International Journal of Educational Development 86 (2021) 102466

Contents lists available at ScienceDirect

International Journal of Educational Development

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

Where do you come from, where do you go? Assessing skills gaps and labour market outcomes for young adults with different immigration backgrounds

Alison Cathles [a], Dongshu Ou [b], Simone Sasso [a], Mary Setrana [c], Tom van Veen [d,*]

[a] UNU/MERIT Maastricht University, the Netherlands
[b] Department of Educational Administration and Policy, Faculty of Education, the Chinese University of Hong Kong, Hong Kong
[c] Center for Migration Studies, University of Ghana, Ghana
[d] School of Business and Economics, Maastricht University, the Netherlands

ARTICLE INFO

JEL classification:
I24
I25

Keywords:
School quality and education performance
Migration and education

ABSTRACT

Using PISA and PIAAC data from twelve OECD countries, we examine the gap in cognitive skills among immigrants and natives and evaluate how those differences have evolved over time. We also consider how demographics, family background and school quality explain the development of cognitive skills of young people with immigrant backgrounds. The results show, first, that some convergence in skills occurs between second-generation immigrants and natives over time. Second, demographics, family background and school quality variables all contribute to the achievement gaps across different groups.

1. Introduction

The economic and social integration of immigrants and their children has garnered increasing attention in policy discussions. One critical input to immigrants' integration and success in European countries is access to education, as educational achievement serves as an important marker of how well immigrants adapt to opportunities that arise in a more mobile global society (OECD/European Union, 2015). In this context, large and persistent achievement gaps between native and immigrant youth can indicate ineffective host country institutions aimed at reducing inequalities.

This paper aims to address three questions related to immigrant education. First, how do educational skills vary between natives and immigrants? Second, do these differences decrease as immigrants remain in the country over time? Third, to what extent do family background and school quality mitigate the differences?

In addressing these questions, we apply an empirical framework that includes demographic, family background and school quality variables to determine differences in cognitive skills. Using the Oaxaca decomposition technique, we explore which factors explain existing skills gaps. The paper makes two contributions to the literature. First, by combining PISA and PIAAC data, and by analysing gaps between natives and first- and second-generation immigrants, we investigate the evolution of migrant disparities in literacy and numeracy proficiency between the teenage years and young adulthood and their determinants in a sample of countries that took part in both skills' assessment surveys. Second, we investigate country group patterns to identify countries where policies or institutional arrangements contribute to narrowing the gaps.

2. Comparing the achievement of young people with and without immigrant backgrounds: literature review

While numerous studies have assessed skill gaps between natives and migrant groups, very few have examined the evolution of those gaps and the role that some key factors play in reducing or magnifying them over time. Andon et al. (2014) present a comprehensive overview of achievement gaps between natives and immigrants. They use various tests, taken at different ages and in various countries, to assess the existence of an achievement gap between natives and immigrants. They show the existence of positive achievement gaps between natives and immigrants, in reading, mathematics and science. They also demonstrate that the gaps tend to be lower among older pupils. Our study relates to this research on the assessment of achievement gaps. However, contrary to Andon et al. (2014), in our analysis, we explicitly distinguish between foreign born immigrants (first-generation immigrants) and native-born pupils with at least one foreign-born parent

* Corresponding author at: Maastricht University, School of Business and Economics, the Netherlands.
E-mail address: t.vanveen@maastrichtuniversity.nl (T. van Veen).

https://doi.org/10.1016/j.ijedudev.2021.102466
Received 14 February 2020; Received in revised form 9 July 2021; Accepted 13 July 2021
Available online 23 August 2021

(second-generation immigrants). Second, we combine data from comparable tests taken by pupils of different ages. This enables a more precise assessment of the development of achievement gaps over time. Third, whereas Andon et al. (2014) determine unconditional skills gaps, we explicitly aim to investigate the drivers of the achievement gaps, such as family background, school quality and country of destination. Our review of the literature therefore focuses on the dynamics of skills gaps and on these drivers.

2.1. Skills gaps over time

A cohort analysis for a group of pupils provides a precise estimate of skills gaps over time. However, a lack of appropriate data precludes this type of analysis. Instead, skills gaps are often measured using differences between various age groups or by comparing the gaps between first-generation and second-generation immigrants. The results are mixed. Azzolini et al. (2012) and Portes and Fernández-Kelly (2008) find that the achievement gap with native students is more sizable for first-generation immigrants. Duong et al. (2016) conduct a meta-analysis that investigates academic outcomes among immigrant youth, using studies limited to US immigrants. They conclude that while second-generation immigrants perform better than first-generation immigrants, third-generation immigrants perform worse than their second-generation counterparts. Dustmann et al. (2012), conversely, find that intergenerational mobility of knowledge is very low, and that the differences in educational outcomes between second-generation immigrants and natives are consistent with the differences among their parent's generations. Flisi et al. (2016) reach a similar conclusion: gaps in performance for both first- and second-generation immigrants persist even after controlling for family background characteristics. An alternative approach is to investigate whether the age of the immigrant at arrival matters for their educational performance. Smith et al. (2015) conclude that first-generation immigrants who arrive at secondary school age are at a disadvantage in comparison to those who arrived at a younger age. Flisi et al. (2016) also find evidence that first-generation immigrants who arrive before the age of 15 perform better than those who arrive after that age. Still another approach is to look at the scores of immigrant children on a specific test over a specific period. In a study based on the Netherlands, Huijnk and Andriessen (2016) investigate the performance of children with a non-Dutch, non-Western background. They conclude that over the period 1994–2014, the scores of these children on the final tests taken at elementary school have improved both for reading and math.

2.2. Skills acquisition and family background

Family background and its impact on student achievement is well established in the literature. Research using cross-national data from the 2000, 2003, and 2009 PISA assessments identify immigrant achievement gaps in literacy and numeracy ranging between 30 and 80 points, relative to the OECD average of 500 points (Azzolini et al., 2012; Levels et al., 2008; Marks, 2005; Sori et al., 2011). In these studies, the primary determinants of the immigrant achievement gap are family background variables, including socio-economic status, home-language,[1] parental education, and family structure. Fuchs and Wößmann (2007) find using PISA data that among a variety of factors, family background is significantly associated with math, science and reading achievement. In a comparative study among ten OECD countries with high immigration flows, Schnepf (2007) finds that family background and school segregation explain the low educational achievement of immigrant groups in continental Europe, whereas in the US and the UK, insufficient language skills are the main determinants of lower achievement scores among immigrants. Marks (2005), employing a comparative study of twenty countries, also concludes that immigrant students' performance is dependent on family and socio-economic factors, while socio-cultural and school factors play only a minor role. These results are partly confirmed by Dustmann et al. (2012) and Huijnk and Andriessen (2016), and those studies also find that the language spoken at home is a major determinant of differences in test scores between immigrants and natives. Flisi et al. (2016) assess skills differences between 15-year-old immigrants and their native counterparts. They conclude that family background, including parental education characteristics and the socio-economic index of the parents' occupational status, remains the main determinants of performance for different immigrant groups.

2.3. School quality and skills acquisition

Gaps in performance between natives and both first- and second-generation immigrants persist even after controlling for family background characteristics. Other variables such as school quality also influence educational achievements among different generations of migrants and their native counterparts (Fuchs and Wößmann, 2007; Woessmann, 2016). School quality is a broad and ambiguous term, which is reflected in the measurement of school quality in various studies.[2] Despite the clear rationale for the inclusion of school quality as a determinant of skills acquisition, empirical studies have not arrived at consistent results. Cobb-Clark et al. (2012) use PISA data and find a significant relationship between school quality variables (such as the percentage of private schools, quality of teacher education, and ability tracking) and migrant-native achievement gaps. However, other school quality variables do not correlate significantly with the achievement gap. In some studies, interaction terms between school quality variables help to draw a more complete picture. Fuchs and Wößmann (2007) find that student performance is higher with external exams, budget formulation and autonomy in textbook choice, and that autonomy in teacher hiring is also positively associated with student performance. Autonomy, furthermore, is more positively associated with performance in school systems that have external exit exams. Students perform better in privately operated schools, but private funding is not decisive. Woessmann (2016) finds that some school quality variables (related to school autonomy, in particular) correlate negatively with the achievement gap, but that the interaction between external exit exams and these autonomy variables results in a strong positive correlation. The quality of the school system may also have an indirect long-term effect. Lack of educational resources for immigrant parents can result in poorer performance among second-generation immigrants compared to their native counterparts (Schnell and Azzolini, 2015). Finally, incomes of immigrant parents tend to be lower, which can determine the educational resources they are able to afford at home.

2.4. Country groups and skills acquisition

A number of studies suggest that the relationship between skills acquisition and labour market performance among immigrants is country-dependent. These studies show that immigration laws and integration policies are key factors in the educational achievement of immigrants. Dustmann et al. (2012) confirm this result for a broader group of countries and conclude that traditional immigrant-receiving countries such as Australia, the US and the UK perform better in closing the skills gap. In a further refinement, the OECD has grouped its member countries into seven categories according to criteria such as

[1] Zinovyeva et al. (2014) investigate the role of language in educational achievement for students who migrate to Spain using PISA 2003−2009. They find initial gaps that decrease as the students continue living in Spain.

[2] Studies use numerous variables to measure school quality. These variables relate to school resources, school autonomy and school accountability. In Section 3, we elaborate on our measurement. Here we list results from previous research.

Table 1
Similarities and differences between PISA and PIAAC tests.

Similarities between PISA and PIAAC	Differences between PISA and PIAAC
Both are designed to have large stratified random and nationally representative samples.	PISA and PIAAC use different testing methodologies in different contexts. PISA only assesses 15-year-olds enrolled in secondary school and the test is taken in a class. PIAAC assesses adults ages 16–65 in their homes. The timeframe of the PIAAC test is shorter; it is approximately 60 min (30 min for a survey about background characteristics and 30 min for the assessment).
Both comply with accepted standards of sampling and international tests.	The scales are different. PISA's scale is from 0 to 1000 (or, 200–800). PIAAC's scale is from 0 to 500. The continuous scales can be standardized and therefore can be compared.
Both are designed to be internationally comparable	The PISA and PIAAC reading and numeracy scores are provided in the data as plausible values, PISA uses 5 plausible values and PIAAC uses 10 plausible values.
Both (a) directly assess cognitive skills in literacy and numeracy and (b) collect demographic information and other variables of interest as well.	PISA tests mathematical literacy. PIAAC tests numeracy. Mathematical literacy tests mathematical knowledge that fits in math programmes of secondary schools. Numeracy is a broader concept and includes real world applications of mathematical concepts.
Despite differences between these, they are both based on a similar conceptual framework Testing constructs are conceptually similar. That means both assessments are based on constructs that are testing for 'real world' applications of cognitive skills.	

Source: Gal and Tout (2014)

immigrants' population size, length of residence, age, educational level, language predominant entry categories, and share coming from high-income countries (OECD/European Union, 2015). Among these criteria, education represents an important determinant of immigrant integration into the labour market (OECD/European Union, 2015).

3. Matching PISA and PIAAC data and measuring school quality

To analyse skills gaps over time, we connect data from the OECD's Programme for International Students Assessment (PISA) and Programme for International Assessment of Adult Competencies (PIAAC). Both programmes test cognitive skills in literacy and numeracy; PISA tests individuals at age 15, and PIAAC tests the skills of adults. These data enable us to explain whether test outcomes differ systematically for young people with immigrant backgrounds and how any differences in cognitive skills found at age 15 shift as cohorts enter adulthood (research questions 1 and 2). Both databases also contain information on family background, and PISA has data on school quality (question 3); we are also able to examine whether the destination country matters for the results. In this section, we elaborate on the matching of PISA and PIAAC data and on our measurement of school quality from these data.

3.1. Matching PISA and PIACC data

To investigate the dynamics of skills gaps, we connect the PISA data to the PIAAC results. The connection between the PISA and PIAAC tests is not a one-to-one relationship: the adults tested by PIAAC are not necessarily the same pupils tested by PISA, which limits our ability to conduct a real cohort analysis. In addition, there are differences in the set-up of the tests, measurement models and item delivery. Gal and Tout (2014) have extensively investigated similarities and differences between mathematical literacy (PISA) and numeracy (PIAAC) tests.

Table 2
Mapping of PISA and PIAAC sample.

PISA taken in…. by 15 year old pupils	Age in PIAAC 2011–12	Age in PIAAC 2014–15 (Greece only)
2000	26–28	29–31
2003	23–25	26–28

Note: Greece is the only country in our sample to implement PIAAC in 2014/2015. The other 13 countries in the sample implemented PIAAC in 2011/2012.

Table 1 gives an overview of the similarities and the differences between the PISA and the PIAAC tests. Gal and Tout (2014) extensively reflect on these differences in particular when the results are used for comparison purposes. They conclude that it is possible to compare average results from PISA and PIAAC tests, but comparisons must be cautious and they recommend using age cohorts of people in PIAAC who can be matched with cohorts of students who participated in earlier PISA assessment. That is exactly our approach.

Following the recommendations of Gal and Tout (2014), we have to 'match' the cohorts from the PISA data with the cohorts from the PIAAC data in our sample. For this, we follow the decision rules described in Table 2. For example, the PISA test given in 2000 to 15-year-old pupils, can be 'matched' with the nationally representative data from PIAAC test administered in the same country roughly 11 years later (in 2011/2012) with people who range in age from 26–28. We select three possible ages because, depending on the exact date of birth (including month) and the exact date of PIAAC assessment (including month), it is feasible that the cohort in the range of ages 26–28 corresponds with the cohort tested by PISA in 2000. A similar matching methodology for the same two datasets has recently been used by the OECD in an analysis comparing skills of teenagers and young adults (Borgonovi et al., 2017).

Our variables of interest for measuring achievement include cognitive skills in reading and math (PISA), cognitive skills in literacy and numeracy (PIAAC), and tertiary education attainment (PIAAC). Both assessments use a standardized methodology to facilitate international comparisons.[3] To investigate how skills gaps among natives, first-generation immigrants and second-generation immigrants have changed over time, we compare the reading and math test scores of 15-year-olds from PISA in the 2000 and 2003 waves with the literacy and numeracy skill scores of the adults who participated in PIAAC in 2011–12 and 2014–15 and who belonged to the same birth cohort as the PISA test takers.[4] We thus select the PIAAC test takers in the age range of 23–28 years. To ensure that our samples are as comparable as

[3] In PIAAC, to estimate the competency score, a set of ten plausible values for individual respondents is derived from the skill assessment and a background questionnaire. These values can be used to generate a competency distribution rather than an individual score. In fact, no single proficiency score (neither the first plausible value nor the average of all ten plausible values) can be assigned to a specific respondent, but rather to a group. In our econometric models, we follow the literature (e.g., Hanushek et al., 2015; Allen et al., 2013) and use the first plausible values. For more information regarding the PIAAC methodology and the use of plausible values, see OECD (2013) Technical Report of the Survey of Adult Skills (PIAAC); von Davier et al. (2009); Perry et al. (2014) and Jacob and Rohstein (2016). Similarly, for PISA, "using one plausible value or five plausible values does not really make a substantial difference on large samples" (OECD, 2009, p. 46) and a significant number of studies using the first plausible value or a single plausible value rather than using all five proved to have no impact on key research findings (Jerrim et al., 2017). The only drawback of using one plausible value as dependent variable is the imputation error (Jerrim et al., 2017).

[4] PIAAC data are collected via surveys of people aged 16–65 in 33 countries. The PIAAC surveys were conducted in 2 rounds. The first round was collected in 2011 and 2012 in 24 countries. The second round was collected in 9 additional countries in 2014 and 2015. All countries in our sample, except Greece, participated in the first round of the survey. (see: http://www.oecd.org/skills/piaac/about/)

possible, we include in the PIAAC sample only those first-generation immigrants who migrated to the destination country before the age of 15. Second-generation immigrants and natives were, by definition, educated in the destination country's school system. This means that all three groups spent at least some time in the school system of the destination country.[5]

In total, 106,090 students sat for the PISA test in the sample we use: 92,307 natives, 4,312 first-generation immigrants and 9,463 second-generation immigrants. For the PIAAC data, our sample includes 7,441 adults who participated in the cognitive skills assessment: 5,614 natives, 222 first-generation immigrants and 525 second-generation immigrants.[6] The PIAAC data includes fewer observations because the age range is restricted.

In order to examine patterns in country-level policies and institutions that may contribute to narrowing the achievement gaps, we divide the fourteen countries for which we have common pertinent variables from PISA and PIAAC[7] into four country groups based on immigrant population characteristics.[8] We draw on the OECD/European Union (2015) for these classifications, though we make some modifications.[9] We consolidate the country groupings as follows:

1 *Longstanding destinations*: United Kingdom, Belgium, France and Netherlands;
2 *Significant recent migration and humanitarian countries*: Denmark, Finland, Sweden and Norway;
3 *New destinations*: Ireland, Greece, Italy and Spain;
4 *Eastern European destinations*: Czech Republic and Russian Federation.

[5] Immigrants can be defined in numerous ways. Consistent with the US Census Bureau, we define foreign-born pupils with at least one foreign-born parent and who came to the destination country before the age of 15 and who spent at least some time in secondary school in the destination country as first-generation immigrants. Native-born pupils with at least one foreign-born parent are defined as second-generation immigrants. Native-born pupils with no immigrant parents are referred to as natives; this category represents our reference group. These definitions have also been used by organizations such as the Dutch Bureau of Statistics.

[6] We checked for the robustness of our results if we would use an alternative definition of the 2nd generation immigrant group, e.g. the one that Flisi uses. If we adopt the Flisi definition for immigrant status, the overarching patterns for literacy remains the same, but the gaps between second and first-generation immigrants are smaller across the board and some patterns change slightly for numeracy. For example, the standardized gap between first and second-generation immigrants is smaller in young-adulthood (PIAAC) than in high school (PISA). We also checked the analytical results and while qualitatively similar, we consider them less reliable, particularly for the Oaxaca decomposition.

[7] Austria, Canada, Germany, New Zealand and the United States do not report information on respondents' age in the PIAAC PUF (Public Use Files) data. Chile, Cyprus, Estonia, Israel, Lithuania, Singapore, Slovak Republic, Slovenia and Turkey did not participate in the PISA waves (2000 and 2003) we are analyzing. For Indonesia, no PIAAC information is publicly available. Japan, Korea, and Poland do not report information on the PISA respondents' country of origin or include only very few individuals who are first- or second-generation immigrants.

[8] These include language, predominant entry categories, length of residence, share with education, share coming from high/low income countries of origin, and size of the immigrant population.

[9] Considering that our sample includes only 14 of the 41 countries in the original classification and that we do not have any countries for some country groups, we merged some of the groups proposed in the OECD/European Union (2015) study. In our classification, we have distinguished longstanding destinations from new destinations. The OECD/European Union (2015) study further distinguishes country groups based on the percentage of immigrants that is highly educated. We were unable to do this, as our sample of countries is much smaller. However, we control for individual-level differences in parental education in both the PISA and PIAAC analyses.

3.2. Measuring school quality

As we discussed in Section 2, studies rely on numerous variables to measure school quality. We follow convention by including three groups of school quality variables: school resources, school autonomy and school accountability (Hanushek and Woessmann, 2007; Hanushek et al., 2013).[10]

School resources variables:

In the first set of school quality variables, we include measures for privatization of schools, teacher quality and school size. The school principals were asked "about what percentage of your total funding for a typical school year comes from government", and we use this continuous variable to reflect the degree of privatization of schools, with the assumption that private schools perform better (Meroni et al., 2015). For teacher quality, we include two dummy variables which directly relate to students' learning in literacy and math skills: shortage/inadequacy of language/math teachers (equal to 1 if the school reported "Not at all", i.e. no shortage of teachers in language or math subject). We use no shortage in language teachers in the regressions for reading and no shortage in math teachers in regressions for math. We also include the percentage of full-time equivalent teachers who have a higher degree qualification in pedagogy; we presume that the higher the percentage, the higher the quality of teaching and the higher school quality. School size is measured by the total number of students (in log form).

School autonomy variables:

The second set of school quality variables includes schools' autonomy in decision-making. Researchers have found that school autonomy has important implications for students' academic performance (Hanushek and Woessmann, 2007; Woessmann et al., 2007). The extent of autonomy reflects the degree to which stakeholders in the schooling system have incentives to improve student outcomes, including addressing potential academic performance differences among immigrant and native students (Woessmann, 2016). Studies have shown positive effects from school autonomy in budgeting and staffing in terms of improving student achievement in developed and high-performing countries (Hanushek et al., 2013; Hanushek and Woessmann, 2007). However, evidence also suggests that while the effect of school autonomy is positive in areas with informational advantages at the local level, school autonomy can have negative consequences in areas where local rent-seeking activities are common (Fuchs and Wößmann, 2007; Wößmann, 2003). However, generally speaking, the ability of schools to recruit their own teachers and to formulate their budgets helps to ensure the right teacher coverage and an appropriate allocation of existing limited resources that together serve the schools' needs, for example, by providing teachers with training and support to increase their competencies in working in diverse and multicultural classrooms (European Commission/EACEA/Eurydice, 2019).

In 2000 and 2003, school principals were asked, "In your school, who has the main responsibility for …" Principals needed to tick as many boxes as appropriate to indicate whether a particular item (for example, hiring teachers) is "Not a main responsibility of the school", or "the school's governing board", "principal", "department head" or "teachers". Provided that the principals did not report that hiring teachers is "not a school responsibility", we create a dummy variable coded 1 to indicate that the school has autonomy of, for example, hiring teachers. We include three school-based decision-making types following the existing literature (Wößmann et al., 2007):[11] 1)

[10] In the comprehensive set of school quality variables, there are 18 variables, which is quantitatively demanding. Based on the literature described above, we zeroed in on 10 school quality variables.

[11] As discussed in Wößmann et al. (2007), there are also survey items on autonomy in firing teachers. However, these measures are highly collinear with autonomy in hiring teachers. Therefore, we use just one autonomy variable to capture decisions in staff hiring/firing.

Table 3
Descriptive statistics based on PISA (2000, 2003) and PIAAC (2011/12, 2014/15).

	PISA			PIAAC		
	Natives	1 st Gen.	2nd Gen.	Natives	1 st Gen.	2nd Gen.
Test scores Mean Reading/Literacy score	510.39	462.57	498.88	286.95	262.55	284.26
	(93.76)	(106.74)	(99.59)	(44.38)	(49.23)	(46.88)
Mean Math/Numeracy score	513.23	469.78	502.73	282.48	259.33	276.09
	(94.37)	(101.85)	(96.80)	(48.43)	(51.01)	(52.51)
Demographic variables						
Female (%)	48.82	51.46	48.28	52.28	53.15	54.86
Test language same as the language at home (%)	90.40	56.61	80.21	98.68	70.27	91.62
Family variables						
Parent's education:						
Uncompleted secondary	24.94 %	19.04 %	23.53 %	22 %	30 %	25 %
Secondary	31.26 %	26.65 %	26.53 %	46 %	25 %	33 %
Tertiary	43.80 %	54.31 %	49.93 %	32 %	45 %	42 %
Books at home:						
10 and below	14.93 %	29.01 %	18.08 %	10 %	26 %	13 %
11−100	49.34%	45.18 %	46.63 %	43 %	42 %	42 %
101−500	31.91 %	22.96 %	30.75 %	37 %	25 %	35 %
more than 500	3.83%	2.85 %	4.53 %	9%	7%	10 %
ESCS (SES) - PISA Only	0.08	−0.08	0.08			
	(0.91)	(1.00)	(0.97)			
School resources						
% of government funding	89.90 %	90.97 %	90.35 %	91 %	93 %	91 %
No shortage of language teachers*	64.01 %	57.63 %	58.09	66%	63 %	64 %
No shortage of math teachers*	58.99 %	50.97 %	51.29 %	61 %	59 %	58 %
% of full-time teachers with a higher degree in pedagogy	46.80 %	53.93 %	54.40 %	53 %	63 %	60 %
School size (natural log)	6.24	6.23	6.35	6.16	6.07	6.27
	(0.66)	(0.65)	(0.65)	(0.41)	(0.4)	(0.38)
School autonomy						
Hiring teachers	66.63 %	70.85 %	77.34 %	70 %	77 %	74 %
Formulating budget	71.43 %	71.20 %	72.41 %	78 %	82 %	78 %
Determining course content	71.82 %	67.53 %	70.65 %	71 %	67 %	68 %
School accountability						
Assessment is used to:						
Inform parents of child's progress	97.63 %	97.43 %	97.31 %	95 %	90 %	94 %
Decide grade Retention/promotion	76.65 %	74.84 %	74.59 %	72 %	52 %	68 %
Group students	44.36 %	47.66 %	50.38 %	44 %	41 %	51 %
Number of observations	92,307	4,312	9,463	5,614	222	525

Note: Source: PISA 2000 and 2003; PIAAC 2011 and 2014. Standard deviations are presented in parentheses. PISA scores range from 0 to 1,000, whereas PIAAC scores range from 0 to 500; ESCS (SES) stands for student socio-economic background and is only available in the PISA dataset. Data about whether the language of the test is spoken at home are not available for the Russian Federation for either survey year, therefore these country-year combinations are not included in the descriptive statistics, nor the analysis.

determining course content; 2) hiring teachers; and 3) formulating the school budget. We thus rely on three dummy variables to measure school autonomy.

School accountability variables:

The third set of school quality variables measures school accountability. Holding schools accountable, which includes informing parents about student performance, creates an incentive for schools to close the educational performance gap for immigrant students (OECD, 2010a). Of course, the effect of using assessments to make decisions about the retention or promotion of students may be two-sided: on one hand, students may study harder if they fear the consequences of failing and not being promoted. On the other hand, an increased threat of grade retention could have negative effects on students, especially by undermining the motivation of academically disadvantaged students (Battistin and Schizzerotto, 2019). Furthermore, grouping students by ability for instructional purposes could increase teaching effectiveness by reducing variation within classrooms and thus better accommodating students' academic needs at a proper instructional pace (OECD, 2010b). However, evidence also suggests that academic tracking and ability grouping could limit opportunities for students—particularly immigrant students—to learn and fully develop their educational potential (Schofield, 2010).

In the 2000 and 2003 data sets, six common items address school accountability (i.e. "In your school, are assessments of 15-year-old students used for any of the following purposes"):

a) to inform parents about their child's progress
b) to make a decision about retention or promotion
c) to group students for instructional purposes
d) to compare the school to < district or national > performance
e) to monitor the school's progress from year to year
f) to make judgments about teachers' effectiveness

We created dummy variables for the first three items[12], coded 1 if the principals reported "Yes" on that particular statement.

After merging the school-level variables with the data for the students who took the PISA test with the appropriate weights, these factors can be used at the student level like any other student attribute (OECD, 2009). For the PIAAC data, country means for each school quality variable are calculated separately, again using the appropriate weights, and merged with the PIAAC participant data according to the mapping in Table 2. We cannot map the individual school-level information to the individuals in the PIAAC sample, as we follow an approach adopted by

[12] We did not use all six common items after checking for multicollinearity; we reduced the dimensions of school accountability by focusing on the first three. We also conducted analyses using only b, c, and d as in Wößmann et al. (2007), and the results are qualitatively the same. In addition, we checked for multicollinearity in these predictors, using Variance Inflation Factor (VIF) values. The VIF values are all lower than 4 which suggests low/moderate multicollinearity.

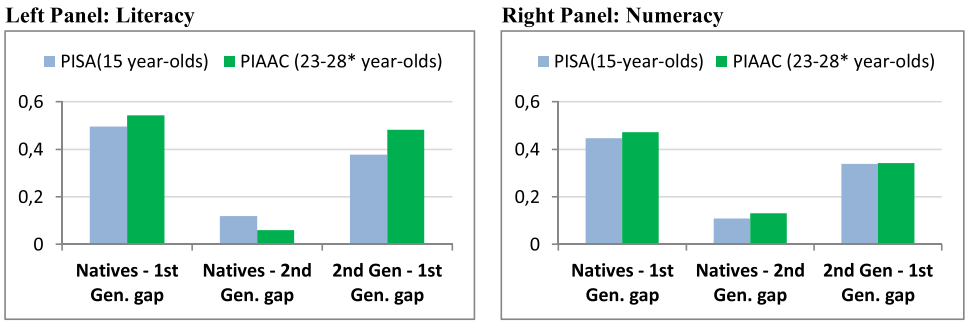

Fig. 1. Standardized skill-gaps between natives, first- and second-generation migrants.
Notes: We follow Borgonovi et al. (2017) in calculating 'standardized skill gaps'. The gap refers to the difference in means for each immigrant group divided by the standard deviation for the entire sample. A pairwise means test shows that for PIAAC, gaps between groups are statistically significant, except for the literacy gap between natives and second-generation immigrants.

Meroni et al. (2015) that generates a country-level variable for various school quality measures. Adding the country-level school characteristics can control for potential differences in school quality at the national-level that may be attributed to observed variation in PIAAC skill scores. It is worth noting that the interpretation of the school quality variables differs across the PISA and PIAAC data: whilst the coefficients for the school variables in the PISA data reveal the association of school characteristics and individual performance, the school quality coefficients in the PIAAC data reflect the average association between several dimensions of national school quality and individuals' literacy and numeracy skills.

4. Descriptive statistics

Table 3 presents the descriptive statistics for the variables we employ in this study.[13] We report the average scores of the PISA and the PIAAC tests, and averages of the control variables that we use in our empirical research, including school quality variables. We have also run these descriptive statistics by country. These results are available from the authors on request.

Table 3 confirms the results of other research: cognitive skills assessment scores differ across the groups of young people with and without an immigrant background. We want to stress three points. First, first- generation immigrants under-perform in comparison to *both* second-generation immigrants and natives.

Second, Table 3 also reveals that in our sample, first and second-generation immigrants are coming from different socio-economic backgrounds than their native peer group. First-generation immigrants' parents tend to have completed more tertiary education in comparison to natives' parents. This might be due to positive selection in immigration policies in most countries. However, these immigrant families tend to have fewer books at home. The overall Socio Economic Status (SES) composite score in PISA for the first-generation immigrants is negative, which indicates that they have lower SES than the other two groups. It is worth noting that second-generation immigrants tend to have more highly-educated parents than natives and tend to have more books at home. It is thus important to control for these variables to disentangle whether the differences in testing scores are driven by socio-economic backgrounds[14] as opposed to immigrant backgrounds. The first-generation immigrants tend to have a slightly higher percentage of females in the PISA sample, but the gender distribution is comparable across the three groups in PIAAC. Given that, contrary to first-generation immigrants, second-generation immigrants are born in the test country, and whether the test language is the language spoken most often at home is, as expected, much lower for first-generation immigrants than for second-generation immigrants, which is again lower than for natives. The third point worth stressing is that, in terms of the school resources variables, natives tend to be in schools with no shortage of either language or math teachers even though their schools have a relatively lower percentage of teachers with higher degrees in pedagogy.[15] Second generation immigrants tend to be in schools with a larger student population in the PISA sample. The percentage of government funding for school resources is the highest for the first-generation immigrants. Second-generation immigrants tend to be in schools that have greater autonomy in hiring teachers, and first-generation immigrants tend to be in schools with the least autonomy in deciding course content.[16] It is also interesting to see that for native students, even though their schools receive a relatively lower percentage of government funding, their schools also tend to have a lower degree of autonomy in teacher hiring and are more likely to use assessments to inform parents of child's progress and determine grade retention, but less likely to practice ability grouping using assessments.[17]

5. Dynamics of skills gaps

To address question 1 and 2, we present the data from Table 3 in a slightly different manner to elaborate on the dynamics of skills gaps. Fig. 1 illustrates the standardized skills gaps between young people with different immigrant backgrounds. The left-hand panel compares the standardized PISA score literacy gaps of 15-year-olds (light blue bars) with the standardized PIAAC score literacy gaps (green bars). Reading from left to right, the first gap comparison is between natives and first-generation immigrants, the second between natives and second-generation immigrants and the third between second-generation and first-generation immigrants. The right-hand panel replicates this set-up

[13] Note that PISA collects information on the socio-economic background of the students, but we cannot compare that information with information from the PIAAC data. Therefore, we do not use this information in our set-up.

[14] PISA has an index of economic, social and cultural status (ESCS) that we include in our analysis of PISA. Previous studies have shown that this indicator has a strong correlation with student academic achievements. Even though we did not report the statistics of this variable in most of our tables, our findings are consistent with previous research that ESCS is a statistically significant predictor of skill scores. Results are available upon request. However, PIAAC does not include a similar measure of socio-economic status. It is also worth noting that we consider age effects in skill formation, controlling for month and year of birth in our analysis using PISA and age in our analysis using PIAAC.

[15] For the set of school resources variables, the t-statistics (not reported) show that the first-generation and the second-generation immigrants have no difference in terms of their shortage of language/math teachers and percentage of full-time teachers with a higher degree in pedagogy. School size is no different for natives and first-generation immigrants. All others are statistically different between any two groups.

[16] For the set of school autonomy variables, all three groups are statistically no different in school autonomy in "formulating budget".

[17] For the set of school accountability variables, even though first-generation and second-generation immigrants are not statistically different in terms of using assessment to inform parents of child's progress, and natives and first-generation immigrants are not statistically different, either, but there is a marginal significance for the mean differences of natives and second-generation immigrants (t-test statistics is significant at 10%). All other group means are significantly different statistically.

for numeracy skills. According to the literature, differences in standardized gaps are considered small if the difference is less than 0.3, medium if the difference is between 0.3 and 0.5, and large if it is greater than 0.5.

Fig. 1 shows first that the literacy gap for the first-generation immigrants widens over time. Second, the difference in literacy between natives and second-generation immigrants narrows and disappears almost completely by early adulthood. Indeed, a pairwise means test for the gaps in PIAAC literacy proficiencies shows no statistically significant difference between natives and second-generation immigrants. This finding suggests convergence in literacy between second-generation immigrants and natives over time. Third and consistent with the first two findings, the gap in literacy skills between the first-generation and the other two groups has increased over time.

In the right-hand panel, we observe similar patterns for numeracy, though the size of the gaps differs. For PISA, the standardized numeracy gaps between natives and first-generation immigrants is slightly lower than for literacy (just shy of 0.5) and widens (to just over 0.5) in early adulthood, but the widening is less dramatic than for literacy. The change in the standardized gaps between natives and second-generation groups is also less dramatic for numeracy: the gaps in PISA and PIAAC are both statistically significant with similar magnitudes, implying that the numeracy skill gap between natives and second-generation immigrants remains over time.

Further, in line with Andon et al. (2014), Fig. 1 reveals that the gaps for numeracy skills and for literacy skills appear to be similar between natives and first-generation immigrants. However, numeracy skills gaps are lower (closer to 0.4 than to 0.5). Similarly, the standardized literacy skill gaps between first- and second-generation immigrants is much bigger (hovering around 0.4) compared to the corresponding gaps in numeracy skills (closer to 0.3).

In short, gaps in skill levels between natives and second-generation immigrants are relatively small, and the gaps do not appear to be increasing over time. Natives have the highest level of skills, while first-generation immigrants tend to have the lowest level of skills among all three groups. First-generation immigrants are being left behind in secondary schools, and they do not appear to catch up in young adulthood. We further explore the skill gaps among the three groups controlling for demographics, family and school quality variables, as described in the next section.

6. Determinants of skills gaps: empirical framework

To address the remaining research question, we develop an empirical model based on an education production function framework, as presented in Eq. (1). The model is similar to other studies that use international assessments to analyse the underlying determinants of cognitive skills (e.g., Hanushek and Woessmann, 2017).

$$T = \alpha_0 + \alpha_1 IM + \alpha_2 D + \alpha_3 FB + \alpha_4 SQ + \alpha_5 A + \alpha_6 CG + \varepsilon \quad (1)$$

where T is the outcome of the process of educational production measured by test scores in Literacy and Numeracy; IM indicates the categorical group (natives, first- or second-generation immigrants); and D is a vector of personal traits listed under demographic variables in Table 3, that may impact cognitive skills. FB is a vector of Family Background characteristics; SQ represents a vector of measures for School Quality; and A represents individual ability, which remains unobservable in our analysis. CG indicates the country group based on the OECD/European Union (2015) classification described in Section 3. The error term ε captures unmeasured variables and the randomness of learning.

Further, to investigate how the school quality variables can have different effects across migrant groups, we modify Eq. (1) to allow more flexibility in the specification and to add on a set of interaction terms between immigrant status (IM) and school quality (SQ). Doing so may shed light on the potential mechanisms that explain how these variables widen or reduce the performance gaps.[18]

We use demographic and family variables to control for family background. The demographic variables include gender, month and year of birth (for PISA only) or age (for PIAAC only) and whether the test language is the same as the language spoken most often at home.[19] To proxy family background, we use two socio-economic variables that PISA and PIAAC have in common, parental education and the number of books in a household.[20] The number of books in the home can be a proxy for income, or for family 'culture' towards learning. We include a SES composite index in the PISA analysis as well.

The empirical set-up allows for two types of analyses. First, we analyse differences in educational performance and attribute these outcomes to family background and school quality (question 3). Second, we analyse whether the results of the previous analyses differ by country grouping.

To analyse differences in educational performance, we first estimate Eq. (1) using OLS. In the first specification, we include IM as a categorical variable with natives as the reference group. The results of this regression are presented in Table 4 and are discussed in the next section. Next, we employ a Blinder-Oaxaca decomposition to assess achievement gaps between groups of young people with and without immigrant backgrounds, or with different immigrant backgrounds, and to decompose these gaps into components of interest: family background, school quality, demographics, and country group. For this decomposition, we first re-estimate Eq. (1) for each of the categorical groups in IM (using OLS). Second, since the fitted line ($\overline{X}\widehat{\beta}$) passes through the means (T), we know that for each group it holds that:

$$\overline{T}_N = \overline{X}_N \widehat{\beta}_N \quad (2)$$

$$\overline{T}_{1stGen} = \overline{X}_{1stGen} \widehat{\beta}_{1stGen} \quad (3)$$

$$\overline{T}_{2ndGen} = \overline{X}_{2ndGen} \widehat{\beta}_{2ndGen} \quad (4)$$

where $\overline{T}$ and $\overline{X}$ are the means for each group, with the subscript N representing natives, the subscript 1st Gen representing first-generation immigrants, and the subscript 2nd Gen representing second-generation immigrants. X represents a vector of all independent variables used in the regression. In each equation, $\widehat{\beta}$ represents the estimated coefficients from the OLS regressions for the specific groups.[21] Third, from Eqs. (2) to (4) we can determine the achievement gaps by calculating the differences between the dependent variables in the equations, using the estimated coefficients to disentangle the differences. Using the difference between the native and first-generation immigrant groups as an example, to derive the achievement differentials between the two groups, we subtract Eq. (3) from Eq. (2). Rewriting, we can express this difference as follows:

$$\overline{T}_N - \overline{T}_{1stGen} = \left(\overline{X}_N - \overline{X}_{1stGen}\right)\widehat{\beta}_N + \overline{X}_{1stGen}(\widehat{\beta}_N - \widehat{\beta}_{1stGen}) \quad (5)$$

The first term on the right-hand side of Eq. (5) represents the difference in educational achievement due to variation in the independent variables. The second term on the right-hand side shows the difference in the marginal effects of the independent variables on educational achievement across the groups. We focus on the first term. We show the

[18] Note that the results of the School Quality variables may only be informative in the PISA data analysis, because PIAAC uses country average school quality variables that do not vary over native/immigrant groups.

[19] In PIAAC, this variable is derived and coded by the OECD (2013).

[20] PISA respondents were asked about the number of books in the household at age 15 and PIAAC respondents were asked about the number of books in the household at age 16.

[21] Note that these coefficients are not directly reported in Table 4.

Table 4
OLS Regression Family inputs, School inputs and Literacy/Numeracy Skill Scores.

	Panel A: PISA scores				Panel B: PIAAC scores			
	Literacy		Math		Literacy		Numeracy	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Immigration Group (Natives are the reference)								
1 st-generation immigrants	−19.57***	−19.62***	−19.91***	−19.60***	−19.48***	−16.23***	−20.04***	−17.86***
	(2.70)	(2.70)	(2.77)	(2.75)	(3.01)	(2.93)	(3.18)	(3.10)
2nd-generation immigrants	−3.83**	−3.63**	−6.66***	−6.15***	−3.41*	−2.80	−6.71***	−6.11***
	(1.78)	(1.75)	(1.82)	(1.79)	(1.91)	(1.87)	(2.14)	(2.09)
Demographic variables:								
Female	−30.94***	−30.41***	10.59***	10.78***	−0.61	−0.41	−11.02***	−10.59***
	(1.06)	(1.05)	(1.11)	(1.10)	(1.01)	(0.99)	(1.10)	(1.08)
Test language same as the language at home	20.23***	19.02***	11.19***	10.72***	12.29***	13.28***	5.87*	7.44**
	(1.79)	(1.78)	(1.98)	(1.98)	(3.25)	(3.27)	(3.54)	(3.51)
Family variables:								
Either parent has a higher education degree	−11.43***	−11.84***	−8.56***	−9.10***	14.99***	14.98***	16.03***	14.92***
	(1.41)	(1.40)	(1.48)	(1.46)	(1.45)	(1.42)	(1.62)	(1.58)
Home owned								
11−100 books	27.36***	25.00***	28.18***	25.98***	22.03***	22.43***	23.19***	24.74***
	(1.69)	(1.67)	(1.67)	(1.66)	(1.90)	(1.88)	(2.08)	(2.03)
101−500 books	51.33***	47.58***	54.89***	51.31***	40.70***	40.77***	42.46***	44.06***
	(1.91)	(1.89)	(1.95)	(1.94)	(1.99)	(1.97)	(2.17)	(2.13)
More than 500 books	52.65***	48.15***	58.21***	53.89***	44.43***	45.23***	47.30***	49.96***
	(3.03)	(2.98)	(3.17)	(3.13)	(2.49)	(2.49)	(2.74)	(2.72)
ESCS (SES) - PISA Only	32.14***	30.41***	29.33***					
	(0.81)	(0.82)	(0.84)					
School resources								
% of government funding in total school funding		−0.23***		−0.24***		0.21		0.24
		(0.02)		(0.03)		(0.19)		(0.21)
No shortage of language/math teachers[†]		7.18***		11.46***		14.80**		−12.13
		(1.07)		(1.12)		(6.42)		(7.61)
% of full-time teachers with higher degree in pedagogy		0.97*		−3.47**		−17.00		−12.06***
		(0.96)		(1.03)		(2.33)		(2.64)
School autonomy								
Hiring teachers		12.66***		11.36***		41.07***		41.96***
		(1.23)		(1.28)		(3.90)		(4.39)
Formulating budget		−6.66***		−10.20***		−50.76***		−26.24***
		(1.25)		(1.35)		(5.86)		(6.98)
Determining course content		0.22		0.70		−43.98**		−40.67****
		(1.23)		(1.24)		(4.03)		(4.71)
School accountability								
Assessment is used to:								
inform parents of child's progress		0.23		1.04		−34.39***		−29.95**
		(2.74)		(2.68)		(12.98)		(13.67)
Decide grade retention/promotion		10.00***		9.72***		40.43***		41.24***
		(1.41)		(1.58)		(3.61)		(4.05)
Group students		2.48**		1.42		7.08		−2.64
		(1.16)		(1.22)		(4.73)		(5.24)
Country groups								
Significant recent migration and humanitarian countries	−4.26***	15.13***	−14.31***	0.73	3.52***	21.22***	9.91***	33.57***
	(0.99)	(1.46)	(0.98)	(1.53)	(1.36)	(3.41)	(1.49)	(4.16)
New destinations	−19.70***	−10.19***	−44.92***	−41.03***	−21.12***	−15.28***	−19.99***	−12.15***
	(1.04)	(1.30)	(1.05)	(1.34)	(1.33)	(2.27)	(1.45)	(2.52)
Eastern European destinations	−55.75***	−55.58***	−42.22***	−44.30***	−10.52***	−16.60***	−2.34	−1.372
	(1.27)	(1.54)	(1.32)	(1.66)	(1.61)	(3.33)	(1.75)	(3.87)
R-squared	0.27	0.28	0.20	0.22	0.21	0.24	0.22	0.25
Number of observations	106,082	106,082	106,082	106,082	6,361	6,361	6,361	6,361

Note: Source: PISA 2000 and 2003; PIAAC 2011−12 and 2014−15. Dependent variables for Panel A are PISA literacy (Columns (1) and (2)) and math scores (Columns (3) and (4)). Dependent variables for Panel B are the first plausible value for PIAAC literacy (Columns (5) and (6)) and numeracy scores (Columns (7) and (8)). Each cell represents the coefficient of the corresponding variable estimated by Eq. (1). For details on the variables please refer to the note under Table 3. We also include SES, month and year of birth in regressions using PISA data and age in regressions using PIAAC, but statistics for these three variables are not reported. Robust standard errors are in parentheses. * significant at 10 %, ** significant at 5%, *** significant at 1%.

[†] We use "no shortage of language teachers" when we estimate Reading or Literacy scores, and we use "no shortage of math teachers" when we estimate Math or Numeracy scores.

determinants of the total achievement gap between each group for both literacy and numeracy in Table 5. We also report the part of the achievement gap that can be explained by all of our independent variables together. We further identify the portion of the gap that can be attributed to each of the independent variables. This can be done separately for each variable or for a group of independent variables (i.e., for family background, school quality, demographic features and country group components). We also calculate the fraction of the explained difference that is accounted for by each independent variable, which we sum (to a subtotal) for each component of interest. We repeat this procedure for the decompositions of the achievement gaps between natives and second-generation immigrants, and between first and second-generation immigrants.

Table 5
Blinder-Oaxaca Decomposition Results.

	Panel A: PISA scores						Panel B: PIAAC scores					
	Native –1 st Gen.		Native –2nd Gen.		2ndGen. –1 st Gen.		Native –1 st Gen.		Native –2nd Gen.		2ndGen. –1 st Gen	
	Read (1)	Math (2)	Read (3)	Math (4)	Read (5)	Math (6)	Lit (7)	Num (8)	Lit (9)	Num (10)	Lit (11)	Num (12)
Explained by Demographic variables												
Female	1.44	−0.51	0.62	−0.22	0.78	−0.28	0.00	0.09	0.01	0.27	−0.06	−0.25
Test language same as the language at home (%)	2.95	1.57	0.72	0.36	3.07	2.34	3.76	1.61	0.49	0.17	4.42	3.33
Subtotal (a)	4.39	1.06	1.34	0.14	3.85	2.06	3.76	1.7	0.5	0.44	4.36	3.08
Explained by Family Variables												
Either parent has a higher education degree	1.07	0.84	0.88	0.66	0.04	0.02	−1.61	−1.6	−0.61	−0.59	−0.81	−0.91
Books at home (All three dummies)	3.85	4.28	−0.03	−0.03	4.61	4.54	6.10	6.69	0.71	0.76	6.34	6.17
Subtotal (b)	4.92	5.12	0.85	0.63	4.65	4.56	4.49	5.09	0.10	0.17	5.53	5.26
Explained by School Resources												
% of government funding	0.30	0.31	0.02	0.03	0.48	0.47	−0.16	0.56	0.01	−0.02	−1.76	−1.26
% of full-time teachers with higher degree in pedagogy	−0.02	0.44	−0.13	0.36	0.02	0.00	1.73	1.19	1.02	0.71	0.8	0.74
No shortage of language/math teachers*	0.25	0.38	0.40	0.64	−0.14	−0.34	0.49	−0.19	0.3	−0.32	0.23	−0.02
School size (log)							−0.04	0.39	0.03	−0.46	−0.96	−2.62
Subtotal (c)	0.53	1.13	0.29	1.03	0.36	0.13	2.02	1.95	1.36	−0.09	−1.69	−3.16
Explained by School Autonomy												
Hiring teachers	−1.46	−1.38	−1.14	−1.05	−0.41	−0.16	−2.6	−2.62	−1.67	−1.69	−1.26	−1.58
Formulating budget	−0.12	−0.18	0.19	0.29	−0.43	−0.76	2.0	0.92	0.08	0.04	2.75	3.08
Determining course content	0.02	0.02	−0.01	−0.01	0.00	−0.18	−1.4	−1.27	−0.93	−0.86	−0.51	−0.57
Subtotal (d)	−1.56	−1.54	−0.96	−0.77	−0.84	−1.1	−2.00	−2.97	−2.52	−2.51	0.98	0.93
Explained by School Accountability												
Inform parents child's progress	0.00	−0.01	0.00	0.00	0.01	0.02	−1.5	−1.14	−0.44	−0.39	−1.86	−2.97
Decide grade retention/promotion	−0.25	−0.26	0.03	0.03	0.06	0.15	8.11	8.02	1.71	1.72	5.75	9.21
Group students	−0.04	−0.02	−0.10	−0.06	0.16	0.13	0.31	−0.05	−0.47	0.2	−0.91	−0.28
Subtotal (e)	−0.29	−0.29	−0.07	−0.03	0.23	0.3	6.92	6.83	0.8	1.53	2.98	5.96
Explained by Country Group												
Significant recent migration and humanitarian countries	0.15	0.02	0.22	0.02	−0.03	0.06	−4.68	−7.17	0.72	1.13	−3.04	−7.17
New destinations	−1.45	−5.51	−1.66	−6.53	0.07	0.81	−1.26	−1.1	−1.04	−0.83	−0.06	−0.02
Eastern European destinations	10.90	8.67	4.17	3.36	6.63	5.08	−1.47	−0.09	−0.6	−0.05	−1.61	−1.27
Subtotal (f)	9.6	3.18	2.73	−3.15	6.67	5.95	−7.41	−8.36	−0.92	0.25	−4.71	−8.46
Total Achievement Gap = (g)	40.38	30.10	5.26	1.33	35.12	28.78	24.4	23.14	2.69	6.39	21.71	16.76
Demographic variables (a) / Total Gap (g)	0.11	0.04	0.25	0.11	0.11	0.07	0.15	0.07	0.19	0.07	0.20	0.18
Family variables (b) / Total Gap (g)	0.12	0.17	0.16	0.47	0.13	0.16	0.18	0.22	0.04	0.03	0.25	0.31
School resources variables (c) / Total Gap (g)	0.01	0.04	0.06	0.78	0.01	0.00	0.08	0.08	0.51	−0.01	−0.08	−0.19
School autonomy (d) / Total Gap (g)	−0.04	−0.05	−0.18	−0.58	−0.02	−0.04	−0.08	−0.13	−0.94	−0.39	0.05	0.06
School accountability (e) / Total Gap(g)	−0.01	−0.01	−0.01	−0.02	0.01	0.01	0.28	0.30	0.30	0.24	0.14	0.36
Country Groups (f) / Total Gap (g)	0.24	0.11	0.52	−2.37	0.19	0.21	−0.30	−0.36	−0.34	0.04	−0.22	−0.50
Total explained / Total Gap (g)	0.51	0.34	0.30	−3.71	0.58	0.55	0.34	0.21	−0.24	−0.02	0.37	0.26
Total Explained	20.62	10.37	1.57	−4.92	20.40	15.73	8.26	4.89	−0.64	−0.14	8.06	4.33

Note: Source: PISA 2000 and 2003; PIAAC 2011−12 and 2014−15. The decomposition of group differences in change in PISA or PIAAC scores are between the two groups specified on the top row of the Table. Dependent variables for Panel A are PISA literacy and math scores. The gap between native and second-generation for literacy is not significant, so the results in that column are not statistically meaningful.

* We use "no shortage of language teachers" when we estimate Reading or Literacy scores, and we use "no shortage of math teachers" when we estimate Math or Numeracy scores.

7. Results and discussion

Table 4 shows the results of the estimation of Eq. (1) with simple OLS. Panel A presents the results for PISA and Panel B for PIAAC, for both literacy and math scores. Columns (1), (3), (5) and (7) report the results without the school quality variables, while Columns (2), (4), (6) and (8) include these. The dependent variable is achievement scores in math and reading for PISA and numeracy and literacy for PIAAC. We specify the IM (immigrant group) with a categorical variable that uses the group of natives as a reference. A negatively signed coefficient thus indicates that the result for the immigrant group is lower than for natives.

In general, the regression results confirm our observations in Fig. 1: first-generation immigrants are more disadvantaged than second-generation immigrants relative to natives, and for second-generation immigrants, literacy skill gaps are smaller than math/numeracy skill gaps. Second, being a first-generation immigrant is strongly negatively associated with lower performance on both the PISA and PIAAC tests. This is also true for second-generation immigrants, with the exception of PIAAC literacy scores, where the coefficient loses statistical significance when we introduce the school quality variables. For PISA, both literacy and math scores are about 20 points lower for first-generation immigrants compared to natives, which is about 21 percent of one standard deviation in the PISA scores for natives.

Controlling for school quality variables reduces the gaps slightly for both first- and second-generation immigrants. On average, PIAAC numeracy scores for first-generation immigrants are 20.04 points lower than for natives, without controlling for school quality variables. When we include the school quality controls, this difference declines to 17.86, or one-third (36 percent) of the standard deviation in PIAAC numeracy scores for our total sample. The numeracy scores for second-generation immigrants are 6.71 points lower than the adult numeracy scores for

natives, on average. Once we control for school quality variables, this difference decreases to 6.11.

As noted in Section 3, we measure school quality by school resources, school autonomy and school accountability. For school resources, having no shortage of language/math teachers correlates strongly and positively with students' literacy and numeracy skills. The share of public funding is negatively correlated with students' test scores on PISA, which is consistent with existing studies. This may reflect a tendency for academically weaker students to go to schools with a higher share of government funding. Further, different from our conventional assumption that more teachers who possess strong pedagogical content knowledge might be more effective in raising students test scores, our OLS regression results do not confirm this: although Column (2) shows that the percentage of full-time teachers with higher degrees in pedagogy is positively related to literacy scores in PISA, the coefficient is only marginally significant at the 10 % level. Further, Column (4) suggests that having more teachers with a higher degree in pedagogy is negatively correlated with PISA math scores. The negative magnitude is even larger in Column (8) when using PIAAC data, which implies that a higher percentage of teachers with a higher degree in pedagogical training at national level is not associated with higher levels of numeracy skills. Our results also show a positive and persistent relationship between other variables that proxy school quality and educational performance. We find that school autonomy in hiring teachers is strongly and positively associated with cognitive skills at age 15 and in young adulthood. The finding that school autonomy in allocating the budget is negatively correlated with students' performance is consistent with findings from previous studies (Fuchs and Wößmann, 2007; Hanushek et al., 2013). This result supports the assumption that the effect of school autonomy depends on specific decision-making areas. School accountability in decision-making regarding holding students back a year or advancing them ahead a year (retention/promotion) is also strongly and positively associated with cognitive skills. School policies that group students by ability show a positive and significant sign for literacy test scores in the PISA data.

As we might expect from our descriptive analysis, second-generation immigrants modestly underperform vis-à-vis natives (significant at the 10 percent level) on the PIAAC measurement for literacy, without controlling for the different qualities of the national education systems. The difference in literacy scores between second-generation immigrants and natives, however, becomes statistically insignificant, once the school quality variables are introduced. This does not hold for numeracy skills; the coefficients for second-generation immigrants remain significantly negative. For the second-generation immigrants, the size of the coefficient is about one-third the size of the coefficient for the first-generation immigrants. Comparing these results with the significant gaps between second-generation immigrants and natives, as assessed by the PISA data, suggests that some improvement occurs at the pseudo-cohort level that may also be attributable to different qualities of the national education systems. These results suggest that second-generation immigrants start with an arrear in literacy but seem to catch-up during their time in school. However, since the school quality variables can be calculated only at the country level, the results should be interpreted cautiously.

These regressions also control for the socio-economic background variables that are common in the two datasets. Consistent with the existing literature, we find that family background variables—which may indicate socio-economic status or proxy for the family's 'culture' towards education performance—show a strong and significant correlation to achievement scores. It is often suggested that this effect wanes as teenagers leave the family home and enter adulthood[22] and that the effect of school quality may emerge later in the education lifecycle. Our results confirm this to some extent. The number of books at home is positively and significantly related to achievement, with bigger magnitudes in the PISA data than in the PIAAC data. Parental education is also positively and significantly associated with higher PIAAC scores. Even though coefficients for parental education show statistically negative signs in the PISA data, the SES indicators have much larger and positive magnitudes.[23]

The country peer-groups have large and statistically significant coefficients, although the variability of these coefficients is quite high. The country-group variable is constructed as a categorical variable with 'longstanding destination' countries as a reference category; the results suggest that immigrants in 'recent migration and humanitarian' countries have higher achievement scores than those in 'longstanding destination' countries.[24] On the other hand, immigrants in 'new destination' countries have lower scores in literacy and numeracy, but the size of the coefficients decreases once school quality variables are introduced. Immigrants to Eastern European[25] countries have even lower literacy and numeracy scores. These results suggest that where immigrants go bears importantly on their educational attainment.

Finally, it is interesting to note that in our sample female students tend to have lower literacy scores when they were in school, but this gender disadvantage in literacy skills largely vanishes and becomes statistically insignificant in our adult pseudo-cohort sample. On the contrary, while females tend to do well in math assessment at age 15, their numeracy skills in the PIAAC data are much weaker compared to their male counterparts, whether or not we control for school quality variables. The results also show that when the language spoken at home corresponds to the test language, this has a positive and statistically significant effect on all achievement scores.

We also tested interaction effects. The results are shown in Table A1 in Appendix A. We interact the immigrant status dummies with the

[22] Cameron and Heckman (1998) found that the effect of family background diminished with higher levels of education, but a corollary may be that the family background effect simply fades with age.

[23] When we exclude ESCS in our OLS regressions, the coefficient of parental education becomes positive and statistically significant. The negative coefficients on parental education when including ESCS in the regressions also reflect the fact that immigrants have a higher percentage of parents who have tertiary education (see Table 3). It is worth noting that the variance inflation factor (VIF) test is done for all our regressions to test for multicollinearity and the VIF scores of all independent variables are lower than 3. Moreover, we prefer the models with ESCS because the adjusted R-square is much higher with the inclusion of the ESCS variable, indicating that SES is an important indicator in explaining the variation in students' PISA scores after controlling for parental higher education attainment. Nonetheless, we acknowledge that ESCS is an error-prone measure created using a latent response model. Because ESCS consists of parent education, it is likely that parental education and ESCS are collinear and biased, which may create a potential bias in other covariates.

[24] For PISA, the negative coefficients of this country group became statistically insignificant or even positive after controlling for school quality variables.

[25] Note that since the language variable is not available for Russia, this peer country group is represented by the Czech Republic only. The results in Table 4 exclude Russia. It is also possible that the distribution of immigrants in different countries in our sample may influence the country-group results as follows. For *Longstanding destinations:* the number of observations of first-generation immigrants in France is quite low with respect to the other countries in the group, therefore the results for first-generation immigrants are probably driven by the UK, Belgium and the Netherlands (the distribution of first gen immigrants among these three countries is fairly even). For *Significant recent migration and humanitarian countries*: the number of observations in Finland for first generation immigrants is quite low and the number of first-generation immigrants in Denmark is higher than Sweden, or Norway. The results for that country group are perhaps more driven by Denmark, and then by Sweden and Norway and less by Finland. For *New destinations*: there are very few observations for first generation immigrants in Italy, and more in Greece, Ireland and Spain. Therefore, Italy is not likely to be driving the results for this group, rather the three other countries. While there are some differences in the distribution of second-generation immigrants (for example, relatively lower numbers in Finland and Italy and Norway), the distribution is relatively even across countries, with no clear country driving the results with respect to second generation immigrants.

school quality variables to understand how the effects of the school quality variables vary across immigrant groups, particularly for the PISA sample. We find that both first- and second-generation immigrants under-perform with respect to their native counterparts when they are in schools with a higher percentage of public funding. If weaker students are more likely to be sorted into schools with more government funding, this suggests that immigrants, and especially first-generation immigrants, are more likely to be sorted. Interestingly, having no shortage of math teachers seems to correlate strongly with better math performance in PISA for second-generation immigrants. It is also worth noting that the threat of grade retention seems to be negatively associated with first-generation immigrants' performance in both literacy and math in PISA, which confirms that such school accountability policies can exacerbate the achievement gap between academically disadvantaged students and other students (Battistin and Schizzerotto, 2019).

Table 5 reports the results of the Oaxaca linear decomposition of the achievement gaps. This decomposition reveals the proportion of each gap that can be explained by the various (groups of) variables. Negative signs in the decomposition results indicate that the variable narrows the gap, and positive signs indicate a widening of the gap. Panel A presents the results from the PISA skills assessments, and panel B shows results from the PIAAC skills assessments. The first two columns in panel B are based on Eq. (5) and represent the differentials for the predicted achievement scores between natives and first-generation immigrants on the PIAAC test. As an example, the total mean difference in scores between natives and first-generation immigrants (found in the row labelled 'Total Achievement Gap = (g)' in Table 5) is 24.4 points in literacy and 23.14 points in numeracy. The decompositions indicate that for literacy, virtually none (0.0009) of the difference between natives and first-generation immigrants can be explained by being female. On the other hand, for numeracy, 0.09 points of the total achievement gap between natives and first-generation immigrants can be explained by being female. A much bigger part of the gap is explained by whether the test language is the language spoken at home, though this finding is more pronounced for literacy (3.76 points) than for numeracy (1.61 points).

In Table 5, Subtotal (a) sums the part of the gap explained by our demographic variables, by adding the coefficients in the linear decomposition to calculate the aggregate effect. The demographic characteristics explain 17 percent of the gap in literacy and 10 percent of the total achievement gap in numeracy. We also sum subtotals for family variables (subtotal b), school inputs (subtotal c), school autonomy (subtotal d), and school accountability (subtotal e), and at the bottom of the table we present the share of the gap that can be explained by each sub-total. For the achievement gap between natives and first-generation immigrants (as measured by PIAAC scores), 17 percent of the literacy gap and 18 percent of the numeracy gap can be explained by family background variables. School resources and the autonomy variables contribute only negligibly to the gap between natives and first-generation immigrants. A large part of both the literacy gap (28 percent) and the numeracy gap (30 percent) is explained by our school accountability variables. Country groupings do not seem to explain much of these particular gaps.

The row labelled 'Total explained'[26] indicates the proportion of the gap in achievement scores that can be explained by all of the independent variables together. In total, they explain 8.26 points or 34 percent of the total achievement gap between adult natives and first-generation immigrants in literacy scores (which totals 24.40 points) and 4.89 points, or around 20 percent, of the total gap in numeracy scores (23.14 points). Our results are consistent with McEwan and Marshall's decomposition results, in which all independent variables account for approximately 30 percent of the gaps in scores in Cuba and Mexico.

Our independent variables account for a much larger share of the gap in the PISA scores. For PISA, the family variables and country groups explain the largest portions of the total gaps for all groups. Among three sets of school quality variables—resources, autonomy and accountability—school resources and accountability represent part of the story in explaining achievement gaps. Nevertheless, the results suggest that other unobserved determinants of achievement explain a portion of the gaps between groups with and without immigrant backgrounds, as well.

To summarize, our empirical results suggest that specific elements of school quality, as represented in national education systems are important, especially given that those benefits may take some time to emerge. For example, in our OLS regressions (Table 4), adding the school quality variables slightly exacerbates the negative coefficients at age 15 (PISA). However, at the stage of young adulthood (PIAAC), once we control for the quality of the different national education systems of the pseudo cohort, the size of the negative coefficient for first-generation immigrants (vis-à-vis natives) decreases and for second-generation immigrants PIAAC literacy scores, the coefficient becomes statistically insignificant. Taken together with our decomposition results, these findings suggest that the effects of different qualities in education systems on educational achievement may be persistent. While it remains beyond the scope of available data to conduct a proper longitudinal analysis, the present 'screenshot' analysis at two points in time (based on birth cohorts) represents an improvement over cross-sectional studies analysing either only teenagers (i.e., relying on PISA data) or only (young) adults (i.e., relying on PIAAC data). Additional research might examine why the school quality variables relate to achievement scores in the way that they do, and why some of these factors might help to narrow the gaps while others seem to exacerbate the gaps.

8. Conclusion

The integration of immigrants into the labour market is key to a successful immigration policy. In this paper we analyse the dynamics of skills gaps between natives and immigrants and whether policies that target school quality can effectively decrease those gaps. We investigate differences in educational performance between first-generation immigrants, second-generation immigrants and natives, which may highlight barriers for immigrants entering the labour market (if negative) or may suggest added value from immigrants coming to the host country (if positive). Our results first show some convergence in skills between second-generation immigrants and natives over time. Second, the gap in literacy skills between first-generation immigrants and natives, and among first-generation and second-generation immigrants, has increased over time. Our decomposition results show that demographics and family background contribute to the achievement gaps between different groups. Further, we find that school quality variables, such as school autonomy and school accountability, contribute to decreasing the skills gap for young adults with different immigrant backgrounds, in particular regarding the numeracy gap. The implications of this study are clear and important. While first-generation immigrant pupils increasingly struggle relative to their native and second-generation immigrant peers, the quality of education that young immigrants receive can help to combat the challenges they face. Policies that improve school quality may thus help host countries to reap rewards from their immigrant populations.

Author statement

Alison Cathles: Conceptualization, Methodology, Formal analysis, Writing.

Dongshu Ou: Conceptualization, Methodology, Formal analysis, Writing.

[26] Typically, the total should be equal to the sum of the subtotals. In Table 4, however, for the sake of presentation, we do not show the decomposition results for all of our variables (for example, month and year of birth and SES composite score are not reported in results using PISA). Therefore, the total explained is taken from the regression output and cannot be easily related to the subtotals in the table.

Simone Sasso: Conceptualization, Methodology, Formal analysis, Writing.

Mary Setrana: Conceptualization, Visualization, Writing.

Tom van Veen: Supervision, Writing.

Acknowledgement

We thank the *Worldwide Universities Network (WUN)* for their financial support of this research. We thank three anonymous referees for their very helpful comments.

Appendix A

Table A1
OLS Regression Family inputs, School inputs and Literacy/Numeracy Skill Scores, with interaction terms for school quality variables and immigration group dummies.

	Panel A: PISA scores		Panel B: PIAAC scores	
	Literacy	Math	Literacy	Numeracy
Immigration Group (Natives are the reference)				
1 st-generation immigrants	6.81	0.25	−19.96***	−14.05***
	(9.37)	(9.39)	(5.01)	(4.53)
2nd-generation immigrants	−6.00	−3.88	0.34	−0.92
	(6.13)	(6.49)	(2.94)	(3.10)
Demographic variables:				
Female	−30.43***	10.86***	−0.51	−10.65***
	(0.82)	(1.10)	(0.99)	(1.08)
Test language same as the language at home	19.31***	10.88***	13.38***	7.14***
	(1.79)	(2.00)	(3.27)	(3.52)
Family variables:				
Either parent has a higher education degree	−11.95***	−9.11***	14.97***	14.90***
	(1.40)	(1.46)	(1.43)	(1.59)
Home owned				
11−100 books	24.97***	25.84***	22.47***	24.83***
	(1.67)	(1.66)	(1.88)	(2.04)
101−500 books	47.56***	51.20***	40.78***	44.05***
	(1.88)	(1.93)	(1.98)	(2.15)
More than 500 books	47.87***	53.44***	45.37***	50.03***
	(2.98)	(3.13)	(2.50)	(2.74)
ESCS (SES) - PISA Only	30.43***	27.85***		
	(0.82)	(0.84)		
School resources				
% of government funding in total school funding	−0.19***	−0.21***	0.17	−0.32
	(0.03)	(0.03)	(0.19)	(0.22)
No shortage of language/math teachers*	6.74***	10.18***	15.41***	−10.57
	(1.16)	(1.22)	(6.63)	(7.83)
% of full-time teachers with higher degree in pedagogy	0.28 (1.03)	−4.56*	−16.62***	−11.11***
		(1.14)	(2.41)	(2.75)
1 st-generation immigrants * % of government funding in total school funding	−0.45***	−0.37***	0.40 (0.74)	0.28 (0.63)
	(0.12)	(0.11)		
2nd-generation immigrants * % of government funding in total school funding	−0.14**	−0.13*	0.34 (0.51)	0.56 (0.57)
	(0.07)	(0.07)		
1 st-generation immigrants * No shortage of language/math teachers	−5.05	3.19 (5.37)	32.78	13.93
	(5.17)		(30.49)	(29.71)
2nd-generation immigrants * No shortage of language/math teachers	5.56 (3.51)	8.92**	−2.56	−11.56
		(3.55)	(16.82)	(18.85)
1 st-generation immigrants * % of full-time teachers with higher degree in pedagogy	1.13 (6.19)	6.40 (6.24)	−19.03	−18.38
			(11.75)	(13.09)
2nd-generation immigrants * % of full-time teachers with higher degree in pedagogy	4.13 (2.97)	5.26*	0.36 (7.72)	−5.10
		(2.69)		(8.53)
School autonomy				
Hiring teachers	12.56***	12.23***	41.12***	42.25***
	(1.28)	(1.32)	(4.03)	(4.50)
Formulating budget	−5.94***	−8.66***	−50.35***	−24.02***
	(1.33)	(1.44)	(6.10)	(7.28)
Determining course content	0.12	−0.74	−43.44***	−39.69***
	(1.31)	(1.32)	(4.16)	(4.88)
1 st-generation immigrants * Hiring teachers	0.25	−6.34	−1.55	−2.06
	(6.02)	(5.97)	(19.48)	(12.84)
2nd-generation immigrants * Hiring teachers	0.55	−7.14*	3.74	2.72
	(3.92)	(3.88)	(11.16)	(12.84)
1 st-generation immigrants * Formulating budget	−4.20	−6.58	24.07	−22.08
	(5.28)	(5.62)	(33.06)	(32.60)
2nd-generation immigrants * Formulating budget	−5.45	−11.21***	−19.87	−35.79*
	(3.80)	(3.81)	(19.73)	(21.27)
1 st-generation immigrants * Determining course content	−11.02*	0.52 (6.59)	−8.43	−5.70
	(6.27)		(19.37)	(23.75)
2nd-generation immigrants * Determining course content	6.59 (4.35)	13.50***	−6.84	−12.00
		(4.27)	(12.76)	(14.96)
School accountability				
Assessment is used to:				
inform parents child's progress	−0.19	0.70	−34.44**	−27.15*
	(2.72)	(2.66)	(13.79)	(14.71)
Decide grade retention/ promotion	11.84***	11.79***	41.46***	41.22***
	(1.44)	(1.58)	(3.72)	(4.17)
Group students	1.88	0.92	7.11	−4.43
	(1.25)	(1.31)	(4.93)	(5.47)
1 st-generation immigrants * Assessment is used to inform parents child's progress			16.34	−30.65
			(54.27)	(54.76)

(*continued on next page*)

Table A1 (*continued*)

	Panel A: PISA scores		Panel B: PIAAC scores	
	Literacy	Math	Literacy	Numeracy
2nd-generation immigrants * Assessment is used to inform parents child's progress			−18.96	−47.55
			(41.48)	(44.51)
1 st-generation immigrants * Decide grade retention/ promotion	−17.20***	−17.46**	−8.70	9.42
	(6.60)	(6.85)	(18.23)	(16.93)
2nd-generation immigrants * Decide grade retention/ promotion	−7.76*	−9.70*	−6.14	5.74
	(4.37)	(5.01)	(11.56)	(12.25)
1 st-generation immigrants * Group students	5.50	2.83	13.43	29.55
	(5.59)	(5.88)	(26.16)	(26.22)
2nd-generation immigrants * Group students	2.31	3.18	−4.23	18.34
	(3.73)	(3.77)	(14.91)	(16.76)
Country groups				
Significant recent migration and humanitarian countries	14.94***	0.76	20.71***	33.31***
	(1.46)	(1.53)	(3.43)	(4.18)
New destinations	−10.21***	−40.99***	−15.88***	−12.99***
	(1.30)	(1.35)	(2.28)	(2.53)
Eastern European destinations	−55.49***	−44.16***	−17.44***	−2.36
	(1.54)	(1.67)	(3.36)	(3.89)
R-squared	0.28	0.22	0.25	0.25
Number of observations	106,082	106,082	6,361	6,361

Note: Source: PISA 2000 and 2003; PIAAC 2011−12 and 2014−15. Dependent variables for Panel A are PISA literacy (Columns (1) and (2)) and math scores (Columns (3) and (4)). Dependent variables for Panel B are PIAAC literacy (Columns (5) and (6)) and math scores (Columns (7) and (8)). Each cell represents the coefficient of the corresponding variable estimated by Eq. (1) plus additional interaction terms for the two immigrant group dummies (first-generation and second-generation immigrants) and three sets of school quality variables. Note that the interaction term under Accountability 1 using PISA data is not included due to multicollinearity. For details on the variables please refer to the note under Table 2. We also include SES, month and year of birth in regressions using PISA data and age in regressions using PIAAC, but statistics of these three variables are not reported. Robust standard errors are in parentheses. * significant at 10 %, ** significant at 5%, *** significant at 1%.
*We use "no shortage of language teachers" when we estimate Reading or Literacy scores, and we use "no shortage of math teachers" when we estimate Math or Numeracy scores.

References

Allen, J., Levels, M., van der Velden, R., 2013. Skill mismatch and skill use in developed countries: evidence from the PIAAC study. ROA Research Memorandum, ROA-RM-2013/17. Research Centre for Education and the Labour Market (ROA), Maastricht.

Andon, A., Thompson, C., Becker, B., 2014. A quantitative synthesis of the immigrant achievement gap across OECD countries. Large-scale Assess. Educ. 2, 7. http://www.largescaleassessmentsineducation.com/content/2/1/7.

Azzolini, D., Schnell, P., Palmer, J., 2012. Educational achievement gaps between immigrant and native students in two "new immigration" countries: italy and Spain in comparison. Ann. Am. Acad. Pol. Soc. Sci. 643 (1), 46–77.

Battistin, E., Schizzerotto, A., 2019. Threat of grade retention, remedial education and student achievement: evidence from upper secondary schools in Italy. Empir. Econ. 56, 651–678. https://doi.org/10.1007/s00181-018-1443-6.

Borgonovi, F., Pokropek, A., Keslair, F., Gauly, B., Paccagnella, M., 2017. Youth in transition: how do some of the cohorts participating in PISA fare in PIAAC? OECD Education Working Papers, No. 155. OECD Publishing, Paris.

Cameron, S., Heckman, J., 1998. Life cycle schooling and dynamic selection Bias: models and evidence for five cohorts of American males. J. Polit. Econ. 106 (2), 262–333.

Cobb-Clark, D.A., Sinning, M., Stillman, S., 2012. Migrant Youths' Educational Achievement: The Role of Institutions. Ann. Am. Acad. Pol. Soc. Sci. 643 (1), 18–45.

Duong, M.T., Badaly, D., Liu, F., Schwartz, D., McCarty, C., 2016. Generational differences in academic achievement among immigrants youths: a meta-analytic review. Rev. Educ. Res. 86 (1), 3–41.

Dustmann, C., Tommaso, F., Gianandrea, L., 2012. Educational achievement of second-generation immigrants: An international comparison. Econ. Policy 27 (69), 143–185.

European Commission/EACEA/Eurydice, 2019. Integrating Students From Migrant Backgrounds in to Schools in Europe: National Policies and Measures, Eurydice report. Publications Office of the European Union, Luxembourg.

Flisi, S., Meroni, E., Vera-Toscano, E., 2016. Educational outcomes and immigrant background. JRC Technical Reports No. 102629. Joint Research Centre, EC.

Fuchs, T., Wößmann, L., 2007. What accounts for international differences in student performance? A re-examination using PISA data. Empir. Econ. 32, 433–464. https://doi.org/10.1007/s00181-006-0087-0.

Gal, I., Tout, D., 2014. Comparison of PIAAC and PISA frameworks for numeracy and mathematical literacy. OECD Education Working Papers, No. 102. OECD Publishing, Paris. https://doi.org/10.1787/5jz3wl63cs6f-en.

Hanushek, E., Woessmann, L., 2007. The role of education quality for economic growth (English). Policy Research Working Paper No. WPS 4122. World Bank, Washington, DC. http://documents.worldbank.org/curated/en/260461468324885735/The-role-of-education-quality-for-economic-growth.

Hanushek, E., Woessmann, L., 2017. School resources and student achievement: a review of cross-country economic research. In: Rosén, M., Hansen, K.Y., Wolff, U. (Eds.), Cognitive Abilities and Educational Outcomes: A Festschrift in Honour of Jan-Eric Gustafsson. Springer, pp. 149–171.

Hanushek, E., Link, S., Woessmann, L., 2013. Does school autonomy make sense everywhere? Panel estimates from PISA. J. Dev. Econ. 104, 212–232.

Hanushek, E., Schwerdt, G., Wiederhold, S., Woessmann, L., 2015. Returns to skills around the world: evidence from PIAAC. Eur. Econ. Rev. 73, 103–130.

Huijnk, W., Andriessen, I., 2016. *Integratie in Zicht*? (Integration in View) Report by the Dutch Social and Cultural Bureau of Policy Analysis (in Dutch with an English Summary).

Jacob, B., Rothstein, J., 2016. The measurement of student ability in modern assessment systems. J. Econ. Perspect. 30 (3), 85–108.

Jerrim, J., Lopez-Agudo, L., Marcenaro-Gutierrez, O., Shure, N., 2017. What happens when econometrics and psychometrics collide? An example using the PISA data. Econ. Educ. Rev. 61 (C), 51–58.

Levels, M., Dronkers, J., Kraaykamp, G., 2008. Immigrant children's educational achievement in Western countries: origin, destination and community effects on mathematical performance. Am. Sociol. Rev. 73, 835–853.

Marks, G., 2005. Accounting for immigrant non-immigrant differences in reading and mathematics in twenty countries. Ethn. Racial Stud. 28 (5), 925–946. https://doi.org/10.1080/01419870500158943.

Meroni, E., Vera-Toscano, E., Costa, P., 2015. Can low skill teachers make good students? Empirical evidence from PIAAC and PISA. J. Policy Model. 37, 308–323.

OECD, 2009. PISA Data Analysis Manual: SPSS, second edition. OECD Publishing, Paris. https://doi.org/10.1787/9789264056275-11-en.

OECD, 2010a. Closing the Gap for Immigrant Students: Policies, Practice and Performance, OECD Reviews of Migrant Education. OECD Publishing, Paris. https://doi.org/10.1787/9789264075788-en.

OECD, 2010b. Equal Opportunities? The Labour Market Integration of the Children of Immigrants: The Labour Market Integration of the Children of Immigrants. OECD Publishing, Paris.

OECD, 2013. Technical Report of the Survey of Adult Skills (PIAAC). OECD Publishing, Paris.

OECD/European Union, 2015. Indicators of Immigrant Integration 2015: Settling In. OECD Publishing, Paris. https://doi.org/10.1787/9789264234024-en.

Perry, A., Wiederhold, S., Ackermann-Piek, D., 2014. How can skill mismatch be measured? New approaches with PIAAC. Methods Data Anal. 8 (2), 137–174.

PISA (Programme for International Student Assessment) database, 2000. Database and supporting documentation available at: https://www.oecd.org/pisa/data/databasepisa2000.htm.

PISA (Programme for International Student Assessment) database, 2003. Database and supporting documentation available at: https://www.oecd.org/pisa/data/databasepisa2003.htm.

Portes, A., Fernández-Kelly, P., 2008. No margin for error: educational and occupational achievement among disadvantaged children of immigrants. Ann. Am. Acad. Pol. Soc. Sci. 620, 12–36.

Schnell, P., Azzolini, D., 2015. The academic achievements of immigrant youths in new destination countries: evidence from southern Europe. Migr. Stud. 3 (2), 217–240.

Schnepf, S.V., 2007. Immigrants' educational disadvantage: an examination across ten countries and three surveys. J. Popul. Econ. 20, 527–545.

Schofield, J.W., 2010. International evidence on ability grouping with curriculum differentiation and the achievement gap in secondary schools. Teach. Coll. Rec. 112 (5), 1490–1526.

Smith, W., Brezicha, K., Persson, A., 2015. Factors influencing the immigrant achievement gap among traditional Settlement countries: a look at within and across group assimilation. In: Islam, K., Iftekhar, S. (Eds.), International Perspectives on Comparative Education Policy. Bloomsbury, London, pp. 151–167.
Sori, I., Susteric, N., Gaber, S., 2011. Immigrant students' achievements in Croatia, Serbia and Slovenia in context. CEPS J. 1 (3), 31–51.
von Davier, M., Gonzalez, E., Mislevy, R., 2009. What are plausible values and why are they useful?. In: IERI Monograph Series: Issues and Methodologies in Large Scale Assessments, 2, pp. 9–36.
Woessmann, L., 2016. The importance of school systems: evidence from international differences in student achievement. J. Econ. Perspect. 30 (3), 3–32.
Wößmann, L., 2003. Schooling resources, educational institutions and student performance: the International evidence. Oxf. Bull. Econ. Stat. 65 (2), 117–170.
Wößmann, L., Lüdemann, E., Schütz, G., West, M., 2007. School Accountability, Autonomy, Choice, and the Level of Student Achievement: International Evidence From PISA 2003, OECD Education Working Papers, No. 13. OECD Publishing. https://doi.org/10.1787/246402531617.
Zinovyeva, N., Felgueroso, F., Vazquez, P., 2014. Immigration and student achievement in Spain: evidence from PISA. SERIEs (Berl) 5 (1), 25–60.