Cite this article

Porcu, M., Sulis, I., Usala, C. *et al.* Will the gap ever be bridged? A cross-national comparison of non-native students' educational achievements. *Genus* **79**, 19 (2023). https://doi.org/10.1186/s41118-023-00199-5

- DOI
  https://doi.org/10.1186/s41118-023-00199-5

.

Genus

ORIGINAL ARTICLE Open Access

# Will the gap ever be bridged? A cross-national comparison of non-native students' educational achievements

Mariano Porcu[1], Isabella Sulis[1], Cristian Usala[1*] and Francesca Giambona[2]

*Correspondence:
cristian.usala@unica.it

[1] Dipartimento di Scienze Politiche e Sociali, Università di Cagliari, Viale S. Ignazio 78, 09123 Cagliari, Italy
[2] Dipartimento di Statistica, Informatica e Applicazioni "G. Parenti", Università di Firenze, Viale Morgagni, 59, 50134 Firenze, Italy

**Abstract**

A crucial element in evaluating the success of immigrant integration policies is to compare the school performances of immigrant students with that of natives. According to large-scale international assessments, immigrant students tend to underperform their native peers even after controlling for socioeconomic conditions, with relevant differences depending on countries' migration histories. This article analyses the evolution in the skills gap between native students and students with an immigrant background by comparing traditional (France, Germany, and the United Kingdom) and new immigration countries (Italy and Spain) using data collected from the Program for International Student Assessment survey from 2009 to 2018. We model students' learning outcomes in mathematics and reading in a multilevel setting by clustering students at the high school level. This approach allows us to compare countries' relative performances concerning the immigrant integration process by accounting for the role played by high schools and students' backgrounds. Results show that the differences between native and immigrant students in both sets of countries are narrowing but still present.

**Keywords:** Immigration, Educational inequalities, PISA, Multilevel models

## Introduction

Every year, millions of people migrate across national boundaries, driven by the hope for a better life and the need to escape adverse conditions. This phenomenon has grown substantially in the last 30 years and now occupies a prominent place on the global political agenda. Factors such as improved transportation, the globalisation of economies, population ageing and demographic decline of Western populations will make this topic even more prominent in the future. Thus, the integration of immigrants into host countries is a crucial matter in question, both for economic systems and long-term growth in social welfare. Educational systems play a crucial role in this integration process, particularly as the number of students with a migratory background continues to rise. Indeed, according to the Organisation for Economic Cooperation and Development (OECD), on average, in the last two decades, the percentage of 15-year-old students with a migratory background has increased by more than two points (OECD, 2012b, 2015a).

Previous research has consistently shown a substantial performance gap between immigrant students and natives in several contexts, with immigrants generally underperforming their native peers. Factors such as students' socioeconomic background, families' economic and cultural endowments, language proficiency, and country of origin contribute to exacerbating inequalities among students with different migration backgrounds [see, among others, Azzolini et al. (2012); Gabrielli and Impicciatore (2022); Heath and Brinbaum (2007); Schnell and Azzolini (2015)], especially when considering less-advantaged students (Gabrielli et al., 2022).

Building upon this literature, this paper aims to classify countries according to their pathway towards integrating immigrant students into their educational systems. In particular, we rely on data collected from the OECD Program for International Student Assessment (PISA) surveys conducted in 2009, 2012, 2015, and 2018 and on a multilevel approach. This data allows us to analyse students' performances by accounting for the role played by students' and families' characteristics, as well as the attended high schools. Combining this data with our analytical approach, we provide descriptive evidence on the trajectories towards reducing educational achievement inequalities between students with different migratory backgrounds. Specifically, we analyse the evolution of the performance gap between native and immigrant students over time in five European countries: France, Germany, Italy, Spain, and the United Kingdom. These countries are selected due to their distinct historical perspectives. Italy and Spain, previously emigration nations, have become destination countries after the fall of the Iron Curtain, making them relatively new destinations. The other three countries have a more consolidated, albeit different, migration history. France and the United Kingdom have experienced incoming migrant flows reflecting their colonial history. In contrast, Germany's rampant industrialisation process has acted as a beacon for foreign workers since the mid-twentieth century.

Besides the differences between their migratory histories, these countries exhibit substantial heterogeneity in the organisation of their educational systems and integration policies. While these aspects greatly affect students' achievements, our paper focuses on measuring and comparing the educational inequalities in these countries and monitoring their road map towards a more inclusive system, rather than explaining the effectiveness of specific school policies and practices. Nonetheless, we utilise the general information on educational systems, policies, and practices observed in the selected countries to provide a qualitative description of the main differences that contribute to explaining countries' pathways toward reducing the inequalities related to students' migratory backgrounds. Thus, we focus on the differences in students' performances between and within countries accounting for students' demographic characteristics and the heterogeneity at the high school level to capture changes over time between native students and students with immigrant backgrounds. These findings allow both to make comparisons between and within countries and highlight the observed road map of each country towards integrating students with a migratory background. Moving from this framework, this paper deals explicitly with the following research questions:

– How do educational outcomes in mathematics and reading differ among students with immigrant backgrounds in new and old destination countries?

- Have countries experienced a reduction in inequalities linked to immigrant status over time?
- Are there countries that tend to fill the gap in academic performances between first- and second-generation immigrants?

This paper contributes to the previous literature by providing a measure of the educational gap between native and immigrant students by relying on (i) observing five countries with different immigration histories; (ii) information provided by multiple waves of an international large-scale assessment survey; (iii) using a robust approach which combines multilevel models with Rubin's rules for multiple imputations to account for the role played by schools and students' characteristics. Moreover, to the best of our knowledge, this is the first time that country × time × immigrant status interaction terms are used to depict the trajectories of countries towards equity between students with different immigrant backgrounds in order to spot the effects of the immigrant conditions in each country over time and to differentiate between first- and second-generation immigrants, providing robust insights on the position of countries in their road map towards integrating immigrants. That allows us to evaluate countries' positions in absolute and relative terms. The results show that the differences between native and immigrant students are relevant, especially if we compare first-generation immigrants and natives, while the gap between natives and second-generations has narrowed in the time frame considered.

The paper is organised as follows: Sect. 2 provides an overview of the existing literature related to the analysis of the gaps in educational achievement between students with different immigrant backgrounds. Section 3 provides information on migration trends and the main features of educational systems in the countries selected for this study. Section 4 presents the data, and Sect. 5 outlines the model fitted. Results are presented in Sect. 6. Sections 7 and 8 contain discussion and final remarks.

## Theoretical framework

The vast literature on the determinants of students' competencies acknowledges family socioeconomic conditions as the main predictor of inequalities observed among students, even though important differences are observed between schools and countries (Bertoletti, 2023; Masci and Ieva, 2016; Sulis et al., 2020) that are related to teachers' practices, schools' characteristics and management (Agasisti et al., 2020; Bertoletti, 2023; Masci et al., 2018). Immigrant students' low academic performances are regularly observed in many economically developed countries as highlighted by many international reports (OECD, 2010, 2011, 2012b, 2013b, 2016, 2019).

This gap in educational achievements is influenced by many factors related to language and cultural barriers, migration-related costs, discrimination, and socioeconomic inequalities between natives and immigrants [see, for example, Chiswick and DebBurman (2004); Hartung (2015); Gabrielli et al. (2022)]. Immigrant students typically underperform compared to natives and are characterised by poorer socioeconomic conditions that exacerbate these differentials. Nevertheless, the educational gap remains even after controlling for students' socioeconomic status (OECD, 2012b). Indeed, immigrants' educational disadvantages are also affected by the national context. For example, Hillmert

(2013) compared migrants' educational performances in five traditional immigration countries (France, Germany, the Netherlands, Sweden, and the United Kingdom), concluding that countries' institutions, policies, and education interact in shaping immigrants' educational situation. Moreover, a considerable degree of heterogeneity in educational performances is explained by students' countries of origin and the level of social integration achieved in the host country.

Despite the extensive literature on educational achievement gaps between natives and immigrants, most studies focus on traditional destination countries for immigration, such as North America or the North Central European countries. Their findings indicate that the educational disadvantage of young immigrants accounts for the lower endowment of parental resources; often, once this under-supply is controlled, the educational gap shrinks or even disappears. Focusing on Germany, Lüdemann and Schwerdt (2013) observed how immigrant students are disadvantaged in terms of grades, teacher recommendations, and standardised IQ test scores, which then influence the wage gap between native and immigrant workers. For North America countries, Rothon et al. (2009) investigated the educational attainments of minority groups in the United States, Canada, and Britain highlighting the pivotal role of students' socioeconomic background and countries of origin, with students from China and India often outperforming natives. Comparing the data on 8 industrialised countries, Schnepf (2008) found that countries characterised by high levels of educational dispersion among natives (such as New Zealand and Australia) also exhibit high dispersion among immigrants. Moreover, they provide evidence of the importance of student's language skills, especially for low achievers. In the case of new immigration countries, Schnell and Azzolini (2015) examined PISA data for Greece, Italy, Portugal, and Spain, showing that immigrants' disadvantages are primarily linked to limited economic resources rather than parents' educational credentials, which align to those of natives. This suggests that the low level of integration of adult immigrants in the job market leads to lower-paid occupations, negatively impacting family well-being and students' academic performances.

Regarding immigrants' integration, two other important elements affect their educational attainment: generational status and the time spent in the country. Second-generation immigrants tend to outperform first-generation immigrants on average (Schnepf, 2004; OECD, 2012b), although this effect varies depending on the country of origin. For example, studying the participation of immigrants in secondary education in Italy and using labour force survey data, Azzolini and Barone (2013) observe how for certain national groups (e.g., East Asians) the differential with respect to natives disappears in the second-generations. However, other groups (e.g., Northern Africans) still exhibit lower performances even in the second-generation. Concerning the time spent in the country, previous literature indicates that the gap reduces as immigrants stay in the country longer. For example, examining PISA data for Spain, Zinovyeva et al. (2014) observe that immigrant students initially underperform natives, but the performance gap narrows the more the time spent in Spain. The study's results emphasise that although those of Latin American origin have an initial linguistic advantage among immigrants, this advantage does not help them to fill the gap compared with natives. According to the authors, most of the gap is due to individual and family features, and only a limited amount is due to academic characteristics. Similar evidence is also reported by Vaquera

and Kao (2012). Comparing Italy and Spain using PISA data, Azzolini et al. (2012) analyse the educational achievement gap between immigrants and native students. Also in this case, socioeconomic background and linguistic skills prove to be significant predictors for the observed gaps in both nations, although the language spoken at home is more important in Italy.

Lastly, students' school tracks play a significant role in creating an educational gap for immigrants. For instance, Contini and Azzolini (2016) found that the gap arises from immigrants' lower academic performance (primary effect) and the different decision models that influence families' choices regarding their children's school paths (secondary effect). Analysing the transitions from lower to upper secondary schools in Italy, and controlling for socioeconomic background, the authors show that children of immigrants tend to enrol in vocational paths more than natives, contributing to widening the existing gaps. Gabrielli et al. (2022), studying the differences in academic resilience between natives and immigrant students in traditional (France, Germany, United Kingdom and Netherlands) and new migration countries (Italy and Spain), provide evidence that the policies aimed at improving the school environment with high-quality resources in terms of services, teachers' qualification, parental involvement and extra curricular activities have the greatest benefits for immigrant students and those with a vulnerable socioeconomic backgrounds. Results also highlight that variations in school policies and practices are the main factor contributing to differences in students' resilience. Moreover, resilient immigrant students are more prevalent in North-western European countries and among those who migrate in early childhood. Similar findings are discussed by Gabrielli and Impicciatore (2022), who emphasise that negative school-related attitudes create barriers between natives and non natives since the primary and low secondary school, as shown by the strong association between students' performance and disadvantaged socioeconomic backgrounds [see also Sulis and Porcu (2015); Triventi et al. (2022)].

Therefore, the literature depicts a complex picture in which socioeconomic inequalities and immigration policies in the hosting nations shape the gap in learning outcomes between students with different immigrant backgrounds. Moreover, it shows a divide between countries that tend to reproduce disparities across generations and those that succeed in bridging the gap between second-generation immigrants and natives.

## Setting the background

As shown in Fig. 1, in 2018, the percentage of foreign citizens and stateless in the total population ranged between 7.4% for France to 11.7% for Germany. Italy, Spain and the United Kingdom register, respectively, a percentage equal to 8.5%, 9.8% and 9.5%. Considering the evolution between 2007 and 2018, we can observe a rise in the percentage of foreigners in all the countries except for Spain, which registered a reduction of 0.2 percentage points. In the time frame considered, the percentage of foreign citizens increased by 3.5 points in Italy and the United Kingdom and 3 and 1.6 points in Germany and France, respectively.

This increase in the foreigner population has been met with a variety of policies implemented by host nations and international treaties. Since the early twentieth century, migration flows have also been regulated to select migrants based on their skills or to

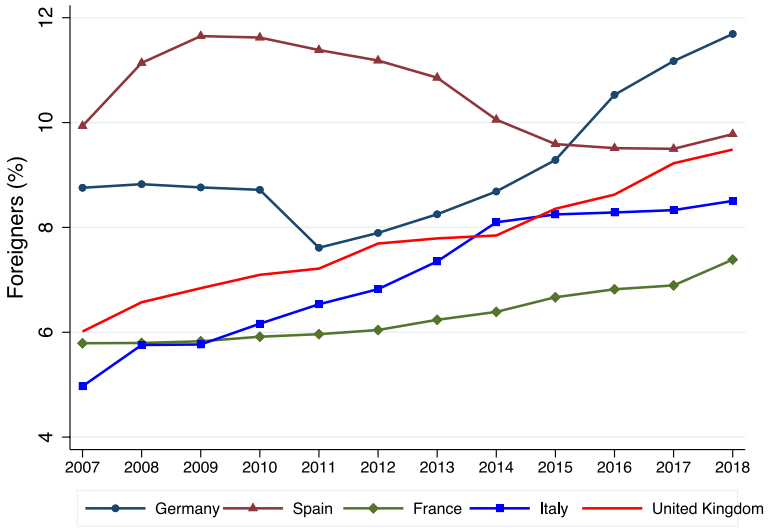


**Fig. 1** Percentage of foreigners and stateless in the resident population. (*Source*: EUROSTAT)

give immigrants the right to live with their families. As noted by Schnepf (2004) and OECD (2012b), these features have a direct and relevant impact on the issues related to the education of immigrant students.

Apart from differences in the evolution of immigrant populations, these countries also show notable variations in their composition. On one hand, Italy and Spain can be classified as new destination nations, with a large population of low-educated immigrants. On the other hand, France and Germany are classified as long-standing destination countries with many low-educated immigrants, while the United Kingdom is a long-standing destination country with many recent and highly educated immigrants (OECD, 2016).

Despite having compulsory education from 9 to 11 years starting from the age of 6, these countries exhibit significant variations in the organisation of their school systems. In Germany, the education system is jointly governed by the federal government and the 16 Länder. Each Länder has the authority to make laws on educational issues and its own ministries (OECD, 2020b). The Spanish education system is quite decentralised, with the 17 regions responsible for schools maintenance, funding, and organisation. Upper secondary education provides both a general path for direct access to tertiary education and vocational training. The proportion of students enrolled in private schools exceeds the OECD average, and grade repetition is common (about 35% of students repeat a grade by age 15) (OECD, 2017b). The French school system is highly centralised. The curriculum is the same nationwide up to the secondary level. After five years of primary school, students progress to *collèges*, which cover the first four years of secondary education, followed by *lycées* that offer a 3-year course of secondary education (OECD, 2020a). In the United Kingdom, the education policy and governance are the responsibility of its four countries (England, Scotland, Wales, North Ireland); vocational education is provided at both secondary and tertiary levels, with large differences across countries (OECD, 2015b). Italy's Ministry of Education oversees the governance of the education system and sets national minimum standards, while some responsibilities are shared with the regions. Schools have some autonomy in resource allocation, and any

secondary educational path allows access to tertiary education. Italy has higher school dropout rates than the OECD average and notable regional disparities in student outcome measures (OECD, 2017a).

The diverse composition of migrant populations and distinct characteristics of the high school system at the national and local level have important consequences on students' educational attainment and the differences among students with immigrant backgrounds. Thus, it is crucial to consider these differences when interpreting measures of educational attainment gaps between natives and immigrants over time.

## Data

This analysis considers data collected in the last four PISA surveys (2009, 2012, 2015, and 2018). The PISA target population is that of students aged between 15 years and 3 months and 16 years and 2 months at the time of the survey who have completed a minimum of 6 years of formal education. For many countries, the age of 15–16 represents the time for transition from basic education to a more advanced one (OECD, 2012a, 2014, 2017c, 2020c). We consider students' performances in reading and mathematics as dependent variables. We focus on these subjects as reading skills (literacy) are essential for learning and understanding information across all subjects, while math skills (numeracy) are crucial for success in many fields, such as science, engineering, finance, and technology.

In PISA surveys, to minimise the assessment burden on each student and to avoid the scaling of skills being influenced by the "booklet effect", each student is asked to handle only a part of the whole test within the three domains assessed (reading, mathematics, and science), following a systematic booklet assembly and rotation procedure.[1] For this reason, PISA provides different plausible values (PV) of student scores rather than one measure of achievement. This strategy allows accounting for the uncertainty associated with the estimates of each student's achievements (Monseur and Adams, 2009; OECD, 2017c). Thus, as these PVs are considered as belonging to different outcome variables, they have been used as multiple response variables to monitor achievement in mathematics and reading.

To compare students' performances between immigrant and native students, we considered the following information for each wave:

- Country of the PISA assessment (COUNTRY): France (FR), Germany (DE), Italy (IT), Spain (ES), United Kingdom (UK).
- Immigrant status (IMMSTAT): according to OECD and PISA, immigrant students are defined as those whose both parents were born in a different country than the one where the student takes the test. Among immigrants, PISA distinguishes between second-generation (IMMSTAT $= 1$) and first-generation students (IMMSTAT$= 2$). Second-generation students are born in the country of PISA assessment, while first-

[1] The "booklet effect" refers to effects on students' performances in specific questions related to the number of items and the position of these items in the test booklet. See, for example, Gonzalez and Rutkowski (2010); Veldkamp and Sluijter (2019); OECD (2005) for details.

**Table 1** Descriptive statistics

| Wave | Average | Germany | Spain | France | UK | Italy |
|---|---|---|---|---|---|---|
| Number of students | | | | | | |
| 2009 | 13,690 | 3808 | 23,886 | 3707 | 10,539 | 26,510 |
| 2012 | 13,792 | 3497 | 23,693 | 4027 | 10,892 | 26,849 |
| 2015 | 7612 | 4983 | 6016 | 5384 | 11,003 | 10,674 |
| 2018 | 12,990 | 4177 | 33,244 | 5648 | 11,029 | 10,851 |
| Number of schools | | | | | | |
| 2009 | 572 | 225 | 886 | 168 | 482 | 1097 |
| 2012 | 610 | 227 | 901 | 226 | 507 | 1187 |
| 2015 | 344 | 255 | 197 | 250 | 547 | 470 |
| 2018 | 512 | 220 | 1,089 | 251 | 467 | 534 |
| % Non-native students (% 2nd generation) | | | | | | |
| 2009 | 7.02 (2.15) | 12.66 (8.46) | 8.54 (1.14) | 10.33 (7.77) | 5.89 (2.90) | 4.84 (1.07) |
| 2012 | 8.03 (2.36) | 9.84 (8.29) | 9.53 (0.98) | 12.17 (8.29) | 7.37 (3.04) | 6.11 (1.63) |
| 2015 | 10.49 (4.96) | 14.17 (11.36) | 9.89 (1.68) | 11.03 (7.54) | 11.76 (4.32) | 7.54 (3.18) |
| 2018 | 11.83 (5.88) | 19.58 (14.56) | 11.51 (4.58) | 13.85 (9.30) | 11.54 (5.78) | 9.07 (4.86) |
| Read: average reading score (std. dev.) | | | | | | |
| 2009 | 498.9 (86.07) | 511.6 (87.67) | 490.3 (83.32) | 512.7 (93.74) | 504.7 (86.42) | 500.5 (86.26) |
| 2012 | 505.4 (86.35) | 525.6 (83.16) | 498.7 (84.68) | 522.8 (96.71) | 509.8 (83.39) | 504.2 (86.92) |
| 2015 | 508.9 (86.15) | 529.1 (87.32) | 505.2 (78.36) | 517.5 (98.85) | 507.9 (84.03) | 498.2 (83.00) |
| 2018 | 493.8 (90.38) | 516.6 (97.92) | 487.5 (86.40) | 496.0 (98.00) | 511.3 (92.02) | 485.4 (89.59) |
| Math: average mathematics score (std. dev.) | | | | | | |
| 2009 | 499.9 (84.92) | 526.7 (90.67) | 495.8 (86.75) | 512.3 (89.34) | 499.8 (79.45) | 498.0 (83.00) |
| 2012 | 501.8 (85.68) | 530.2 (90.77) | 499.4 (83.97) | 509.7 (89.64) | 500.5 (84.41) | 499.6 (85.66) |
| 2015 | 505.1 (81.34) | 521.7 (81.69) | 495.3 (76.47) | 508.0 (85.11) | 502.8 (78.24) | 503.6 (83.83) |
| 2018 | 499.1 (81.23) | 516.5 (87.71) | 495.6 (78.17) | 496.5 (88.32) | 504.7 (79.97) | 498.8 (84.06) |

generation students are born in a different country. Natives (IMMSTAT= 0) refer to the remaining population.

Concerning students' immigrant backgrounds, PISA also provides information on how old first-generation immigrants were when moving to the destination country. Considering 2018, the average age of migration was 7.6 years, ranging from 7.3 in Italy to 9.7 in Germany. This information may be used to further classify first-generation students into other groups based on the time spent in the country or on whether they have attended at least their primary school in their destination country. However, since we are interested in the overall differences between native students and those with an immigrant background, we prefer to not consider the classification of immigrant students based on the so-called fractional generations in the analysis.[2]

Table 1 reports descriptive statistics calculated at the country level for each wave with respect to students' immigrant status and scores. The data set counts 240,417 students,

[2] To explore the role of the time spent in the country by students, we have estimated several specifications in which first-generation immigrants are divided according to their age of arrival. As expected, the results indicate that the lower is the age of arrival in the country the lower is the gap between first-generation and second-generation immigrant students. The results are available from the authors upon request.

with an average of 60,104 students for each wave.[3] In total, we have 10,186 schools, with an average of 2546 schools for each wave.

Moreover, to take into account the heterogeneity in student demographic and socioeconomic characteristics, the following control variables have been considered among the predictors:

- Language spoken at home (LHOMEDIFF): this indicator takes value 1 if the language spoken at home differs from the PISA assessment language.[4]
- SEX = M: indicator that takes value 1 for male students.
- Parental educational level (PARED): highest parental education in years of schooling.
- Parental occupational status (HISEI): highest parental occupational status based on the International Socioeconomic Index of occupational status (ISEI) (Ganzeboom & Treiman, 2003). Higher HISEI scores indicate higher occupational status.
- Family possession of cultural items (CULTPOSS): the PISA index of family cultural possessions is derived from what the students report on the availability of specific items at home such as classic literature, books of poetry, works of art, musical instruments, etc. Higher values indicate a higher family endowment of cultural items.
- Family possession of educational resources (HEDRES): the PISA index of home educational resources depends on the availability of items such as a desk to study at, a computer, educational software, a dictionary, etc. Higher values indicate higher availability of educational resources at home.

The LHOMEDIFF indicator helps to control for the variations in students' scores associated with their language proficiency and the level of integration of their families in the country. Additionally, it partially accounts for the influence of students' country of origin, especially for those coming from countries with colonial ties to the destination country. These students, such as immigrants in Spain from Spanish-speaking countries in Central and South America, are more likely to speak the language of assessment of the PISA test. While the best solution would be to use the information on students' country of origin, this data is only available for some countries, and the definition of country of origin varies widely. For instance, in Spain, we observe solely whether a student was born in the country of the PISA test, while Italy provides information on immigrants' origin limited to the European level (i.e. Europeans and non-Europeans).

Additionally, we include information on students' sex and the socioeconomic background of their families. These variables account for the differences in students' performance that may depend on families' socioeconomic endowment that, as we highlight in Sect. 2, are one of the most relevant determinants of students' educational attainment.

In addition to students' and families' socioeconomic factors, the high school attended and its characteristics can also impact students' academic performances. As described

[3] We have 68,450, 68,958, 38,060, and 64,949 records for 2009, 2012, 2015 and 2018, respectively. The decrease in sample size in 2015 is primarily due to a reduction in the number of sampled students in Spain. Unfortunately, we do not have information on the causes of this reduction. However, it is important to note that this reduction does not impact the level of coverage of the target population of enrolled students, as highlighted by OECD (2013a). Furthermore, as indicated in Tables 1 and 2, this change does not affect the average characteristics and performances of students, which remain stable across waves. Therefore, we are confident that our results should not be affected by this element.

[4] Dialects or regional languages are considered equal to the test assessment language.

**Table 2** Students' characteristics by country and wave

| Wave | Average | Germany | Spain | France | UK | Italy |
|---|---|---|---|---|---|---|
| Sex (% females) | | | | | | |
| 2009 | 50.56 | 51.08 | 49.89 | 52.31 | 51.66 | 50.41 |
| 2012 | 50.72 | 51.30 | 50.61 | 53.22 | 50.79 | 50.33 |
| 2015 | 51.33 | 51.11 | 51.03 | 52.34 | 51.66 | 50.74 |
| 2018 | 50.43 | 47.88 | 50.54 | 50.21 | 52.63 | 48.94 |
| LHOMEDIF: % with language spoken at home different from test language | | | | | | |
| 2009 | 3.78 | 8.74 | 3.20 | 4.67 | 3.18 | 3.70 |
| 2012 | 4.19 | 5.32 | 3.43 | 5.86 | 3.62 | 4.70 |
| 2015 | 6.73 | 9.59 | 3.66 | 5.74 | 5.03 | 9.39 |
| 2018 | 7.51 | 15.18 | 4.82 | 8.34 | 9.36 | 10.50 |
| HISEI: socioeconomic status (std. dev.) | | | | | | |
| 2009 | 47.77 (16.62) | 49.89 (15.47) | 46.81 (17.22) | 47.83 (16.97) | 49.63 (15.70) | 47.59 (16.43) |
| 2012 | 49.62 (21.30) | 52.26 (20.38) | 48.23 (21.60) | 53.03 (21.24) | 55.14 (20.84) | 47.76 (20.87) |
| 2015 | 52.54 (21.76) | 52.78 (20.23) | 49.28 (23.44) | 52.11 (21.49) | 56.89 (21.35) | 50.01 (21.26) |
| 2018 | 51.31 (22.50) | 52.71 (20.81) | 50.19 (23.14) | 51.02 (22.17) | 58.15 (21.24) | 47.42 (20.99) |
| PARED: highest parental education (std. dev.) | | | | | | |
| 2009 | 13.17 (3.23) | 14.44 (2.82) | 12.54 (3.71) | 13.00 (1.97) | 14.26 (1.87) | 13.13 (3.24) |
| 2012 | 13.39 (3.15) | 14.28 (3.19) | 12.82 (3.55) | 13.14 (1.95) | 14.29 (1.99) | 13.45 (3.18) |
| 2015 | 13.72 (2.88) | 14.48 (3.22) | 12.77 (3.78) | 13.46 (1.84) | 14.01 (2.11) | 13.73 (3.07) |
| 2018 | 13.79 (3.12) | 14.51 (3.50) | 13.61 (3.46) | 13.65 (1.98) | 14.08 (2.15) | 13.85 (3.11) |
| CULTPOSS: family cultural possession (std. dev.) | | | | | | |
| 2009 | 0.02 (0.89) | – 0.05 (0.94) | 0.22 (0.86) | – 0.18 (1.02) | – 0.23 (0.95) | – 0.02 (0.82) |
| 2012 | 0.15 (0.99) | 0.14 (0.99) | 0.24 (0.96) | – 0.11 (1.02) | – 0.22 (1.01) | 0.26 (0.96) |
| 2015 | 0.06 (1.01) | 0.12 (1.04) | 0.16 (0.98) | – 0.06 (0.99) | – 0.10 (1.05) | 0.21 (0.96) |
| 2018 | – 0.02 (0.94) | 0.16 (1.01) | 0.02 (0.92) | – 0.45 (1.04) | – 0.19 (1.02) | 0.20 (0.75) |
| HEDRES: family educational resources (std. dev.) | | | | | | |
| 2009 | – 0.00 (0.90) | 0.23 (0.86) | – 0.10 (0.90) | – 0.23 (0.95) | – 0.05 (0.92) | 0.11 (0.87) |
| 2012 | 0.08 (0.89) | 0.26 (0.85) | 0.08 (0.87) | – 0.16 (0.81) | 0.07 (1.00) | 0.10 (0.86) |
| 2015 | 0.09 (0.92) | 0.13 (0.86) | – 0.01 (0.89) | – 0.19 (0.82) | 0.07 (1.01) | 0.29 (0.87) |
| 2018 | 0.01 (0.94) | 0.05 (0.90) | – 0.09 (0.90) | – 0.19 (0.90) | 0.10 (1.01) | 0.28 (0.92) |

in Sect. 5, we account for differences among high schools by employing a multilevel approach and including the averages of students' characteristics at the high school level. Although the PISA database provides some information on high schools' characteristics, resources, and policies, we prefer to apply this analytical approach to account for all the characteristics that vary among high schools within the same country and wave. Indeed, we aim to measure the differences attributable to students' country of destination and immigrant status rather than focus on the specific policies implemented at the school level.

Table 2 reports descriptive statistics on control variables calculated at the country level for each wave.

In the following, we will handle five PVs estimates for each student as five different outcomes of the response variables related to the achievement in mathematics and reading. These values are modelled in a regression setting that combines multilevel analysis with Rubin's rule for multiple imputations (Rubin, 1987) to get robust estimates of the countries' position regarding the performance differentials with respect to students'

migratory background. To the best of our knowledge, the use of this approach is new in this framework and provides sensible insights on divergences in the effect of immigrant background across countries over time.

## Model

To analyse the differences in tests scores between native and immigrant students we define $y^m_{(ijgt(s))}$ as the $m$th ($m = 1, \ldots, 5$) PV of the score of student $i$ ($i = 1, \ldots, N$) belonging to school $j$ ($j = 1, \ldots, J$) in country $g$ ($g = 1, \ldots, 5$) at time $t$ ($t = 1, \ldots, 4$) in the $s$th subject ($s = 1$, reading; $s = 2$, mathematics). To account for the clustering of students in schools, we fit 5 two-level regression models considering students as level-1 units and schools as level-2 units (Rabe-Hesketh & Skrondal, 2012). Given this framework, students' scores are modelled as follows:

$$y^m_{ijt(s)} = \alpha_m + \beta'_m X_{ijt} + \gamma'_m Z_j + \zeta'_m D_{igt} + u_{mj} + e_{imjt}, \tag{1}$$

where $X_{ijt}$ is the vector of students' and families' socioeconomic and cultural characteristics shown in Table 1, and $Z_j$ includes the average characteristics of students at the high school level. The vector of dummy variables $D_{igt}$ includes the interactions between students' immigrant status (native, first-generation immigrants, and second-generation immigrants), country, and wave. Thus, the $\zeta^m$ parameters measure the performance gap between natives and immigrants across countries from 2009 and 2018 with respect to the baseline defined as the performance of German natives in 2009. The term $u_{mj} \sim \mathcal{N}(0, \sigma_{u,m})$ denotes the random intercepts that capture the differences in average performances among schools.

The fixed and random components of the model allow us to differentiate between the variability scores due to differences between schools to the one related to within-individual variability. As highlighted by Mundlak (1978), the combination of random intercepts at the high school level $u_{mj}$ and the high school-level means $Z_j$ accounts for time-invariant characteristics at the high school level that may affect students' performances [see also Grilli and Rampichini (2009)]. Therefore, our empirical approach allows us to capture divergences in the effect of the different immigrants' backgrounds across countries and waves that do not depend on high school characteristics and students' socioeconomic background.

Another possible solution to account for the school effect would have been to estimate a fixed-effects model that included interaction terms among student migration status, country, survey wave, and the entire set of fixed effects observed at the school level. Such an approach would have been more robust and required fewer assumptions compared to the multilevel model (Wooldridge, 2010). However, it would not have allowed for direct comparisons of students' performances against the baseline. To assess the robustness of our results, we provide a comparison of the estimated coefficients related to students' socioeconomic characteristics between the multilevel and fixed-effects models in Table I in the supplementary materials. Notably, we observe that the estimated coefficients are nearly identical between the two specifications, providing support for the multilevel approach and indicating that we account for time-invariant characteristics at the school level.

An additional element that may affect our results is the unobserved heterogeneity related to the compositional differences in students' populations across countries and over time such as students' countries of origin, average age of arrival in the country, and reasons for migration. However, two elements support our empirical specification. First, thanks to the multilevel approach, the composition of students population should vary within high schools to affect our results. Second, based on the descriptive statistics in Table 2, we can note that students' average characteristics remain stable in the considered time frame.[5] Moreover, to ensure the robustness of our results, we have fitted five separate models, one for each country to better account for differences in students' composition within the same country.[6] The results obtained are similar to those obtained with the overall model. The only notable difference is found when considering the model estimated using only French data, where the results indicate lower scores in general for all the categories of students and a positive trend in the last two waves.[7] Therefore, we prefer our estimation approach as it allows for a direct comparison of students' performances across the observed countries.

We estimate the model five times for $m = 1, \ldots, 5$ to account for the variability in the PVs associated with student $i$. The parameter estimates $[\alpha, \beta, \gamma, \zeta, \sigma_u^2]^m$ and the related standard errors are pooled using Rubin's rule (Rubin, 1987). Specifically, denoting with $\hat{\theta}^m$ the estimate of parameter $\theta$ for the $m$th model, its pooled estimate $\bar{\theta}$ is obtained as the average estimate over the five models:

$$\bar{\theta} = M^{-1} \sum_{m=1}^{M} \hat{\theta}^m. \tag{2}$$

The related standard error is defined as:

$$\mathrm{SE}_p(\hat{\theta}) = \sqrt{W + B + \frac{B}{M}}, \tag{3}$$

where $\mathrm{SE}_p(\hat{\theta})$ combines the within-estimate variance ($W$) and the between-estimate variance ($B$), defined as follows:

$$W = \frac{\sum_{m=1}^{M} \mathrm{SE}(\theta)^2}{M}, \tag{4}$$

$$B = \frac{\sum_{m=1}^{M} (\hat{\theta} - \bar{\theta})^2}{M - 1}. \tag{5}$$

The combined use of the estimates of the parameters and their related standard errors are considered to inspect the time trends in the selected countries regarding the differences in achievement between native and immigrant students.

The multilevel model described has been fitted with the mixed routine for linear mixed-effects models available in STATA 16.1.

[5] We observe relevant differences only in the last wave with respect to the variables LHOMEDIF and HEDRES in Germany and in CULTPOSS in Spain.

[6] We thank the anonymous reviewer for suggesting this solution.

[7] The results are available from the authors upon request.

## Results

The results regarding students' reading performances are reported in Table 3, while those on mathematics scores are shown in Table 4. In each table, the coefficients related to variables $Z_j$ (i.e. students' average characteristics at the school level) are indicated as school: $X_{ijt}$, where $X_{ijt}$ is one covariate at the student level (i.e. school: PARED for the average of PARED). Results show that the between-school variability is a relevant source of the total variability. Indeed, the estimates show that about 32% of the differences in

**Table 3** Pooled multilevel model: reading

| Variable | Coeff. | se | *p*-value | Variable | Coeff. | se | *p*-value |
|---|---|---|---|---|---|---|---|
| Intercept | 390.23 | 6.14 | 0.000 | 2nd.de.18 | 5.18 | 4.96 | 0.297 |
| PARED | 0.05 | 0.07 | 0.429 | 2nd.es.09 | – 10.97 | 5.60 | 0.050 |
| HISEI | 0.48 | 0.01 | 0.000 | 2nd.es.12 | – 28.20 | 5.89 | 0.000 |
| CULTPOSS | 10.61 | 0.24 | 0.000 | 2nd.es.15 | 4.82 | 9.88 | 0.625 |
| HEDRES | 3.98 | 0.22 | 0.000 | 2nd.es.18 | – 12.39 | 3.70 | 0.001 |
| LHOMEDIF = yes | – 15.09 | 0.96 | 0.000 | 2nd.fr.12 | – 1.66 | 6.08 | 0.785 |
| Sex = m | – 20.63 | 0.38 | 0.000 | 2nd.fr.15 | – 0.77 | 6.31 | 0.902 |
| School: PARED | 1.39 | 0.49 | 0.005 | 2nd.uk.09 | 13.65 | 5.61 | 0.015 |
| School: HISEI | 1.87 | 0.08 | 0.000 | 2nd.uk.12 | – 3.50 | 5.33 | 0.511 |
| School: CULTPOSS | 11.58 | 1.62 | 0.000 | 2nd.uk.15 | – 10.36 | 5.11 | 0.043 |
| School: HEDRES | 35.49 | 1.63 | 0.000 | 2nd.uk.18 | – 15.06 | 4.75 | 0.002 |
| School: LHOMEDIF | – 17.71 | 4.65 | 0.000 | 2nd.it.09 | – 14.36 | 5.47 | 0.009 |
| School: male | – 24.75 | 2.23 | 0.000 | 2nd.it.12 | – 19.58 | 4.86 | 0.000 |
| Nat.de.09 (bas.) | – | – | – | 2nd.it.15 | – 41.84 | 6.24 | 0.000 |
| Nat.de.12 | 2.56 | 3.81 | 0.502 | 2nd.it.18 | – 34.23 | 4.67 | 0.000 |
| Nat.de.15 | 12.27 | 3.96 | 0.002 | 1st.de.09 | – 14.66 | 6.13 | 0.017 |
| Nat.de.18 | 5.45 | 3.90 | 0.162 | 1st.de.12 | – 15.60 | 11.73 | 0.184 |
| Nat.es.09 | – 0.11 | 3.21 | 0.972 | 1st.de.15 | – 19.75 | 10.15 | 0.052 |
| Nat.es.12 | – 3.77 | 3.17 | 0.234 | 1st.de.18 | – 38.06 | 6.68 | 0.000 |
| Nat.es.15 | 5.16 | 3.93 | 0.189 | 1st.es.09 | – 38.31 | 3.74 | 0.000 |
| Nat.es.18 | – 10.82 | 3.03 | 0.000 | 1st.es.12 | – 36.79 | 3.55 | 0.000 |
| Nat.fr.09 | 26.47 | 4.15 | 0.000 | 1st.es.15 | – 16.77 | 6.07 | 0.006 |
| Nat.fr.12 | 18.29 | 3.92 | 0.000 | 1st.es.18 | – 28.98 | 3.53 | 0.000 |
| Nat.fr.15 | 13.04 | 4.03 | 0.001 | 1st.fr.09 | 13.03 | 8.85 | 0.141 |
| Nat.fr.18 | 6.30 | 3.81 | 0.098 | 1st.fr.12 | – 9.39 | 7.50 | 0.211 |
| Nat.uk.09 | 8.96 | 3.32 | 0.007 | 1st.fr.15 | – 6.36 | 7.55 | 0.400 |
| Nat.uk.12 | – 3.77 | 3.26 | 0.247 | 1st.fr.18 | – 13.31 | 6.27 | 0.034 |
| Nat.uk.15 | – 9.24 | 3.51 | 0.009 | 1st.uk.09 | 4.16 | 5.80 | 0.473 |
| Nat.uk.18 | –7.98 | 3.44 | 0.020 | 1st.uk.12 | – 1.18 | 5.15 | 0.819 |
| Nat.it.09 | – 4.09 | 2.99 | 0.172 | 1st.uk.15 | – 23.91 | 5.54 | 0.000 |
| Nat.it.12 | – 7.50 | 3.03 | 0.013 | 1st.uk.18 | – 9.61 | 4.68 | 0.040 |
| Nat.it.15 | – 22.04 | 4.16 | 0.000 | 1st.it.09 | – 29.65 | 3.80 | 0.000 |
| Nat.it.18 | – 27.17 | 3.22 | 0.000 | 1st.it.12 | – 26.27 | 3.92 | 0.000 |
| 2nd.de.09 | – 9.04 | 4.72 | 0.055 | 1st.it.15 | – 57.56 | 5.63 | 0.000 |
| 2nd.de.12 | – 1.24 | 5.87 | 0.833 | 1st.it.18 | – 38.63 | 4.89 | 0.000 |
| 2nd.de.15 | 3.72 | 5.71 | 0.515 | | | | |

| Random-effect parameters | Estimate | se |
|---|---|---|
| Between schools std. dev. | 34.41 | 0.14 |
| Residual std. dev. | 71.49 | 0.07 |

**Table 4** Pooled multilevel model: mathematics

| Variable | Coeff. | se | *p*-value | Variable | Coeff. | se | *p*-value |
|---|---|---|---|---|---|---|---|
| Intercept | 353.40 | 5.85 | 0.000 | 2nd.de.18 | – 16.33 | 5.21 | 0.002 |
| PARED | 0.22 | 0.07 | 0.002 | 2nd.es.09 | – 19.75 | 5.52 | 0.000 |
| HISEI | 0.50 | 0.01 | 0.000 | 2nd.es.12 | – 27.15 | 6.09 | 0.000 |
| CULTPOSS | 8.58 | 0.23 | 0.000 | 2nd.es.15 | – 19.35 | 8.74 | 0.027 |
| HEDRES | 5.31 | 0.23 | 0.000 | 2nd.es.18 | – 26.94 | 4.82 | 0.000 |
| LHOMEDIF = yes | – 7.83 | 1.21 | 0.000 | 2nd.fr.12 | – 28.55 | 5.64 | 0.000 |
| Sex = m | 19.33 | 0.43 | 0.000 | 2nd.fr.15 | – 19.80 | 5.76 | 0.001 |
| School: PARED | 2.59 | 0.47 | 0.000 | 2nd.uk.09 | – 10.47 | 5.96 | 0.079 |
| School: HISEI | 1.82 | 0.08 | 0.000 | 2nd.uk.12 | – 28.03 | 5.36 | 0.000 |
| School: CULTPOSS | 5.56 | 1.57 | 0.000 | 2nd.uk.15 | – 26.22 | 5.15 | 0.000 |
| School: HEDRES | 37.79 | 1.55 | 0.000 | 2nd.uk.18 | – 32.73 | 5.79 | 0.000 |
| School: LHOMEDIF | – 12.00 | 4.12 | 0.004 | 2nd.it.09 | – 29.13 | 5.49 | 0.000 |
| School: male | – 3.09 | 2.08 | 0.138 | 2nd.it.12 | – 27.63 | 4.64 | 0.000 |
| Nat.de.09 (bas.) | – | – | – | 2nd.it.15 | – 43.57 | 5.61 | 0.000 |
| Nat.de.12 | – 4.45 | 3.88 | 0.251 | 2nd.it.18 | – 30.14 | 6.14 | 0.000 |
| Nat.de.15 | – 6.40 | 3.91 | 0.101 | 1st.de.09 | – 16.62 | 7.14 | 0.020 |
| Nat.de.18 | – 9.39 | 3.84 | 0.014 | 1st.de.12 | – 25.28 | 11.54 | 0.028 |
| Nat.es.09 | – 1.84 | 3.17 | 0.561 | 1st.de.15 | – 37.02 | 8.39 | 0.000 |
| Nat.es.12 | –11.49 | 3.12 | 0.000 | 1st.de.18 | – 33.68 | 6.80 | 0.000 |
| Nat.es.15 | – 12.83 | 3.90 | 0.001 | 1st.es.09 | – 52.78 | 3.69 | 0.000 |
| Nat.es.18 | – 12.01 | 3.18 | 0.000 | 1st.es.12 | – 54.67 | 3.54 | 0.000 |
| Nat.fr.09 | 15.98 | 4.14 | 0.000 | 1st.es.15 | – 42.09 | 5.78 | 0.000 |
| Nat.fr.12 | – 4.21 | 3.93 | 0.284 | 1st.es.18 | – 42.84 | 4.64 | 0.000 |
| Nat.fr.15 | – 5.06 | 3.74 | 0.177 | 1st.fr.09 | 0.67 | 9.44 | 0.943 |
| Nat.fr.18 | – 8.46 | 3.79 | 0.026 | 1st.fr.12 | – 30.04 | 7.22 | 0.000 |
| Nat.uk.09 | – 9.20 | 3.32 | 0.006 | 1st.fr.15 | – 23.44 | 6.74 | 0.001 |
| Nat.uk.12 | – 27.42 | 3.22 | 0.000 | 1st.fr.18 | – 23.07 | 8.80 | 0.009 |
| Nat.uk.15 | – 27.95 | 3.29 | 0.000 | 1st.uk.09 | – 12.34 | 5.95 | 0.038 |
| Nat.uk.18 | – 28.75 | 3.44 | 0.000 | 1st.uk.12 | – 15.72 | 5.01 | 0.002 |
| Nat.it.09 | – 17.25 | 2.98 | 0.000 | 1st.uk.15 | – 35.32 | 4.36 | 0.000 |
| Nat.it.12 | – 21.88 | 3.03 | 0.000 | 1st.uk.18 | – 21.91 | 6.63 | 0.001 |
| Nat.it.15 | – 29.44 | 3.51 | 0.000 | 1st.it.09 | – 40.07 | 3.86 | 0.000 |
| Nat.it.18 | – 26.99 | 3.43 | 0.000 | 1st.it.12 | – 34.45 | 3.93 | 0.000 |
| 2nd.de.09 | – 12.34 | 4.61 | 0.007 | 1st.it.15 | – 53.02 | 5.84 | 0.000 |
| 2nd.de.12 | – 17.56 | 5.65 | 0.002 | 1st.it.18 | – 37.74 | 5.55 | 0.000 |
| 2nd.de.15 | – 22.21 | 4.90 | 0.000 | | | | |

| Random-effect parameters | Estimate | se |
|---|---|---|
| Between schools std. dev. | 34.16 | 0.15 |
| Residual std. dev. | 69.62 | 0.14 |

students' scores that are not explained by the heterogeneity of students in terms of socioeconomic characteristics and immigrant status are explained by differences between schools.

Concerning students' reading competencies, female students score nearly 20 points higher than males. In line with the previous literature, higher reading performance is positively associated with the students' socioeconomic status, educational resources,

and the possession of cultural items. A small positive association is also observed for the level of parental education. These positive effects are also found in relation to students' average characteristics at the high school level. Conversely, performances tend to decline when the language spoken at home differs from the test language, as well as in high schools with a higher proportion of males and students who do not speak the test language at home.

These findings remain consistent when examining students' mathematical competencies. Indeed, from Table 4 we can see comparable effects for parental education, socioeconomic status, cultural possessions, and educational resources both at the individual and school level. We also observe a small negative effect related to the differences between the language spoken at home and the assessment language. A significant difference is related to the effect of students' sex: males outperform females by about 19 points in mathematics (the opposite was observed for reading).

To sum up, all indicators reflecting students' access to socio-cultural resources have the expected impact on their competencies. On average, parental education and educational resources have a pronounced effect on mathematics scores, while the impact of cultural possessions is stronger in determining divergences in reading.

As outlined in Sect. 5, we measure the differences between native and immigrant students by including the set of country $\times$ year $\times$ immigrant status interaction terms. To allow the comparison of students' performances across countries and waves we consider the average achievements of native Germans in 2009 as the baseline. Consequently, each interaction coefficient identifies the distance in performance for a particular category of students from the average performance of a native German in 2009. Positive values highlight higher values of learning outcomes, whereas negative values indicate lower performances compared to the baseline.[8]

Based on the results in Table 3, it is evident that immigrant students in Spain and Italy are those that obtain the lowest scores, along with second-generation immigrants in the United Kingdom in the 2018. Indeed, all the coefficients estimated for first-generation students in Italy and Spain are significantly lower than the baseline, with the lowest performances observed among students in Italy in 2015. In France, only students in 2018 registered a performance significantly lower than the baseline, while the worst performances in Germany and United Kingdom are observed in 2018 and 2015. Overall, the point estimates show that first-generation students have performances that are below the baseline, with the best performance observed among students in France in 2009. Similarly, the results for second-generation immigrants indicate that students in Italy and Spain have the worst performances, except for those in Spain in 2015, where the point estimate is above the baseline. The highest performances were observed among students in the United Kingdom and France in 2009. When considering natives, Italy exhibits the worst performance within the time frame considered, with all the parameter estimates that are significantly lower than the baseline, except for students in the 2009 wave. For Spain, only the 2018 wave registers performances significantly lower than the baseline, while the same holds true for the last two waves in the United Kingdom. The

[8] To provide a graphical comparison of students' performances Figs. S1 and S2 in Additional file 1 display the parameters estimated for students' scores in reading and mathematics with the associated 95% confidence interval.

analysis of the differences between countries in the reading competencies shows that the highest ranks are held by natives in France between 2009 and 2015 and in Germany in 2015. Among immigrants, only second-generation students in the United Kingdom in 2009 registered a significant positive difference with respect to the baseline.

The coefficients estimates regarding mathematics performances in Table 4 show a different picture with respect to those observed for reading. If we consider the natives, we can note that most of coefficients are significantly lower than the baseline, with the lowest ranks occupied by students in Italy and United Kingdom. Concerning first-generation students, most estimated parameters are significantly below the baseline. The only group that exhibits values close to the baseline is the one estimated for the 2009 wave in France. Conversely, students in Italy and Spain consistently achieve lower rankings, while the best performances are observed in the early waves in France, the United Kingdom, and Germany. As for second-generation students, the highest rankings are observed among students in France and the United Kingdom in 2009. Indeed, the parameter estimates for these students do not significantly differ from the baseline. However, all the other estimated coefficients for second-generation students are significantly lower than zero.

Focusing on the analysis of students' performance trends, Figs. 2, 3, 4, 5, 6 outline, for each country and dependent variable, the variation in students' achievements across waves measured by the immigrant status–country–wave fixed effects.

From Fig. 2, we observe the following trends in Germany. Natives' performances in reading had an increasing trend from 2009 to 2015 and then decreased in 2018, although it remained above the value registered in 2009. Second-generation students' performances improved over time, reaching the same level as natives in 2018. However, this improvement is primarily due to a decline in natives' performances rather

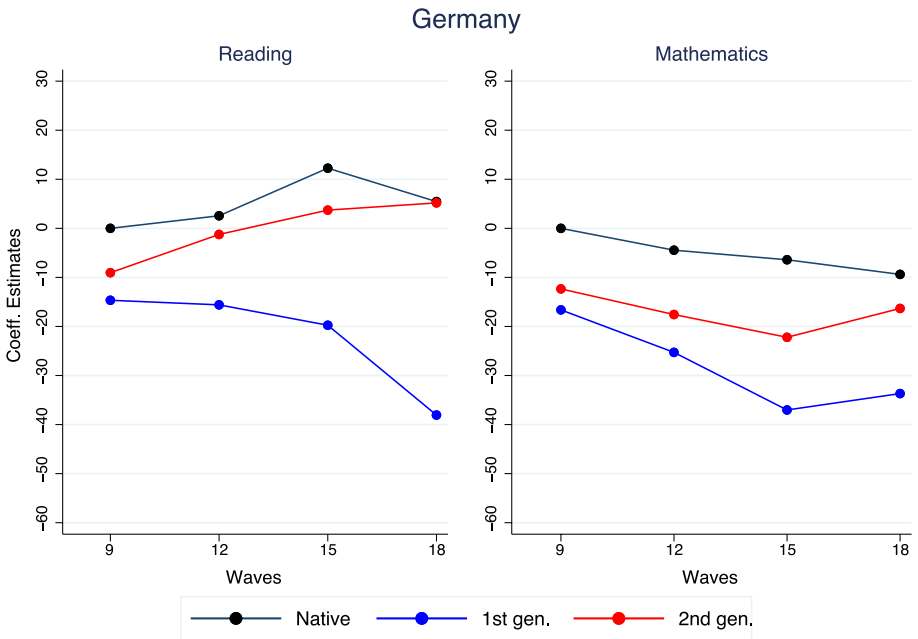


**Fig. 2** Germany: estimated parameters by immigrant status and wave. The figure shows immigrant status–country–wave fixed effects estimated via the pooled multilevel models in Tables 3 and 4

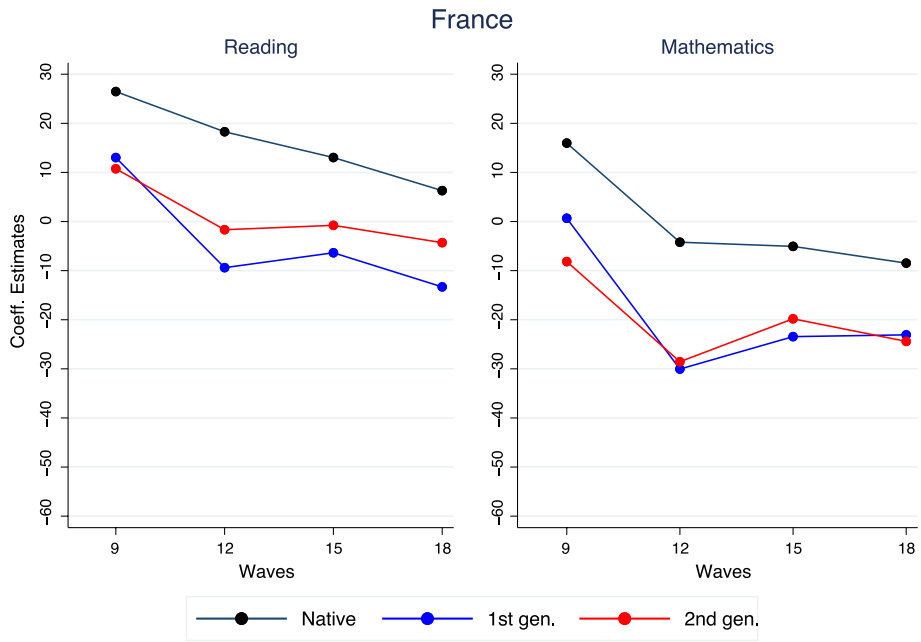


**Fig. 3** France: estimated parameters by immigrant status and wave. The figure shows immigrant status–country–wave fixed effects estimated via the pooled multilevel models in Tables 3 and 4

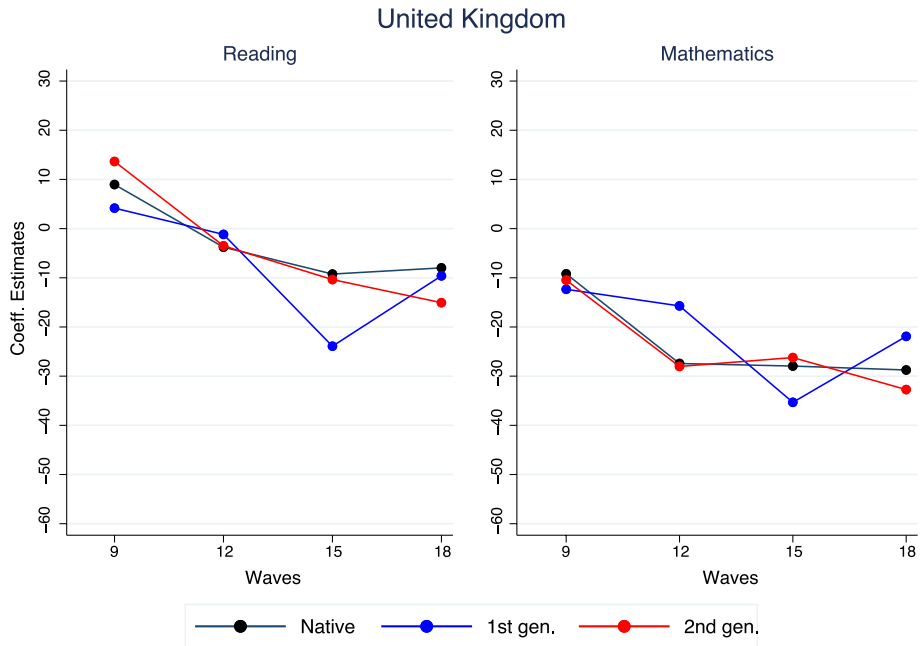


**Fig. 4** United Kingdom: estimated parameters by immigrant status and wave. The figure shows immigrant status–country–wave fixed effects estimated via the pooled multilevel models in Tables 3 and 4

than an increase in scores among second-generation students. First-generation students consistently underperform natives and second-generation students with a substantial reduction in performances in the last wave. This reduction may be attributed to changes in the composition of first-generation students in the last wave. Notably,

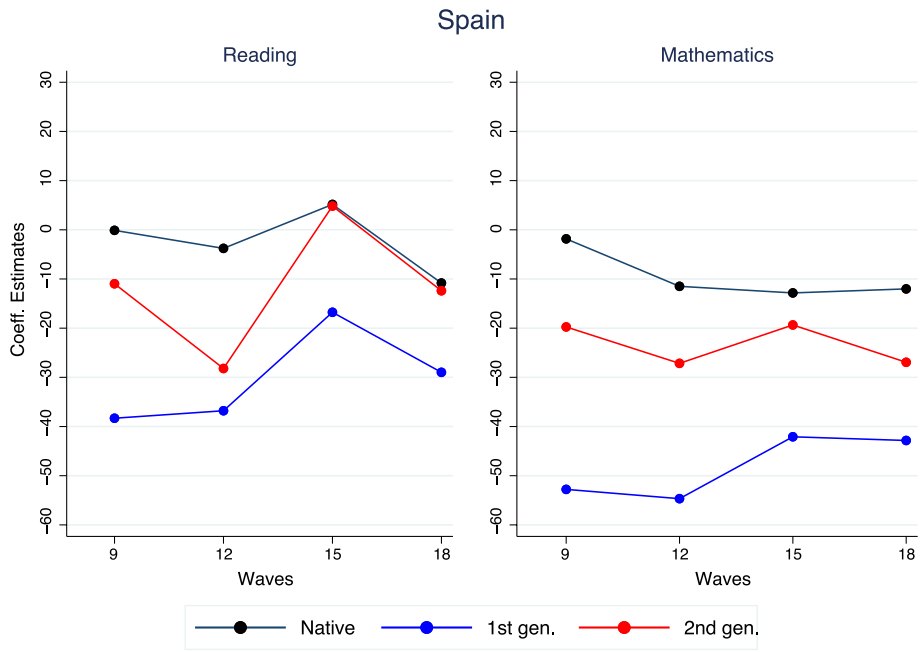


**Fig. 5** Spain: estimated parameters by immigrant status and wave. The figure shows immigrant status–country–wave fixed effects estimated via the pooled multilevel models in Tables 3 and 4

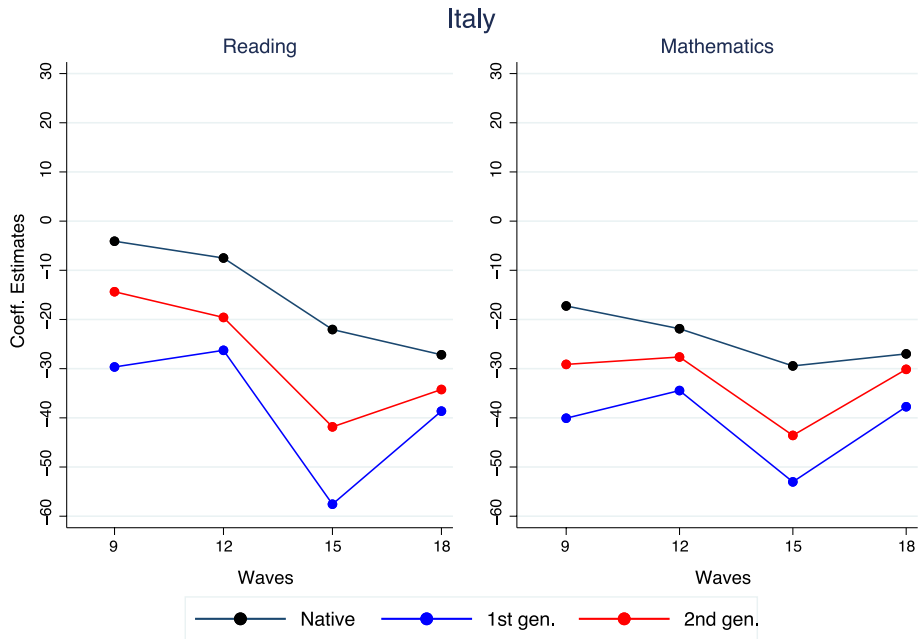


**Fig. 6** Italy: estimated parameters by immigrant status and wave. The figure shows immigrant status–country–wave fixed effects estimated via the pooled multilevel models in Tables 3 and 4

as shown in Table 2, we observe an increase in the share of students that do not speak German at home (from 9.59% in 2015 to 15.18% in 2018) along with a decrease in families' educational resources as measured by the HEDRES indicator (from 0.13 in 2015 to 0.05 in 2018). These elements may be related to events such as the Syrian

refugee crisis and Germany's role as destination country for asylum seekers. However, due to incomplete data on students' country of origin, we cannot control for this element in the analysis.[9] In mathematics, natives' performances remain relatively stable over time. This trend is also observed among second-generation students until 2015. Between 2015 and 2018, second-generation students' performances experienced growth and approached levels similar to native students. For first-generation students, we can note a decreasing trend that reached its minimum in 2015 while staying stable in 2018. The results suggest that the German system has reduced the gap between native and second-generation students. In contrast, the one between natives and first-generation has considerably widened. However, as previously mentioned, this element may be due to a change in the internal composition of first-generation students' group.

Figure 3 shows the parameter estimates for students' in France. Despite showing a descending trend, French natives' register the best performances in reading among all students, especially when considering early waves. First- and second-generation immigrants saw a decline in performances from 2009 to 2012, but they stabilised thereafter. From 2012 onward, second-generation students attained performances comparable to those of native students in Germany in 2009. Conversely, first-generation students consistently underperformed, reaching their lowest point in 2018. With respect to mathematics scores, we observe a reduction in students' performances between 2009 and 2012, while students' performances remain stable after 2012. In particular, natives exhibit performances similar to the baseline, while the coefficients estimated for immigrant students are markedly below zero. Moreover, similar to the findings for reading scores, these two groups showed comparable performances in the time frame considered. This evidence suggests that first- and second-generation students in France do not differ in terms of their performances in reading and mathematics.

For the United Kingdom, we found the narrowest gap between natives' and immigrants' performances in reading. Except for the estimate for first-generation immigrants in 2015, the estimated coefficients are similar across different groups. However, the observed trends do not show significant improvement in students' performances. These findings are also confirmed for mathematics, as students with an immigrant background do not show significant differences compared to natives. However, the results indicate that, on average, students in the United Kingdom achieve lower scores in mathematics compared to the baseline.

Therefore, in traditional immigration countries such as France and Germany first-generation immigrants consistently lag behind natives and second-generation immigrants while their performance levels are comparable to those of native students in the United Kingdom. Moreover, there is a general decline in students' performances, except for the reading performances of natives and second-generations in Germany. Second-generation students show similar performance levels compared to natives in both Germany and the United Kingdom while showing more similarities with first-generation students in France.

[9] We appreciate the valuable input from the anonymous referees who brought attention to this aspect.

Findings from Spain in Fig. 5 indicate a decline in students' performances compared to the 2009 wave. Regarding reading, second-generation immigrants and natives showed similar performances in 2015 and 2018. This element can be related to the reduction in the share of foreigners in the resident population started in 2013 after the Great Recession (see Fig. 1). Indeed, since immigration in-flows have reduced, we can expect that the population of second-generation immigrants is composed by students that have spent an higher number of years in the host country and, therefore, may benefit from having attended their first-cycle of education in the destination country.[10] Considering first-generation students, the results for both dependent variables show an increasing disparity between first-generation students and the other groups. Unlike previous cases, first-generation immigrants in Spain exhibited a positive trend in performances compared to 2009. However, their scores still fall behind those of native students, especially in terms of mathematics achievement.

In Italy, the differences between natives and students with immigrant backgrounds seem more substantial in the earlier waves. In 2018, immigrants and natives exhibit very similar performances, particularly in mathematics where the differences between natives and second-generation students are minimal. However, this similarity seems to be driven, at least for reading, by a decline in natives' performances. It is important to note that, across all groups considered, students' competencies in Italy fall significantly below the value observed for native students in Germany in 2009. However, we can highlight an improvement in immigrants' performances in the last wave. It is worth mentioning that in 2009, there was a noticeable gap in competencies between native students and first-generation immigrants in both mathematics and reading. In contrast, in 2018, the educational system appears to be less unbalanced. Additionally, the observed gap between the average level of competencies in reading and mathematics, which was particularly evident among natives in 2009 and 2012, appears to have narrowed in the last two waves.

## Discussion

The findings described in the previous section provide insights into countries' pathways that countries have taken to reduce educational inequalities among students with different migratory backgrounds. It is interesting to discuss them moving from our research questions (RQ).

With respect to the first RQ (*How do educational outcomes in mathematics and reading differ among students with immigrant backgrounds in new and old immigrant-destination countries?*), the findings suggest that, on average, native students in both traditional and new destination countries perform better in both learning outcomes, followed by second- and first-generation immigrants. However, in line with previous literature, the results change depending on the country and the skill area being considered.

In new destination countries, Spain shows a different pattern between students' reading and mathematics performances. While the gap between native and immigrant students remains stable in mathematics, there is an improvement in reading performances

[10] We thank the anonymous reviewers for suggesting this element.

of second-generation students, reaching levels similar to those of natives in the last two waves. In Italy, we observe a similar pattern (both for mathematics and reading), with a noteworthy reduction in the performance gap between natives and immigrants in the last wave.

For traditional destination countries, the results show similar patterns in both skills areas for the United Kingdom and France. In both countries, we observed a reduction in students' performances across all the groups considered. In contrast, Germany is the only country where reading performances improve for native and second-generation students, while experiencing a decline in mathematics scores similar to the other countries.

Regarding the second RQ (*Have countries experienced a reduction in inequalities linked to immigrant status over time?*), the results show the persistence of the gap between native and immigrant students. These differences are particularly pronounced in the earlier waves and if we compare natives and first-generation students. However, each country presents a different picture regarding the path toward narrowing the gaps.

In the case of new destination countries, we observe a reduction in the gap in Spain. First-generation students' performances have improved from 2009 to 2018, while native and second-generation students have maintained a performance level similar to that of 2009 in the last wave. Moreover, since 2012, second-generation students' results are in line with those of natives. As for Italy, the gap between immigrant and native students has reduced, especially in the last wave. This reduction can be attributed to an improvement in first- and second-generation students' results and a reduction in the performances of natives. It is worth noting that students in Italy exhibit some of the lowest performances levels, particularly in reading competencies. Nonetheless, this improvement in immigrant students' relative performances may signal an improvement in the integration of immigrant students.

In traditional immigration countries, we find that the United Kingdom, with its colonial history and international language, appears to have a less pronounced imbalance, with students' performances being very similar across the groups. In France, there is a high degree of similarity between the achievements and trends observed for first- and second-generation students, suggesting that immigrant students struggle to benefit from the time spent in the hosting country, possibly indicating a lack of integration process. However, it is noteworthy that students in France register some of the highest performances level observed in the data, along with native and second-generation students in Germany. As for Germany, the results show an enhancement in reading performances for native and second-generation students. However, although the gap between these two groups has reduced, the country continues to exhibit persistent differences between first-generation immigrants and other students, which have worsened over time. Therefore, the results show that even if the educational systems in traditional immigration countries had more time to foster integration, they have different results in terms of effectively reducing inequalities among students with different immigrant backgrounds.

In line with the previous literature [see, for example, OECD (2012a)], the results show that the gap in educational attainment between native and immigrant students is still relevant, even after accounting for students' socioeconomic background, individual characteristics, compositional variables and other unobservable factors shared by student who experience the same school environment. Moreover, the evidence regarding families'

economic and cultural resources confirms the existence of a 'double origin gap' affecting less-advantaged students, which stems from both their migratory and socioeconomic backgrounds [see, for example, Gabrielli et al. (2022)]. However, this gap is less pronounced for second-generation immigrants, as highlighted by Azzolini et al. (2012), who benefit from being raised in the host country.

Looking at the third RQ (*Are there countries that tend to fill the gap in academic performances between first- and second-generation immigrants?*), the results show that, in general, second-generation immigrants achieve better results compared to first-generation ones. This element can be attributed to the additional time spent in the country and the opportunity to complete their initial schooling in the destination country.

Among the traditional destination countries, the United Kingdom show the best results in reducing the gap between immigrant students and natives. First- and second-generation immigrants in the United Kingdom achieve results similar to those of natives within the time frame considered. On the other hand, Germany shows a widening gap in performance between immigrant students, primarily due to an improvement in second-generation students' performances and a decline in the achievements of first-generation students. Germany has the largest performance gap among the countries considered, especially in reading skills in 2018. As discussed in Sect. 6, an explanation for the reduction in first-generation immigrants' performances in 2018 can be found in the modification of the composition of immigrant students populations, as signalled by the reduction of the share of students that speak German at home and families' educational resources. In France there is not significant difference between first- and second-generation immigrants, as both groups consistently lag behind native students over time.

We observe two different patterns between Italy and Spain here considered as new destination countries. Italy shows a reduction in the performance gap between immigrants and natives in the last wave. This reduction is evident for both first- and second-generation students, who have improved their performances compared to natives. However, the gap between students with immigrant backgrounds has narrowed only in reading while remaining stable in mathematics. In Spain, the gap between first- and second-generation students persists across the waves for both educational outcomes. However, from 2009 to 2018, the gap has reduced in magnitude thanks to an improvement in first-generation students' performances between 2012 and 2015.

To sum up, the evidence suggests that four out of the five examined countries are still far from closing the gap between native and immigrant students. Factors such as colonial history, linguistic differences, and the socioeconomic context of each country contribute to the definition of completely different paths toward the integration of immigrant students. These findings, reached using a novel approach that accounts for differences among students' characteristics and high schools, confirm results that have already been discussed in the literature but also provide valuable insights into the educational inequalities and the dynamics of immigrant student integration across countries and over time.

## Conclusions

The response of school systems to migration challenges has a substantial impact on the socioeconomic conditions of countries in terms of inclusion and social welfare. In recent years, tens of thousands of migrants and asylum seekers, including many school-aged

children, have fled their countries of origin to find safety and a better life in Europe. Given this phenomenon, it is crucial to assess whether European educational systems are suited to facilitate the integration of immigrant students (and the future workforce) into their new communities. Can these systems effectively prepare all students for a society in which people are willing to collaborate with others from different cultural backgrounds?

In this regard, the analysis proposed has provided valuable insight into the gaps in academic achievements between native and immigrant students, as well as their evolution over time, shedding light on the position of the countries in the path towards the integration of immigrant students. However, as previously mentioned, these performance differences may be influenced by various factors, including the characteristics of high school systems and integration policies at the country level. Successful immigrants integration requires collaboration across different policy domains and between central and local government authorities. Despite the analysis of the effectiveness of these policies is behind the goals of this paper, this Section offers reflections on the elements that should characterise such policies.

According to OECD (2015a), several aspects must be considered to foster immigrant integration in schools. These include offering high-quality early childhood education, and ensuring that all teachers, not just specialists, are prepared to teach in classrooms with immigrant students. Indeed, skilled and well-supported teachers are crucial for successful integration as they can tailor their instructional approaches to reflect the diversity of their student populations and help all students achieve the educational goals and standards of the host country. Moreover, the prompt provision of language instruction can enhance the integration of immigrants since, as our analysis highlights, language barriers are a relevant cause of academic delays.

Moreover, an examination of the education policies implemented in the selected countries reveals that they pushed forward different priorities to reduce disparities and improve efficiency. Germany has focused on specific policies for integrating disadvantaged students and enhancing external school evaluation practices. Spain has prioritised policies to strengthen school autonomy, principal leadership, and vocational education and training. Italy, has recently undertaken reforms to promote quality by addressing the inefficiencies of its educational system, such as promoting school autonomy and improving staff recruitment and evaluation. France has implemented a wide range of policies to reduce social inequalities by promoting a new culture of evaluation, less academic training of teachers, and improved counselling system between secondary and higher education. The United Kingdom has a long-standing tradition of implementing policies to monitor and support underperforming schools through specialised programmes (OECD, 2017a, 2017b, 2020a, 2020b; Sulis et al., 2020). That said, whatever the policy is, its effect in reducing inequalities could be properly assessed only in the long run period.

In any case, the analysis show that differences in academic achievements between native and immigrant students persist, although they are narrowing. These differences remain even after accounting for students' socioeconomic backgrounds and the characteristics of high schools that may affect students' performances. As expected, first-generation students are those that may benefit more from a more effective integration process by showing a persistent lag in literacy and numeracy performances

in almost all the countries considered. On the other hand, the performances gap between second-generation immigrants and natives is diminishing, particularly in reading performances. However, the widespread and persistent learning differentials among students with migration backgrounds calls for further investigation on the causes behind these gaps and on the set of policies that may effectively improve the integration of immigrant students into the society and the labour market of the destination countries.

To the best of our knowledge, these findings, even if they confirm results already discussed in the literature (see Sect. 2), have the merits that they have been reached (i) observing countries with different immigration histories, (ii) relying on information provided by multiple waves of an international large-scale assessment survey, and (iii) using a robust approach which combines mixture models with Rubin's rules for multiple imputation analysis to provide reliable estimates of countries' parameters by accounting for differences among high schools within the same country. As far as we know, this is the first time that the interaction terms between waves, countries and immigrant backgrounds are modelled to detect the trajectories of countries towards equity between students with different immigrant backgrounds and to discover changes in the observed gaps. Moreover, the approach provides a clear measure of country position in the road map of students' integration, that can be used to make any kind of direct comparison between and within countries and waves.

However, the analysis suffers from some significant limitations. First, we cannot assess the role of specific policies aimed at integrating immigrant students. Indeed, the lack of time-variant information on these policies prohibits such analysis. Second, due to the unavailability of information on students' backgrounds in terms of, for example, country of origin or reasons for migration, we do not control for compositional differences across countries and over time. Therefore, further investigations are needed to assess which policies can counteract education inequalities among students with different immigrant backgrounds and how the changes in the characteristics of students' populations affect countries' performances towards the integration of immigrant students.

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s41118-023-00199-5.

**Additional file 1.** Robustness checks and graphical evidences.

**Acknowledgements**
Not applicable.

**Author contributions**
Mariano Porcu contributed to the study conception and design. Material preparation, data collection, empirical analysis were performed by Cristian Usala. Isabella Sulis worked on the introduction, theoretical background, and results discussion. Francesca Giambona critically reviewed the manuscript. All authors read and approved the final manuscript and contributed to the conclusions.

**Funding**
This paper has been supported from Italian Ministerial Grant PRIN 2017 "From high school to job placement: micro-data life course analysis of university student mobility and its impact on the Italian North–South divide", n. 2017HBTK5P—CUP B78D19000180001.

**Availability of data and materials**
Data are available at PISA-OECD data website https://www.oecd.org/pisa/data/.

## Declarations

**Competing interests**
The author declare that there are no competing interests.

Received: 8 November 2022 Accepted: 6 July 2023
Published online: 17 August 2023

**References**

Agasisti, T., Falzetti, P., & Soncin, M. (2020). School principals' managerial behaviours and students' achievement: An empirical analysis of Italian middle schools. *International Journal of Educational Management ahead-of-print*.
Azzolini, D., & Barone, C. (2013). Do they progress or do they lag behind? Educational attainment of immigrants' children in Italy: The role played by generational status, country of origin and social class. *Research in Social Stratification and Mobility, 31*, 82–96.
Azzolini, D., Schnell, P., & Palmer, J. (2012). Educational achievement gaps between immigrants and native students in two "New Immigration Countries": Italy and Spain in comparison. *The Annals of the American Academy of Political and Social Science, 643*(1), 46–77.
Bertoletti, A., Cannistrà, M., Diaz Lema, M., Masci, C., Mergoni, A., Rossi, L., & Soncin, M. (2023). The determinants of mathematics achievement: a gender perspective using multilevel random forest. *Economies, 11*(32), 2347–2368.
Chiswick, B., & DebBurman, N. (2004). Educational attainment: Analysis by immigrant generation. *Economics of Education Review, 23*(4), 361–379.
Contini, D., & Azzolini, D. (2016). Performance and decisions: Immigrant-native gaps in educational transitions in Italy. *Journal of Applied Statistics, 43*(1), 98–114.
Gabrielli, G., & Impicciatore, R. (2022). Breaking down the barriers: Educational paths, labour market outcomes and well-being of children of immigrants. *Journal of Ethnic and Migration Studies, 48*(10), 2305–2323.
Gabrielli, G., Longobardi, S., & Strozza, S. (2022). The academic resilience of native and immigrant-origin students in selected European countries. *Journal of Ethnic and Migration Studies, 48*(10), 2347–2368.
Ganzeboom, H. B., & Treiman, D. J. (2003). *Three Internationally Standardised Measures for Comparative Research on Occupational Status. Advances in Cross-National Comparison*. Springer.
Gonzalez, E., & Rutkowski, L. (2010). Principles of multiple matrix booklet designs and parameter recovery in large-scale assessments. *IEA-ETS Research Institute Monograph, 3*, 125–156.
Grilli, L., & Rampichini, C., et al. (2009). 'Multilevel models for the evaluation of educational institutions: a review'. en. Statistical Methods for the Evaluation of Educational Services and Quality of Products. In P. Monari (Ed.), *Series Title: Contributions to Statistics* (pp. 61–80). Physica-Verlag HD.
Hartung, A. (2015). Unequal Attainments: Ethnic Educational Inequalities in TenWestern Countries. *European Sociological Review, 31*(6), 798–801.
Heath, A., & Brinbaum, Y. (2007). Guest editorial: Explaining ethnic inequalities in educational attainment. *Ethnicities, 7*(3), 291–304.
Hillmert, S. (2013). Links between immigration and social inequality in education: A comparison among five European countries. *Research in Social Stratification and Mobility, 32*, 7–23.
Lüdemann, E., & Schwerdt, G. (2013). Migration background and educational tracking. *Journal of Population Economy, 26*, 455–481.
Masci, C., De Witte, K., & Agasisti, T. (2018). The influence of school size, principal characteristics and school management practices on educational performance: An efficiency analysis of Italian students attending middle schools. *Socio-Economic Planning Sciences, 61*, 52–69.
Masci, C., Ieva, F., et al. (2016). Does class matter more than school? Evidence from a multilevel statistical analysis on Italian junior secondary school students. *Socio-Economic Planning Sciences, 54*, 47–57.
Monseur, C., & Adams, R. (2009). Plausible values: how to deal with their limitations. *Journal of Applied Measurement, 10*(3), 320–334.
Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica, 46*(1), 69.
OECD. (2005). PISA 2003 Technical Report. Tech. rep. OECD.
OECD. (2010). *PISA 2009 Results* (Vol. 2, p. 4). OECD. Chap.
OECD. (2011). *How are school systems adapting to increasing number of immigrant students? PISA in Focus 11*. OECD.
OECD. (2012a). PISA 2009 Technical Report. Tech. rep. OECD.
OECD. (2012b). *Untapped Skills: Realising the Potential of Immigrant Students*. OECD.
OECD. (2013a). *PISA 2012 Results* (Vol. 1, p. A2). Annex: Paris. Chap.
OECD. (2013b). *PISA 2012 Results* (Vol. 2, p. 3). OECD. Chap.
OECD. (2014). PISA 2012 Technical Report. Tech. rep. OECD.
OECD. (2015a). *Can the performance gap between immigrant and non-immigrant students be closed? PISA in Focus 53*. OECD.
OECD. (2015b). Education policy outlook. United Kingdom. Tech. rep. OECD.
OECD. (2016). *PISA 2015 Results* (Vol. 1, p. 7). OECD. Chap.
OECD. (2017a). Education policy outlook. Italy. Tech. rep. OECD.
OECD. (2017b). Education policy outlook. Spain. Tech. rep. OECD.
OECD. (2017c). PISA 2015 Technical Report. Tech. rep. OECD.
OECD. (2019). *PISA 2018 Results* (Vol. 2, p. 9). OECD. Chap.
OECD. (2020a). Education policy outlook. France. Tech. rep. OECD.
OECD. (2020b). Education policy outlook. Germany. Tech. rep. OECD.
OECD. (2020c). PISA 2018 Technical Report. Tech. rep. OECD.

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and Longitudinal Modeling Using Stata* (3rd ed.). STATA Press.
Rothon, C., Heath, A., & Lessard-Phillips, L. (2009). 'The Educational Attainments of the "Second Generation": A Comparative Study of Britain, Canada, and the United States'. en. Teachers College Record: The Voice of Scholarship in Education 111(6), 1404–1443.
Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. Wiley.
Schnell, P., & Azzolini, D. (2015). The academic achievement of immigrant youths in new destination countries: Evidence from southern Europe. *Migration Studies, 3*(2), 217–240.
Schnepf, S. V. (2004). *'How Different Are Immigrants?* (p. 1398). IZA Discussion Paper Series: A Cross-Country and Cross-Survey Analysis of Educational Achievement'.
Schnepf, S. V. (2008). 'Inequality of Learning amongst Immigrant Children in Industrialized Countries'. IZA Discussion Paper Series 3337.
Sulis, I., Giambona, F., & Porcu, M. (2020). Adjusted indicators of quality and equity for monitoring the education systems over time Insights on EU15 countries from PISA surveys. *Socio-Economic Planning Sciences,69*, 100714.
Sulis, I., & Porcu, M. (2015). Assessing Divergences in Mathematics and Reading Achievement in Italian Primary Schools: A proposal of adjusted indicators of school effectiveness. *Social Indicators Research, 122*(2), 607–634.
Triventi, M., Vlach, E., & Pini, E. (2022). Understanding why immigrant children underperform: Evidence from Italian compulsory education. *Journal of Ethnic and Migration Studies, 48*(10), 2324–2346.
Vaquera, E., & Kao, G. (2012). Educational Achievement of Immigrant Adolescents in Spain: Do Gender and Region of Origin Matter? *Child Development, 83*(5), 1560–1576.
Veldkamp, B. P., & Sluijter, C. (2019). Theoretical and practical Advances in Computer-based Educational Measurement, pp. 271–289.
Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.
Zinovyeva, N., Felgueroso, F., & Vazquez, P. (2014). Immigration and student achievement in Spain: Evidence from PISA. *Journal of the Spanish Economic Association, 5*, 25–60.

## Publisher's Note