens-integration_en (accessed June 4, 2025).

Eurydice - European Commission. (2024, March). *Support measures for learners in early childhood and school education*. Available at: <https://eurydice.eacea.ec.europa.eu/eurypedia/belgium-french-community/support-measures-learners-early-childhood-and-school-education> (accessed June 4, 2025).

Evans, W. N., Oates, W. E., & Schwab, R. M. (1992). Measuring Peer Group Effects: A Study of Teenage Behavior. *Journal of Political Economy*, *100*(5), 966–991.

Frattini, T., & Meschi, E. (2019). The effect of immigrant peers in vocational schools. *European Economic Review*, *113*, 1–22.

Geay, C., McNally, S., & Telhaj, S. (2013). Non-Native Speakers of English in the Classroom: What Are the Effects on Pupil Performance. *Economic Journal*, *123*(570), F281–F307.

Giambona, F., & Porcu, M. (2018). School size and students' achievement. Empirical evidences from PISA survey data. *Socio-Economic Planning Sciences*, *64*, 66–77.

Gould, E. D., Lavy, V., & Paserman, M. D. (2009). Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi- Experimental Evidence. *The Economic Journal*, *119*(540), 1243–1269.

Hakuta, K. (1999). The Debate on Bilingual Education. *Journal of Developmental & Behavioral Pediatrics*, *20*(1), 36–37.

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. *Journal of Labor Economics*, *27*(3), 349–383.

Hanushek, E. A., & Kimko, D. D. (2000). Schooling, Labor-Force Quality, and the Growth of Nations. *American Economic Review*, *90*(5), 1184–1208.

Hanushek, E. A., & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, *46*(3), 607–668.

- Hanushek, E. A., & Woessmann, L. (2010). Education and Economic Growth. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International Encyclopedia of Education* (Vol. 2, pp. 245–252). Elsevier.
- Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4), 267–321.
- Hanushek, E., Kain, J., Markman, J., & Rivki, S. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5), 527–544.
- Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, 54, 3–56.
- Heller-Sahlgren, G., & Jordahl Henrik. (2023). Test scores and economic growth: update and extension. *Applied Economics Letters*, 31(11), 1024–1027.
- Hoxby, C. (2000). *Peer effects in the classroom: learning from gender and race variation* (Working Paper No. 7867). National Bureau of Economic Research.
- Hoxby, C. (2006). The Power of Peers. *Education Next*, 2(2).
- Hunt, J. (2017). The Impact of Immigration on the Educational Attainment of Natives. *Journal of Human Resources*, 52(4), 1060–1118.
- Infor Jeunes Laeken. (n.d.). *Aperçu sur les DASPA, ou classes-passerelles*. Available at: <https://inforjeunes.eu/apercu-sur-les-daspa-ou-classes-passerelles/> (accessed June 4, 2025).
- INVALSI. (2022). *Rapporto Invalsi 2022*. Available at: <https://www.istat.it/it/files/2016/06/Bilancio-demografico-2015-1.pdf> (accessed April 22, 2025).
- Istat. (2016). *Bilancio Demografico Nazionale*.
- Kao, G., & Joyner, K. (2004). Do Race and Ethnicity Matter among Friends? Activities among Interracial, Interethnic, and Intraethnic Adolescent Friends. *The Sociological Quarterly*, 45(3), 557–573.
- Kraay, A. (2018). *Methodology for a World Bank Human Capital Index* (Working Paper No. 8593). World Bank Group.
- Krueger, A. B. (2003). Economic Considerations and Class Size. *The Economic Journal*, 113(485), F34–F63.

- Lavy, V., Paserman, M. D., & Schlosser, A. (2012). Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in Classroom. *The Economic Journal*, 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. *Journal of Labor Economics*, 30(2), 367–414.
- Lazear, E. P. (2001). Educational Production. *The Quarterly Journal of Economics*, 116(3), 777–803.
- Mankiw, R. D., & Weil D. A. (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, 107, 407–437.
- Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3), 531–542.
- Marcińczak, S., Musterd, S., van Ham, M., & Tammaru, T. (2016). Inequality and rising levels of socio-economic segregation Lessons from a pan-European comparative study. In *Socio-Economic Segregation in European Capital Cities: East Meets West* (pp. 358–382).
- McEwan, P. J. (2003). Peer effects on student achievement: evidence from Chile. *Economics of Education Review*, 22(2), 131–141.
- McPherson, M., Smith-Lovin, L., & M. Cook, J. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27, 415–444.
- Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66(4), 281–302.
- MIUR. (2017). *Gli alunni stranieri nel sistema scolastico italiano a.s. 2015/2016*. Available at: https://www.mim.gov.it/documents/20182/250207/Notiziario_alunni_Stranieri_nel+sistema_scolastico_italiano_15_16.pdf/777cc2d8-1bf3-42df-9814-2539f777c3d0?version=1.0&t=1495620164915 (accessed May 15, 2025).
- Nelson, C. A. (2000). Neural plasticity and human development: The role of early experience sculpting memory systems. *Developmental Science*, 3(2), 115–130.
- Sabates, R., & Parsons, S. (2012). *The contribution of basic skills to health related outcomes during adulthood: Evidence from the BCS70* (Working Paper No. 91). Department for Business, Innovation and Skills.

- Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics*, 116(2), 681–704.
- Sacerdote, B. (2011). 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, pp. 249–277). Elsevier.
- Sala-i-Martin, X., Doppelhofer, G., & Miller, R. I. (2004). Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *The American Economic Review*, 94(4), 813–835.
- Seah, K. K. C. (2021). Impact of Immigrant Peers on Native Students' Achievement: Evidence from Australia, Canada and the United States. *The Australian Economic Review*, 54(1), 94–116.
- Shuey, E., & Kankaraš, M. (2018). *The Power and Promise of Early Learning* (Working Paper No. 186). OECD Publishing.
- Tonello, M. (2016). Peer effects of non-native students on natives' educational outcomes: mechanisms and evidence. *Empirical Economics*, 51, 383–414.
- Ministero dell'Istruzione e del Merito. (2023). *La dispersione scolastica aa.ss. 2019/2020 – 2020/2021 e aa.ss. 2020/2021 – 2021/2022*. Available at: https://www.mim.gov.it/documents/20182/7715421/Focus_Dispersione+scolastica+aa.ss._1920_2021+-+2021_2122.pdf/7574e014-b372-d32c-a62c-ddabbd5d7c7c?version=1.0&t=1703760495410 (accessed June 6, 2025).
- United Nations, P. D. (2025). *International Migrant Stock 2024: Key facts and figures* (Report No. UN DESA/POP/2024/DC/NO.13). Available at: <https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesapd2025intlmigstock2024keyfactsandfiguresadvance-unedited.pdf> (accessed June 3, 2025).
- World Bank. (2020). *The Human Capital Index 2020 Update: Human Capital in the time of COVID 19*. Available at: <https://documents1.worldbank.org/curated/en/456901600111156873/pdf/The-Human-Capital-Index-2020-Update-Human-Capital-in-the-Time-of-COVID-19.pdf> (accessed June 3, 2025).

Zimmerman, D. J. (2003). Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *The Review of Economics and Statistics*, 85(1), 9–23.

Appendix

A.4.2 Observations in INVALSI dataset without SIDI code and number of observations lost during the merging process

Table A.4.2.1. Number of observations without SIDI Invalsi code in each dataset retrieved from INVALSI.

Subject	5th grade		8th grade				13th grade			
	Mathematics	Language	Mathematics	Language	English reading	English listening	Mathematics	Language	English reading	English listening
Number of observations without SIDI code	4.411	4.063	4.307	4.318	4.299	4.297	193	195	201	199

Table A.4.2.2. Number of observations lost in each step of the merging process.

Merge step	Datasets	Unmatched in master	Unmatched in using	Matched	Result
1	Mathematics (5 th) + language (5 th)	33.632	15.392	393.288	Dataset on 5 th grade
2	Mathematics (8 th) + language (8 th)	4.369	691	544.937	
3	Merge of step 2 + English listening (8 th)	2.254	4.820	542.683	
4	Merge of step 3 + English reading (8 th)	131	6.029	542.552	Dataset on 8 th grade
5	Mathematics (13 th) + language (13 th)	4.201	2.202	478.882	
6	Merge of step 5 + English listening (13 th)	2.071	6.441	476.811	
7	Merge of step 6 + English reading (13 th)	92	6.892	476.719	Dataset on 13 th grade
8	Dataset (5 th) + dataset (8 th)	28.975	178.239	364.313	
9	Merge of step 8 + dataset (13 th)	86.492	198.898	277.821	Final dataset

A.4.5. Relationship between test scores and share of first-generation immigrants and second-generation immigrants in 5th grade

Figure A.4.5.1. Relationship between average score of natives in mathematics test in 5th grade and share of immigrant students in the school in 5th grade. Panel A uses the share of first-generation immigrants; panel B uses the share of second-generation immigrants.

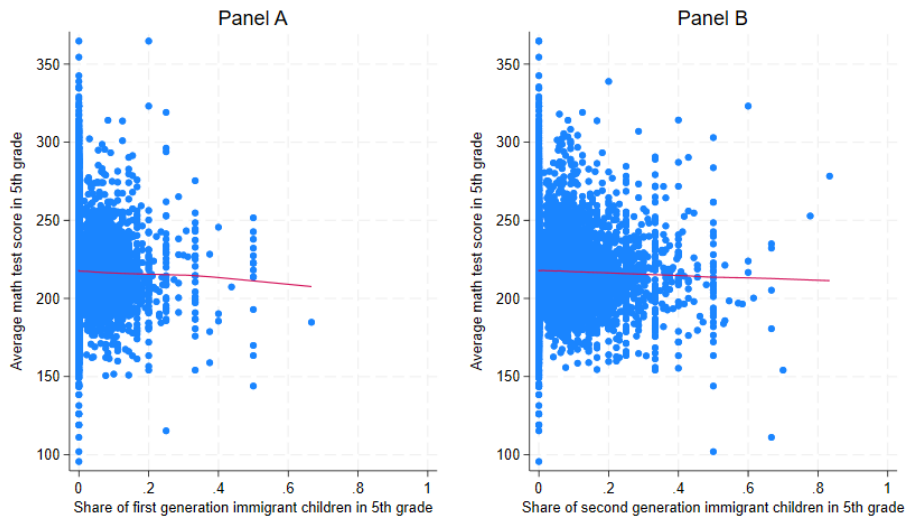


Figure A.4.5.2. Relationship between average score of natives in language in 5th grade and share of immigrant students in the school in 5th grade. Panel A reports the association using the share of first-generation immigrants; panel B using share of second-generation immigrants.

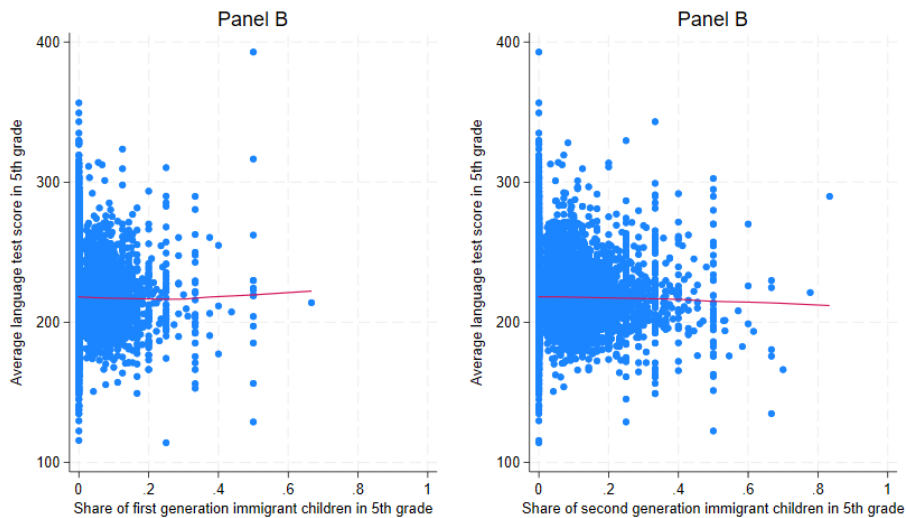


Figure A.4.5.3. Relationship between average score of natives in mathematics in 8th grade and share of immigrant students in the school in 5th grade. Panel A reports the association using the share of first-generation immigrants; panel B using share of second-generation.

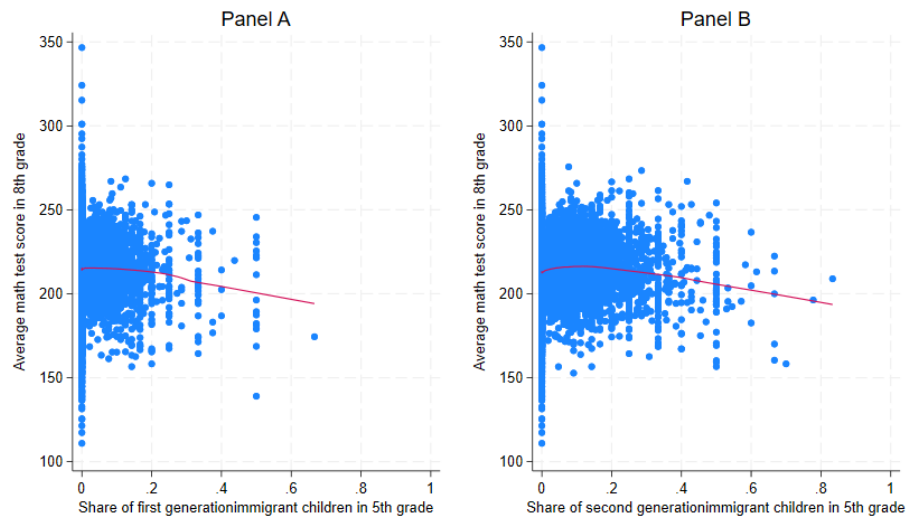


Figure A.4.5.4. Relationship between average score of natives in mathematics in 8th grade and share of immigrant students in the school in 5th grade. Panel A reports the association using the share of first-generation immigrants; panel B using share of second-generation immigrants.

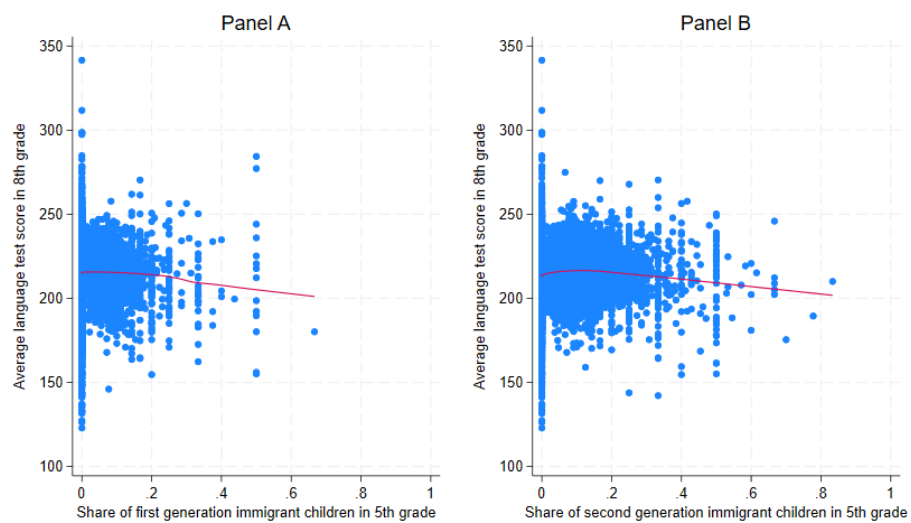


Figure A.4.5.5. Relationship between average score of natives in English (reading) in 8th grade and share of immigrant students in the school in 5th grade. Panel A reports the association using the share of first-generation immigrants; panel B using share of second-generation immigrants.

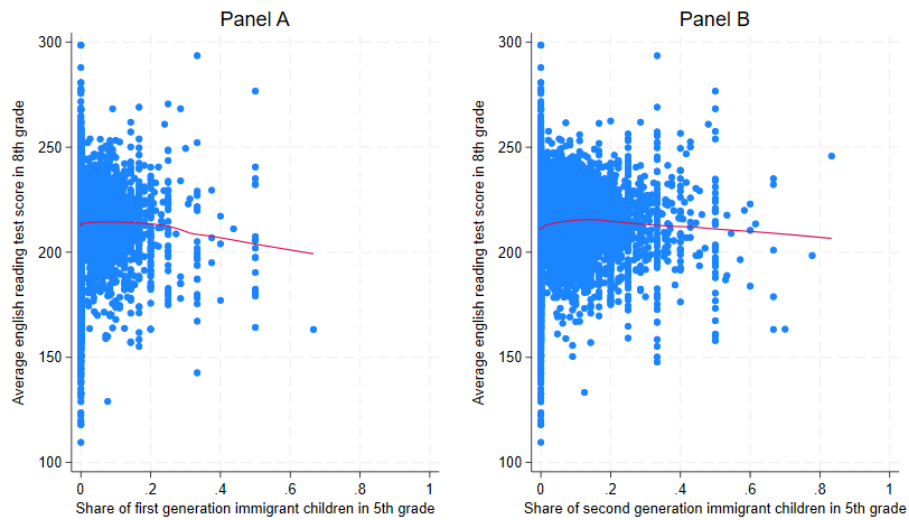


Figure A.4.5.6. Relationship between average score of natives in English (listening) in 8th grade and share of immigrant students in the school in 5th grade. Panel A reports the association using the share of first-generation immigrants; panel B using share of second-generation immigrants.

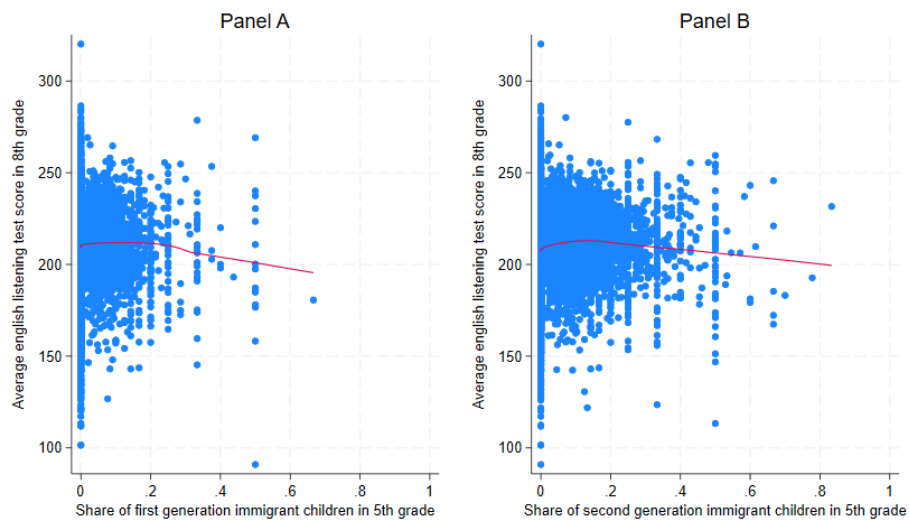


Figure A.4.5.7. Relationship between average score of natives in mathematics in 13th grade and share of immigrant students in the school in 5th grade. Panel A reports the association using the share of first-generation immigrants; panel B using share of second-generation immigrants.

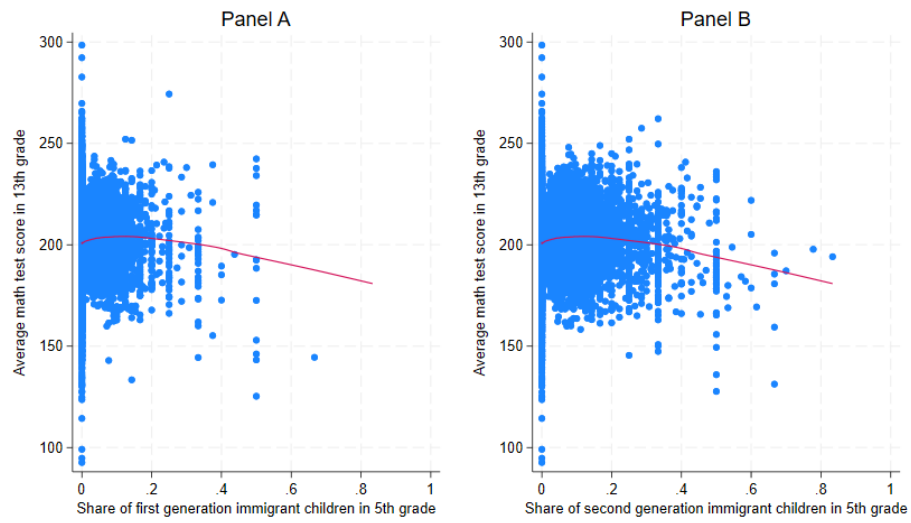


Figure A.4.5.8. Relationship between average score of natives in language in 13th grade and share of immigrant students in the school in 5th grade. Panel A reports the association using the share of first-generation immigrants; panel B using share of second-generation immigrants.

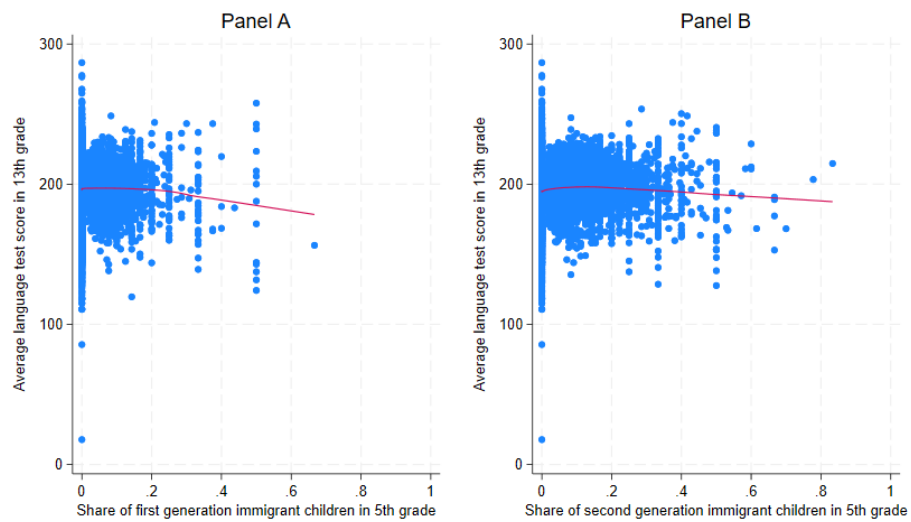


Figure A.4.5.9. Relationship between average score of natives in English (reading) in 13th grade and share of immigrant students in the school in 5th grade. Panel A reports the association using the share of first-generation immigrants; panel B using share of second-generation immigrants.

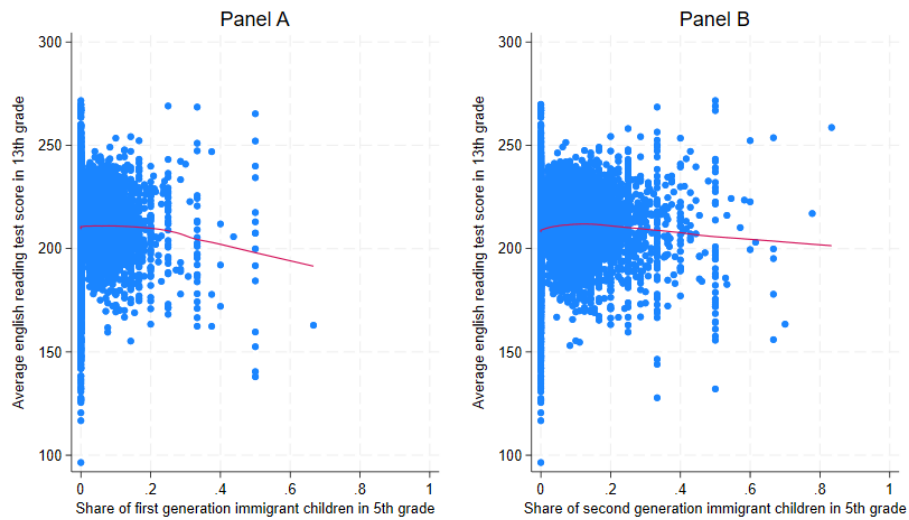
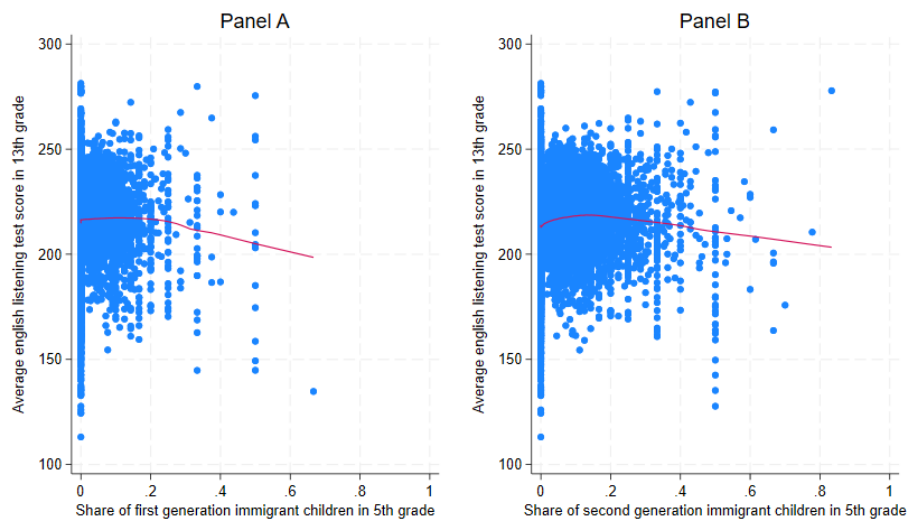


Figure A.4.5.10. Relationship between average score of natives in English (listening) in 13th grade and share of immigrant students in the school in 5th grade. Panel A reports the association using the share of first-generation immigrants; panel B using share of second-generation immigrants.



A.6.1. Distribution of first-generation immigrant children by their age of arrival in Italy.

Table A.6.1.1. Distribution of first-generation immigrants by their age of arrival in Italy.

Age of arrival in Italy	Number of children			
	All origins	European Union	European country (non- UE)	Other
1 year old or younger	706	236	161	308
2 years old	496	165	119	210
3 years old	548	226	109	210
4 years old	495	185	127	181
5 years old	487	133	135	217
6 years old	445	120	141	182
7 years old	373	95	117	159
8 years old	312	87	88	135
9 years old	228	58	56	114
10 years old or older	218	55	54	107
Not available	1889	597	319	835
Total	6207			