Are children of immigrants graded less generously by their teachers than natives, and why? Evidence from student population data in Italy

Moris Triventi

Moris Triventi
University of Trento
Department of Sociology and Social Research
Via Verdi 26, 38122 – Trento
Email: moris.triventi@unitn.it

Author's short bio
Moris Triventi, PhD, is Senior Assistant Professor in the Department of Sociology and Social Research at the University of Trento (Italy). His research interests comprise social stratification and mobility, educational inequalities and quantitative methods. His works have been published, among others, in European Sociological Review, Social Science Research and Work, Employment & Society.

Abstract: This article investigates whether teachers grade students with a migration background (SMBs) less generously than native students with comparable academic skills and examines the sources of such migrants' under-evaluation. We use population data from two whole cohorts of pupils enrolled in Italian primary and lower-secondary school. Using subject-specific standardized test scores as a yardstick, we found that SMBs were graded less generously by teachers than were natives with comparable ability, in both reading and mathematics. Applying the Blinder-Oaxaca method to assess which factors can account for SMBs' disadvantage, we found that the most relevant factors are language spoken at home and family socio-economic resources, but also some students' attitudes towards school matter, especially in lower secondary school. However, observable characteristics are far from accounting for all the teacher grading bias against SMBs, suggesting that unobserved factors and implicit discrimination processes could be at work as well.

Keywords: migrants' penalties, teachers' marks, academic performance, discrimination, education, social inequalities

Acknowledgements
I would like to thank Elisa Pini for excellent research assistance. This paper was presented at the ECSR Annual Conference at Bocconi University in Milan (31st August - 2nd September 2017). We would like to thank participants who made useful comments. Additional valuable feedback has been provided by Raffaele Guetto and Eleonora Vlach. Usual disclaimers apply.

Abstract: This article investigates whether teachers grade students with a migration background (SMBs) less generously than native students with comparable academic skills and examines the sources of such under-evaluation. Using population data from two cohorts of pupils enrolled in Italian primary and lower-secondary school and subject-specific standardized test scores as control variables, we found that SMBs were graded less generously by teachers in comparison to natives with comparable ability, in both reading and mathematics. Applying the Blinder-Oaxaca method to assess which factors can account for SMBs' disadvantage showed that the most relevant factors are language spoken at home and family socio-economic resources, but also some students' attitudes toward school matter, especially in lower secondary school. However, observable characteristics far from account for all teacher grading bias against SMBs, suggesting that unobserved factors and implicit discrimination processes could be at work as well.

Are children of immigrants graded less generously by their teachers than natives, and why? Evidence from student population data in Italy

Introduction

Education is widely conceived by policy makers and academics as a strategy to foster children of immigrants' socioeconomic integration in host countries, one which may have a positive long-term impact on the whole society (Heath, Rothon and Kilpi 2008; Alba et al. 2011). In recent decades, a number of studies have focused on how children with migration backgrounds perform in the school systems (e.g., Heath and Brinbaum 2014), showing that such children have lower achievement test scores (Kao and Thomson 2003), higher likelihood of dropout, and lower final educational attainment compared to children of natives (Schnepf 2007; Alba et al. 2011; Heath and Brinbaum 2014).

Nonetheless, from these studies it is not clear what role teachers play in the (re)production of disadvantages faced by children of immigrants in the early stages of their scholastic experience. Are teachers compensating the academic disadvantages of children of immigrants? Or, conversely, are they inadvertently contributing to their educational-related penalties? In an attempt to answer

these questions, this article investigates a relatively understudied potential source of disadvantage for students with a migration background (SMBs):[1] the way they are graded by teachers. More specifically, it examines whether teachers grade SMBs less generously compared to native students with comparable academic skills and what sources explain such gaps.

Grades are key signs of students' academic abilities (Johnes 2004) and are frequently used as the basis of school-related decisions by families (Boudon 1981; Gambetta 1987). Students' exposure to unfair treatment at school can undermine their development of civic values and trust in institutions (anti-socialization) (Jasso and Resh 2002; Resh 2013). Thus, if teachers tend to under-evaluate children of immigrants, this practice could negatively reverberate on such children's attachment to school, motivation to learn, educational aspirations, and chances to enroll in university (OECD 2015), thereby hindering the process of integration in the destination country. Such longer-term negative consequences of teacher grading bias against SMBs highlight this topic's importance for the international migration literature. Therefore, a closer examination of how teachers assign grades in school and with what effect on SMBs' educational disadvantages is worthy of investigation.

Some existing research analyzes teacher discrimination against children of immigrants, finding contrasting results (e.g., van Ewijk 2010; Sprietsma 2013). Using experimental design on small samples, van Ewijk (2010) reports no signs of discrimination in primary education in the Netherlands, while other scholars find evidence of discrimination in grading against pupils with foreign surnames in Germany (Sprietsma 2013) and Russia (Akifyeva and Alieva 2016). Most studies trying to infer teachers' discrimination from observational data find evidence of discrimination against SMBs (Botelho et al. 2015; Hinnerich et al. 2015), but others find none or

[1] This article uses the expressions "children of immigrants" and "students with a migration background" interchangeably. We use the term "natives" to refer to children born in Italy with both parents born in Italy.

detect even a premium for specific categories of SMBs (Burgess and Greaves 2013; Lindahl 2007). A study based on PIRLS (Progress in International Reading Literacy Study) and PISA (Programme for International Student Assessment) data, for example, found that second-generation immigrants in Germany were affected by grade discrimination in primary school but not in secondary education (Kiss 2013). In the United States, some authors have emphasized that teachers may have "positive stereotypes" of specific categories of ethnic minorities. For instance, Asian-American students have been commended as "model minorities" depicted as hardworking, respectful to teachers, and academically superior, especially in mathematics and science (Chang and Sue 2003). Therefore, it is not clear *a priori* if the grading bias would be against or in favor of children of immigrants, although most studies lean toward the first scenario.

While this overall body of work provides important findings, we argue that its exclusive focus on teacher discrimination neglects other potential sources of bias against SMBs, which could be related to student and parental resources, behavior, and attitudes. For instance, children of immigrants could be graded less generously by teachers compared to equally performing natives if they misbehave in classroom or if their parents are less involved in their scholastic careers. Therefore, as suggested by Burgess and Greaves (2013), a closer look at the overall under-/over-assessment of specific categories of students is warranted.

In this article, we conceive of grades not only as indicators of student ability and mastery of curricula but also as the outcomes of a more complex process in which scholastic evaluations are negotiated between students and teachers, as well as by parents and teachers. Adopting a multidisciplinary perspective and relying on contributions from various research streams (international migration and educational studies, social stratification research, economic studies of teachers' grading practices), we propose an analytical distinction between different potential

sources of grading bias against SMBs. At the empirical level, we use data from the National Agency for the Evaluation of the School System on two whole cohorts of students enrolled in primary and secondary education in Italy in 2012. By integrating grade-equation regression models with the Oaxaca-Blinder decomposition, we assess not only the overall gap between natives and children of immigrants but also its sources, looking at a variety of characteristics (parents' socioeconomic status, family structure, language spoken at home, educational resources, and students' school-related attitudes and behavior). With this technique, we have the possibility to establish the single contribution of each student characteristic to the gap, thereby providing a detailed picture of the sources of teacher grading bias against SMBs. While the Blinder-Oaxaca technique is widely used in the literature on wage gaps by race and gender (e.g., Blau and Kahn 2003; Darity, Guilkey, and Winfrey 1995), it has not yet been applied to the mismatch between teachers' grades and students' academic competencies.

Our analysis of the Italian case can offer interesting insights on a nation with relatively recent immigration flows from a wide number of countries (Santagati 2015) and in which most teachers did not receive specific training to tackle the challenges of teaching in a multicultural educational environment (OECD 2014). The few studies in this area of research have been conducted in Nordic countries or in "old immigration societies," such as continental European countries (e.g., Hinnerich, Höglin and Johannesson 2011; 2015; Kiss 2013; Lindhal 2007; 2016; Lüdemann and Schwerdt 2013). Here, we instead focus on a Southern European country that has gradually become a destination of significant immigration flows (Colombo and Sciortino 2004; Santagati 2015) and in which the share of children with a migration background in schools has been rapidly rising in recent years (Miur 2017).

Moreover, the focus on the Italian case is also motivated by data pointing to potential difficulties in the relationships between teachers and children of immigrants in the country. On the student side, Italy has the highest percentage of foreign early school leavers in Europe (around 35%; Bertozzi 2016). On the teacher side, according to the 2013 OECD Teaching and Learning International Survey (TALIS), Italy has one of the highest shares of lower-secondary teachers indicating that they feel the need for professional development in the area of teaching in a multicultural or multilingual setting (OECD 2014). Thus, the Italian compulsory education represents a critical case study to investigate whether teachers, with limited experience in teaching and assessing the performance of children of immigrants, are affected by systematic biases in their academic evaluations of students.

The article is organized as follows. The next section provides contextual information on the Italian educational system and the presence of SMBs in it, while the third section develops our theoretical framework. The fourth section presents the data, variables, and methods used in the empirical analysis. Section Five presents the empirical results, and the last section discusses the results in relation to our hypotheses and draws some conclusions.

The Italian context

Key features of compulsory school

The Italian education system is composed of primary school (for children between 6 and 11), lower-secondary school (for children between 11 and 14), upper-secondary school (age 14-19), and tertiary education, which formally takes between three and six years. Students start school at the age of six and are expected to stay in school until at least the age of 16. The first cycle of

schooling is organized into primary and lower-secondary education, with a declared goal of providing students with basic learning and tools for active citizenship (Eurydice 2014, 28).

For assessment purposes, the school year is divided into three-month or four-month terms. Periodic and final evaluations of student learning outcomes take place at the end of each term and are expressed in numerical grades, theoretically ranging from 0 to 10. A grade of 6/10 corresponds to a pass. There is no final examination at the end of primary school, and admission to lower-secondary school is based on a final assessment made by all teachers of a given class. Non-admission of a pupil to the following school grade (i.e., failure) is infrequent and is decided unanimously by all class teachers, who must provide specific reasons. At the end of the first cycle (the end of the third year of lower-secondary school), the leaving examination is held.

In primary education, evaluation focuses on pupils' learning process, behavior, and overall learning outcomes. These assessments should be consistent with the learning objectives established in each school's educational offer plan (POF). In the POF, each school's Teachers' Council defines the methods and criteria for assuring that pupil assessment is uniform, transparent, and fair (Eurydice 2014, 27). In lower-secondary education, the Charter of Students in Secondary Schools establishes that each student is entitled to "a transparent and rapid evaluation aimed at starting a process of self-evaluation to identify his/her own strengths and weaknesses and improve his/her own performance" (Eurydice 2014, 38).

Children of immigrants in the Italian school system

As a result of an increased immigrant presence in Italy and the changing composition of migration flows to it (Colombo and Sciortino 2004), students with foreign citizenship[2] in Italian

[2] Given the actual Italian laws on the acquisition of Italian citizenship (http://www.integrazionemigranti.gov.it/en/legal-framework/domestic-law/Pages/The-Italian-citizenship.aspx), these statistics underestimate the incidence of SMBs, as defined in our study.

schools increased by almost 400% between academic years 2001/2002 and 2011/2012, from 196,414 (2.2%) to 755,939 (8.4%) of the overall student population (Miur 2017). The most recent statistics available indicate that in 2015/2016 the number of students with foreign citizenship was almost 815,000, accounting for 9.2% of all students (Miur 2017).

For the cohort analyzed in this article, the presence of foreign students is uneven across school levels, being higher in primary and lower-secondary school (9.5% and 9.3%, respectively, in 2011/2012) and lower in upper-secondary education (6.2%; Miur 2012). With regard to origin countries, Italian schools display a wide heterogeneity: there is a clear prevalence for East Europe (Romania, Albania, Moldova, Ukraine, and Macedonia), followed by Asia (China, India, Philippines, Pakistan, and Bangladesh), North Africa (Morocco, Tunisia, and Egypt), and South America (Ecuador and Peru). Thus, as in other European countries, immigrant groups in Italy are ethnically more heterogeneous compared to those in the United States (Thomson and Crul 2007). This situation suggests that integrating children of immigrants in Italy's educational system could be even more challenging given the heterogeneous backgrounds of children of immigrants, in terms of language spoken, religion, socio-economics, and cultural preferences and norms.

Children of immigrants in Italian schools lag behind native students in terms of academic performance and transitions to higher educational levels (Azzolini 2011). In particular, first-generation immigrant students are frequently enrolled in classes at a lower grade than that of their age group (Azzolini 2011), mainly because of language problems (Mussino and Strozza 2012). They also have higher risks of delayed educational careers (Dalla Zuanna et al. 2009; Mussino and Strozza 2012), choose less ambitious and non-university-oriented paths (Barban and White 2011; Contini and Azzolini 2016), and have higher dropout risks than native students in upper-secondary education (Azzolini and Barone 2013). Students born abroad also have lower academic

performance in terms of standardized test scores (Di Bartolomeo 2011; Azzolini et al. 2012; INVALSI 2017) and grades at the end of educational cycles (Checchi 2010; Barban and White 2011). By contrast, second-generation students' disadvantage compared to that of natives is less marked (Minello and Dalla Zuanna 2014).

Overall, this body of research shows that SMBs in Italy obtain lower educational outcomes than natives but has not paid particular attention to the role of teachers' judgements as a potential source of disadvantage for children of immigrants. One exception is a qualitative study showing that teachers often recommended migrant students to the vocational, rather than academic, track in upper-secondary education, even when these students performed well in lower-secondary education (Bonizzoni, Romito, and Cavallo 2016).

Theoretical framework and hypotheses

In this section, we develop a theoretical framework to better understand the possible mechanisms behind biased grading against SMBs. Teachers' grading practices are the result of a multifaceted assessment process which may reflect not only the objective skill levels achieved by students but also the perceived effort, motivation, and even behavior of such students (OECD, 2012; 2013). Thus, the educational science literature suggests that, when assigning grades, teachers take into consideration not only students' "pure performance" but also their learning trajectory, social background, and educational and occupational prospects (e.g., Tierney et al. 2011).

From our perspective, grades not only are a function of students' ability, motivation, and effort but can also be conceived of as the outcome of a more complex process in which scholastic evaluations are negotiated between students and teachers and between parents and teachers. This idea builds on previous arguments that school grades reflect daily interactions between teachers and students and can be used by these actors as incentives in strategic ways (Costrell 1994). We

expand this notion by including parents as an additional relevant actor in the "grading game," building on research on parental involvement and family-school relationships (Lareau 1987). In this way, we develop a more comprehensive framework to make sense of the possible causes of teacher grading bias. In particular, we use the heuristic model depicted in Figure 1 to understand the role of specific factors that can drive socially biased teacher grading against SMBs. Each factor could help explain SMBs' (putative) disadvantage if two conditions hold: i) the characteristic is unevenly distributed across SMBs and ii) the characteristic is associated with lower grades for equally competent students.

[Fig. 1 about here]

The first explanatory factors refer to pupils' behavior in the classroom and their attitudes toward school and learning. Teachers might be more generous in grading those students who regularly attend classes, are well-behaved in the classroom, actively participate in discussions, demonstrate interest in the topics taught at school, and do their homework regularly (Tierney et al. 2011). If SMBs participate only irregularly in classroom discussions, are less diligent with homework, and are less intrinsically motivated to learn the subject, we could expect that students' attitudes and behavior represent a first source of disadvantage in explaining less generous grading against them.

A second aspect refers to SMBs' language skills and integration in the host society. Language proficiency is an important precondition for integration and a key factor that helps students communicate effectively with teachers (Berry 1997; Schnepf 2007). Pupils with less well-developed syntactical and lexical skills, as well as less ability in managing abstract reasoning,

could have difficulties in communicating effectively with teachers and might fail to meet academic expectations (Bernstein 1961). Globally, many newly arrived immigrant students cannot yet read or adequately speak the predominant language of their host countries (Dustmann, Frattini and Lanzara 2014). The PISA surveys also report that, across economically developed countries, a large portion of SMBs regularly speak a language other than an official host-country language at home (OECD 2015).[3] Such behavior might be a source of disadvantage, since regularly speaking the new language could be considered a sign of acculturation for the origin family and improve pupils' language skills, which are important for success in the classroom (Berry 1997; Schnepf 2007). Indeed, previous cross-national research on Western countries found that a foreign language spoken at home is the single most important factor associated with the achievement gap in test scores between natives and SMBs (Dustmann, Frattini and Lanzara 2014). Therefore, language spoken at home could be a second factor explaining why SMBs are graded less generously by their teachers.

The third set of characteristics that might account for migrant students' under-evaluation in school pertains to their family characteristics. In particular, we argue that the frequency with which parents interact with teachers and the quality of such interactions could shape how teachers perceive students and assess their performance. Even if we do not have direct measures, three family characteristics can act as proxies of family-teacher relationships. The first refers to the educational resources to which students have access at home. While educational resources affect student attainment by providing students with better conditions for learning (e.g., Teachman 1987), in the case of teachers' grading, the mechanism behind the role of educational resources could be

[3] On average, 63% of first-generation immigrant students and 38% of second-generation immigrant students speak a language at home that is different from the language in which the PISA test was conducted (OECD 2015).

indirect. A high level of educational resources at home and the availability of parental help for homework could be proxies for parents' interest (and capacity) to invest in their children's schooling. Furthermore, in the United States, it has been shown that teachers' approach in assigning and grading homework in early educational stages presumes that parents can help children at home (Lareau 1987). Therefore, if SMBs have lower access to such resources, they could be penalized in their daily interactions with teachers and receive lower grades than their native counterparts.

The fourth characteristic that could account for teacher grading bias against SMBs' refers to family structure, in particular to the number of siblings and presence of a single-parent family, which are negatively related to student scholastic achievement (Tanskanen et al. 2016). According to the parental resource dilution model (Downey 1995, 2001), parental resources should influence children's academic outcomes. Since parental resources are finite, the higher the number of children, the lower the level of resources invested in any particular child, all else being equal. One resource that could be important is parents' time to invest in developing a good relationship with their children's school principal and teachers. Families with many children may have less opportunity to be present at school meetings, to regularly meet teachers, and to talk about their children's learning and grades. A similar argument could be put forward in the case of single-parent families. We know that, in European countries, SMBs come from families with a higher number of children than do native students (Mussino and Strozza 2012; Kulu and Hannemann 2016). Moreover, children of immigrants in European countries, especially first-generation pupils, are likely to experience temporary or prolonged family disruption due to differential migration time of their parents (Landale et al. 2011). One can expect that these family conditions provide serious challenges to parents' opportunities to develop and maintain productive and continuous relationships with teachers. Since both cross-sectional (e.g., Grolnick & Slowiaczek, 1994) and

longitudinal (e.g., Miedel & Reynolds, 1999) studies have reported an association between higher levels of parental school involvement and higher academic success for children, this factor could exacerbate biases against SMBs in the evaluation process at school.

The fifth characteristic we consider are parents' socioeconomic and cultural resources, which are proxied by their social class, educational level, and books at home. Work by Lareau (1987) clarifies why these characteristics may matter for the way teachers evaluate students:

> The working-class parents had poor educational skills, relatively lower occupational prestige than teachers, and limited time and disposable income to supplement and intervene in their children's schooling. The middle-class parents, on the other hand, had educational skills and occupational prestige that matched or surpassed that of teachers; they also had the necessary economic resources to manage the child care, transportation, and time required to meet with teachers, hire tutors and become intensely involved in their childrens' [sic] schooling. (10)

Due to difficulties in labor-market integration in European countries, SMBs have a disproportionally higher share of parents with lower-level occupations and socioeconomic status (Heath et al. 2008). Therefore, a part of their disadvantage could be due to parents' lower cultural and socioeconomic resources, which put these students in a subordinate position compared to the teachers' own social status.

Finally, under-evaluation of SMBs could be due to conscious or unconscious forms of discrimination. Economists distinguish between two forms of discrimination: statistical and taste-based discrimination (Guryan and Charles 2013). Statistical discrimination could occur if grading is affected by teachers' beliefs about ability or performance associated with particular foreign backgrounds. Teachers who are unsure of how to grade a student may, for instance, be affected by the perceived ability of the ethnic group to which the student belongs. A second source of discrimination is taste-based discrimination, which involves preferring one group over another on

the basis of internalized social prejudices. If stereotypes enter teachers' judgements and evaluations, specific categories of disadvantaged students could be exposed to substandard treatment in school (James 2004). In line with these arguments, Hinnerich et al. (2015) compared blind and non-blind evaluations of the same test and report substantial discrimination against students with a foreign background in Sweden. In the United States, it was found that teachers tend to assess students from ethnic minorities and low-income families more negatively than white Americans and pupils from higher-income families (Epstein & Dauber, 1991). These findings suggest that discriminatory patterns in school grading practices might be at work, although it is not easy to empirically distinguish the two potential sources of discrimination.

Data, variables and methods

Data

Our empirical analysis uses restricted-access data collected by the National Institute for the Evaluation of the Italian School System on the *whole population* of students enrolled in the 5th grade (primary education) and 6th grade (lower-secondary education) in the 2011/12 academic year (INVALSI 2012). We focus on these grades because they are part of compulsory education in Italy, which is comprehensive (students are not already sorted into different tracks), and because dropouts have not yet occurred.

The INVALSI data are particularly suitable for our purpose because they contain students' scores in standardized tests (administered at the national level) in reading/Italian and mathematics, teachers' grades on the mid-term reporting card in the same two subjects (derived from administrative registers), and information on students derived from both administrative registers and an *ad hoc* student questionnaire. INVALSI tests and questionnaires for the 2011/12 school

year were administered May 9 and May 11, 2012, in the fifth grade (last year of primary school) and May 10, 2012, in the sixth grade (first year of lower-secondary school). The expected time for completing each test was differentiated according to the school level but uniform for all students within the same school level.

The INVALSI enforces a detailed Protocol (i.e., INVALSI 2012; 2017) for test administration and marking to reduce the possibility of teachers' cheating. In line with most national evaluation programs (Eurydice, 2009), the test is not administrated by the class teacher but by teachers of other classes and specialized in a different subject with respect to the one that is tested. All school teachers are simultaneously involved in the grading process, thereby allowing for mutual cross-checking during grading. The School Head, who is responsible for the protocol's correct implementation, supervises the whole process. At the end, an external specialized institution is charged with computing the test scores, using an automated procedure.

In a random sample of targeted schools, the tests took place in the presence of an independent observer, as well as in the presence of school teachers. Thanks to this external control, it was possible to use the random sample data to correct the national assessment for bias due to cheating, according to the INVALSI procedure (INVALSI 2012). The total number of classes surveyed amounted to 29,804 in the 5th grade and 27,402 in the 6th grade. To conduct our analyses, we constructed a pooled dataset of students assessed in the two grades. For the multivariate analyses we restricted the sample, adopting a list-wise deletion of cases with missing information on the variables used as covariates in our regression models. The analytical sample size includes a total of 630,148 cases (309,895 in the 5th grade and 320,253 in the 6th grade).

Variables

Our outcome variables are teachers' grades on the midterm reporting card (February 2012) in two subjects: Italian and mathematics.[4] The original variables range from 1 to 10.[5] The form used by INVALSI to collect mid-term marks from schools asked for both the "written marks" and "oral marks," allowing (if existing) separate grading on the basis of different assessment tools. We preferred the "written marks," due to the fact that the test was administered in a paper and pencil setting. For students in the fifth and sixth grades with missing "written marks," we used the information on the "oral marks," since the correlation between the two types of grades (for those who had both) amounted to around 0.97.[6]

The key independent variable of interest was student migration background – a categorical variable, based on the combination of the mother's, father's, and student's places of birth. The four categories are (1) Natives (students born in Italy with both parents born in Italy); (2) Mixed (students born in Italy with one parent born in Italy and one born abroad); (3) 2GEN (students born in Italy to foreign-born parents); and (4) 1GEN (foreign-born students with at least a foreign-born parent).[7] Unfortunately, the data do not contain information on the specific origin country.

The key control variables considered in our study were student scores on the INVALSI standardized tests of Italian and mathematics. Subject-specific competence was measured by INVALSI, applying Item Response Theory (specifically a Rasch model) to students' answers to

[4] Unfortunately, the datasets do not provide information on the final year marks.
[5] The proportion of students who received a grade between 1 and 3 is less than 0.2%, and excluding them does not affect our results.
[6] The high correlation between written and oral marks is probably due to the fact that it is quite unusual to have separate grades for written and oral assessments in primary and lower-secondary schools; it is plausible that in many cases, schools simply reported the same grade twice.
[7] When the information on origin country was not available, we relied on information about citizenship in order to exclude the smallest number of cases.

the SNV tests;[8] test scores are standardized to have a mean of 200 and a standard deviation of 40 in the whole original sample. We used the scores adjusted for potential cheating, already provided in the restricted-access data. To account for potential non-linearities in the relationship between marks and test scores, we included dummy variables measuring the grade-specific quintile of performance on the INVALSI test. We argue that for our purpose, the INVALSI test scores are a better control of academic proficiency, compared to those provided in the PISA survey (used, for instance, by Borghans and Lebedinski 2014), because they are supposed to measure competencies related to the curriculum taught by teachers in Italian schools instead of more general and valuable competencies for daily-life tasks. By these means, the INVALSI test scores should be more comparable to what teachers' grades are supposed to measure.

[Table 1 about here]

We used as basic control variables: the school's geographical area, students' gender, birth year, and birth quarter, and a dummy variable indicating whether the test was administered in presence of an external observer. In the second part of the analyses, we included a set of individual-level characteristics aimed at explaining teacher grading bias against SMBs. Family structure was measured first considering whether the student came from a family in which at least one parent was not present in the household because of divorced/separated parents or because one parent passed away prematurely (for parsimony we refer to this category as "non-intact family"). Second,

[8] Item Response Theory is a paradigm for the design, analysis, and scoring of tests that aims to measure individuals' abilities and competencies (Lord 2012). It is a theory of testing based on the relationship between individuals' performances on a test item and the test takers' levels of performance on an overall measure of the ability that item was designed to measure. A specific feature of this approach is that it does not assume that each item is equally difficult. For a broad overview on this approach, see Lord (2012).

we used the number of siblings to measure family size. In line with recent contributions in the social stratification literature (Bukodi and Goldthorpe 2013), family sociocultural and economic background was conceived as a multidimensional concept and measured using parental occupation, years of education, and number of books at home. Educational resources at home were captured by an index of material educational resources and a dummy variable measuring the lack of educational support at home.[9] We measured student attitudes and study behavior by relying on four indicators: the frequency of doing homework, an index of the child's study persistence, the student's degree of external motivation (studying to please other or to receive rewards), and an indicator of anxiety in taking the standardized test. This last variable was included because, since we used test scores as a yardstick to which grades were compared, it could be that some categories of students took the tasks involved in the test more seriously than others. Table 1 provides a list of the variables, their "status" in our analyses, and a description of how they were constructed and coded.

Methods

The common way to investigate teachers' grading in a given subject using observational data is to compare grades awarded by teachers with the results obtained by their students in standardized tests administered by independent institutional actors (Dardanoni et al. 2009). Test scores are considered a yardstick to which teachers' grades can be compared to assess the severity of their grading and their fairness in ranking students along the grade distribution (Bonesrønning, 1999; Lindahl, 2007). More generally, this strategy can be used to establish to what extent teachers

[9] Educational resources at home are only weakly correlated to socioeconomic background variables in our data.

take into consideration pupils' academic skills when attributing grades or, conversely, whether they are affected by other aspects as well (OECD 2013; Argentin and Triventi 2015).

The established approach to assessing whether some students are graded differently than others consists in estimating a "grade equation" model comparing a non-blind measure of teachers' evaluation (grades) with a blind and standardized measure of assessment of students' competencies (Dardanoni et al. 2009). We estimated the models of this form (*i* is the individual subscript):

$$GRADE_i = \alpha + \sum_{k=1}^{3} \delta_k SMB_i + \sum_{k=1}^{4} \beta_k \, TSCORE_i + \sum_{k=1}^{8} \varphi_k \, \boldsymbol{Z}_i + \varepsilon_i \qquad (1)$$

In this formula, *GRADE* measures teachers' grades, SMBs is student migration background, *TSCORE* is a vector of dummy variables indicating the quintile of the test score (to capture potential non-linearities), and *Z* is a vector including control variables previously listed. The model is estimated separately by subject (Italian, mathematics) and school level (primary, lower secondary).[10] These models are intended to establish whether and to what extent there is an overall over- or under-evaluation of students with various migration backgrounds (second-generation with both parents born abroad, second-generation with only one parent born abroad, first-generation) compared to natives with the same level of subject proficiency.

In the second part of the analysis, we made use of the Blinder-Oaxaca decomposition method (Blinder 1973; Oaxaca 1973), in particular the "two-fold" decomposition (Jann 2008). By these means, we were able to decompose differences in teachers' grades (net of students' performance in test scores) between natives and SMBs into a part that can be explained by

[10] We preferred to estimate separate models since we had a large number of cases, and, in this manner, we did not constrain the effects of the control variables to be the same across different subjects and school levels.

differences in group characteristics (the "endowments effect") and a part that cannot be explained by these variables. The second component measures the contribution of differences to the coefficients associated with the characteristics included in the model (and differences in the intercept) and is usually attributed to discrimination. It also, however, captures all potential effects of differences in unobserved variables.

Empirical findings

Descriptive statistics

As a first step we looked at the average grades in Italian and mathematics among different groups of students defined by their migration background. Table 2 shows that in the 5th grade, in both subjects, native students received, on average, higher grades (7.7), whereas first-generation migrant students were the worst performing, with an average of 6.8 in Italian and 7.0 in mathematics.

[Table 2 about here]

We also saw that the remaining two categories of students – "Mixed" and "2GEN" - were placed between native and first-generation immigrant students, but with a notable distinction: second-generation children with both foreign-born parents received grades more similar to first-generation immigrant students, while second-generation students with one foreign-born parent were much more similar to natives in their educational performance. This pattern resembled that of lower-secondary education, even if at this level the average grades were lower. The decline in average

grades from primary to lower-secondary education is typical of the Italian school system, in which grading becomes less generous along school levels (Argentin and Triventi 2015).

As a second step, it is useful to inspect the profiles of the four groups of students defined by their migration background (Table 3). Looking at family structure, native students come less frequently from a non-intact family and have fewer siblings compared to first- and second-generation immigrant students. We observed huge differences in terms of parents' occupation, with SMBs much more likely to have parents employed as "blue-collar workers." Differences in the same direction are found when books at home were considered. To the contrary, heterogeneity in parental years of education was much less pronounced. However, this measure does not consider where educational degrees were achieved; it could be that educational qualifications attained abroad have a different meaning than those attained in Italy.

[Table 3 about here]

If we consider the regular use of a foreign language to communicate at home as a proxy for language proficiency and a more general indicator of integration into the destination society, we observe striking heterogeneity across groups. While among second-generation students with one foreign-born parent, only 8% regularly spoke a foreign language at home, this proportion increased to 50% and 60%, respectively, among second-generation students with both foreign-born parents and among first-generation children.

Natives and immigrant students did not differ much in the material educational resources at home but seemed to differ more in terms of relational educational resources. However, the pattern went in an unexpected direction: first-generation students declared slightly less frequently

to have no one who could help them with homework compared to natives and other SMBs. Children of immigrants, especially first-generation immigrant students, displayed lower levels of perseverance in studying, higher levels of extrinsic motivation, and anxiety in doing the test, compared to natives. They also reported doing homework with less frequency than natives, but the differences are modest. These compositional differences provide us with preliminary insights on which factors could matter more in explaining teacher grading bias against SMBs. Indeed, a specific characteristic can be a source of grading bias only if it is unequally distributed among different categories of students.

Regression analysis

In the regression analysis, we first aimed to answer the following question: are SMBs more strictly judged by teachers than natives? If so, does this penalty vary across subjects and school levels? Figure 2 reports the estimated gap in average grades faced by various categories of SMBs compared to natives (reference category) with the same level of performance in subject-specific standardized test. As it shows, in primary and lower-secondary education and in both subjects, on average, SMBs are graded worse than equally competent native students. However, some heterogeneity in the penalty's magnitude is also visible.

[Fig. 2 about here]

Second-generation students with one parent born in Italy receive only slightly lower grades than natives with similar competence level. Larger differences, between 0.3 and 0.6 of a grade, are found when considering, respectively, second-generation students with foreign-born parents and

first-generation pupils. In terms of effect size, measured with standardized coefficients (see Table A1 in Appendix), our estimates point to a bias against SMBs similar to the one found in Germany (Kiss 2013; Lüdemann and Schwerdt 2013) and England (Burgess and Greaves 2013).[11] It is also interesting to compare the magnitude of Italy's native-immigrant gap to gender inequalities in grading, which have been extensively studied in other countries (e.g., Lavy 2008; Hinnerich et al. 2011). In Italy, the grading penalty against SMBs is much larger than the same disadvantage faced by male students, especially in primary education.[12]

The under-evaluation of children of immigrants is slightly larger in primary school than in lower-secondary education, resembling a pattern found by Kiss (2013) in Germany. Moreover, teacher grading bias against SMBs is stronger in mathematics than in Italian. One possible explanation is that teachers of Italian language are more aware of migrant students' language and integration difficulties and, as a consequence, more likely to compensate to some extent for this disadvantage in their attribution of grades. Further studies are needed to better understand the reasons behind the observed differences across school levels and subjects.

Decomposition analysis

Table 4 displays the results of the Oaxaca-Blinder decomposition, which are presented in terms of the percentage of the native-immigrant grading gaps accounted for by the whole set of

[11] For instance, Kiss (2013), analyzing the German case, reports an under-evaluation in reading of second-generation students amounting to -0.271 in primary school (PIRLS 2001) and virtually no differences across tracks in secondary education (PISA 2003). Lüdemann and·Schwerdt (2013), in contrast, report a penalty of 0.187 for second-generation students in German primary education, which shrinks to 0.10 (no longer statistically significant) when adjusting for social background and books at home.

[12] In primary education, SMBs' under-evaluation (compared to natives) is between 3 times (Italian) and 8 times (mathematics) larger than the estimated under-evaluation of males (compared to females). In lower-secondary education, the "migration penalty" is smaller but still comparatively sizable, since it is between 1.5 times (Italian) and 2 times (mathematics) larger than the estimated gender gap.

characteristics considered (family structure, socio-economic and educational resources, language used at home, and student school-related attitudes and behavior). We focus on the two comparisons that appeared the most relevant from the regression results: first-generation versus natives and second-generation with both foreign-born parents versus natives.

[Table 4 about here]

Irrespective of the grade and subject considered, the share of the teacher grading bias explained ranges from a minimum of 29.3% to a maximum of 45.7%. Thus, a large part of SMBs' under-evaluation is not explained by the observed characteristics available in our data. Students' characteristics account for a larger part of the gaps in primary school than in lower secondary education. There is not a clear pattern across subjects: in the 5th grade, the observed characteristics account for a larger share of the gaps in Italian, whereas in the 6th grade pupils' endowments explain slightly more of the gaps in mathematics.

Figures 3 and 4 report in graphical form the results of the detailed decomposition in which each single characteristic's role is estimated (the correspondent estimates are reported in Table A2 in the Online Appendix). This graphical representation is helpful to understand which observed characteristics have more weight in explaining why SMBs are less generously graded by teachers. Similarly to the previous analysis, the results are reported in percentage terms and for the two comparisons between categories of students. Figure 3 presents the results for primary education, whereas Figure 4 presents results for lower-secondary education. In the graph, the grey circles refer to reading, whereas the black squares refer to mathematics.

Looking at primary education, non-intact family did not play a noticeable role in accounting for the gaps between natives and children of immigrants, whereas number of siblings mattered mainly for second-generation students. Among the social background indicators, parental social class had a prominent role, accounting for around 8-9% of the gaps across subjects, followed by parental education (around 5%). Number of books and material educational resources at home, however, did not play important roles, with an estimated contribution between 0% and 2%. Language spoken at home emerged as an important factor accounting for teacher grading bias against SMBs in primary education, alone explaining around 9–11% of the gaps in both subjects.

One could expect that student attitudes toward school and study, as well as their behavior related to homework, could affect teachers' opinion and judgments. The data only partially supported this hypothesis. On the one hand, differences in homework-related behavior between natives and SMBs did not explain teacher grading bias; on the other hand, the degree of children's extrinsic motivation – being higher among children of immigrants and negatively associated with grades – accounted for a substantial part of the gaps, especially in mathematics (7-9%).

[Fig. 3 about here]

Notably, the relative importance of the single characteristics was similar when comparing first- and second-generation students to native students. While previous studies found that, in Italy, second-generation students performed better than first-generation pupils (Azzolini 2011), we showed that the mechanisms behind the disadvantage in grading are pretty similar across the two groups, at least in Italian primary education. Another attention-grabbing finding refers to the

similar patterns observed across subjects: teachers' grading bias against SMBs seems to be more linked to general, rather than to subject-specific, processes occurring in the classroom and school.

Figure 4 reports the results regarding pupils in lower-secondary education (6th grade). Overall, the main pattern is similar to that observed for primary school, but – in line with Table 4 – on average each characteristic accounted for a smaller part of the grading gap. The most notable differences compared to what is observed in primary education were as follows. First, parents' social class is relatively more important than language at home (6-8% versus 3-5%), which mattered much less in lower secondary than in primary education. Second, overall, student attitudes had a larger explanatory power, since study perseverance also accounted for a non-negligible share of the gaps (4%, on average across subjects and migration backgrounds). As found for primary school, the results were very similar across subjects.

[Fig. 4 about here]

To provide a synthetic picture of the main findings, Figure 5 displays the average share of the children of immigrants-natives gap explained by each characteristic examined across the different school levels and subjects, reported in descending order of magnitude (vertical lines). Additionally, all estimates are reported with different markers to convey an idea of the extent to which each factor's contribution is homogeneous or heterogeneous across the various estimates.

[Fig. 5 about here]

This figure clarifies that the most important factor was social background: in particular, social class – an indicator of family socio-economic resources and economic standing – played a relevant and consistent role across subjects and school levels. Language spoken at home and pupils' extrinsic motivations were respectively the second and third most important factors; however, both contributed to explaining SMBs' under-evaluation in primary education, but much less so in lower-secondary school. Other variables, such as pupils' behavior related to homework and help received by parents in doing homework, did not contribute substantially to explaining teachers' grading bias against SMBs in Italian compulsory education.

Robustness checks

To end, we performed sensitivity checks to assess our findings' robustness to alternative methodological choices. The first was to exclude students with grades that were extremely low (1–3), which are rather uncommon, especially in primary education. The second check was to restrict the analysis only to the random subsample of schools in which an external observer took part in administering the test. Third, we used different specifications of the test score variable, introducing it as i) a continuous variable with a linear effect on grades, ii) a continuous variable with a quadratic effect on grades, or iii) a categorical variable measuring students' position in terms of deciles. The results from all these sensitivity checks were identical or substantially similar to the ones presented in this article. Fourth, in our models we accounted for the interdependence of pupils sharing the same classroom by using clustered standard errors. An alternative was to use hierarchical models to account for the data's nested structure. Again, also using multilevel linear regression with the same model specification gave analogous results to the ones presented here.

Finally, we took into consideration a further mechanism that can potentially explain heterogeneity in the degree of children of immigrants' under-evaluation – that is, differences in the distribution of children of immigrants and natives across schools and their allocation to classes within schools. In particular, SMBs could be more often matched with lower-quality teachers or with those with weaker grading skills. In this case, under-evaluation would be the result of a compositional effect due to a process of segregation across schools and classrooms in which teachers were systematically more or less able to provide a fair judgment of their students. To rule out this possibility, we conducted an analysis based on classroom-level fixed-effect regression models. Since we did not have teachers' characteristics in our datasets but knew that most teachers had only one class per grade, introducing the classroom fixed-effects allowed us to take into account all unobserved characteristics of teachers (see Kiss 2013 for a similar strategy). By these means we could establish to what extent the same teacher was giving more or less generous grades to equally competent native and migrant students in the same classroom. Results reported in Table A3 in the Appendix suggest that our findings are not driven by segregation of children of immigrants across schools and classrooms with systematically different types of teachers.

Discussion and conclusions

In this article we proposed to look at a relatively understudied potential source of disadvantage in school for SMBs in the international migration literature: the way these students are graded by teachers in key subjects such as reading (Italian) and mathematics. The main object of interest was not simply how grades were distributed among students with different migration backgrounds but instead how teachers evaluated natives and children of immigrants that were approximately equally competent in a specific subject. To do so, we took advantage of recent data

on the whole population of pupils enrolled in primary (5th grade) and lower-secondary education (6th grade) in Italy. By means of the so-called "grade equation model," we assessed whether and to what extent SMBs were graded more or less generously by teachers compared to native students with analogous individual characteristics and performance in subject-specific standardized test scores. Even if our study's results have mainly a descriptive value, they can be informative regarding the grading disadvantage faced by children of immigrants in school and provide preliminary clues about the drivers of such gaps.

We found that SMBs tend to be under-valued by teachers: children of immigrants with a comparable ability to that of natives (same performance on standardized tests) obtained, on average, lower grades from teachers in both reading and mathematics. In terms of effect size, our results point to substantially large penalties for both second-generation students with parents born abroad and first-generation students. Similar to previous studies on other types of immigrant penalties in the school system (e.g., Barban and White 2011), first-generation students are the most penalized. Differently, and in contrast to the previous literature examining achievement gaps (e.g., Azzolini et al. 2012), we observed that second-generation students with both parents born abroad are in a position more similar to first-generation students than to natives. The large number of cases on which the analysis is based allowed us to distinguish among different migration backgrounds, and this distinction helped show that students with only one foreign-born parent are assessed by teachers in a similar way as native students. Thus, including them with second-generation students with both foreign-born parents could lead to a downward bias of teacher grading bias against second-generation students.

Extending previous studies, our work also tried to shed more light on student characteristics that may drive the tendency to grade SMBs less generously when they possess similar levels of

subject competencies. One may expect that teachers tend to factor into their evaluation not only the pure performance but also other elements such as students' genuine interest in the subject, diligence in doing homework, or approach to studying the subject (Gregory et al. 2010). We found that some of these elements, for instance student attitudes and motivations, played a non-negligible role in explaining teacher grading bias against children of immigrants. However, other homework-related behaviors appear to be less relevant in this respect. The most important factors appear to be language spoken at home and indicators of family socio-economic resources. This finding reinforces our argument that grade attribution not only is a teacher-student affair but also involves parents, who may put pressure on teachers, influencing (directly and indirectly) teachers' opinions about students' academic potential and prospects, or be more present in their children's school life.

This notwithstanding, one must bear in mind that the differences in observed characteristics between students with a migrant background and natives account for less than half the grading gaps. Importantly, while some studies find that the achievement gaps between natives and children of immigrants almost disappear when adjusting for language spoken at home and socioeconomic background (Marks 2005; Schnepf 2007), this is not the case for teacher grading in Italy. Since our models can explain only a minority of teacher grading bias, other factors could play a role in explaining the tendency to attribute lower grades to SMBs than natives with similar academic competencies.

Unfortunately, we were not able to include more direct measures of the relationship between teachers and students, student behavior in classroom, or detailed indicators of relations between parents and teachers or school principals, which in our framework should be important for influencing grade attribution. Yet, with such a large share of unexplained differences between SMBs and natives, it is plausible that some processes of discrimination by teachers, either implicit

or explicit, are at work. One may reasonably think that teachers, like other persons, are affected by daily interactions and exposed to mass media that can foster processes of stereotyping (James 2004). Recent evidence from England, for example, suggests that when assessing a pupil's likely progress, teachers use information on the performance of members of that group in that school from previous years (Burgess and Greaves 2013). It could be that stereotypical or statistical discrimination is an important factor responsible for the unexplained differences in grading between native students and SMBs in Italy, as was found in studies using psychological instruments to detect implicit prejudices (Van den Bergh et al. 2010) or experimental designs (e.g. Botelho et al. 2015).

Despite the limitations just mentioned, this article contributes to the literature on international migration by highlighting the need to study in more depth the sources of disadvantages SMBs face in school, since education is a key arena in which individuals form civic values, attitudes, and opinions. In particular, by showing that teachers tend to grade children of immigrants less generously than natives, we highlight an overlooked source of educational inequalities that can negatively influence SMBs' integration in the destination countries.

Our study may also have policy implications related to schools' daily functioning. Our findings point to the importance of school principals and teachers being aware of how their evaluation standards systematically vary across categories of students, in this case natives and SMBs. Increasing awareness of potential biases against children of immigrants could be done by providing simple and effective reports to principals, which can serve as starting points for discussion within schools (Argentin and Triventi 2016). Furthermore, in Italy there has been strong teacher resistance to the central educational agencies' attempts to introduce standardized tests as part of the final evaluation of student competencies in compulsory school and upper-secondary

education. Contrary to this position, and in line with previous recommendations by educational psychologists (Molouff 2008; Molouff et al. 2013), our results suggest that making students anonymous during grading might contribute to mitigating social inequalities in teachers' judgements, which reflect not only student competencies in the subject but also socially biased evaluations.

References

Akifyeva, R., & Alieva, A. (2016). The influence of ethnicity on teacher expectations and teacher perceptions of student warmth and competence basic research program. National Research University Higher School of Economics, Working papers series: Psychology, WP BRP 65/PSY/2016.

Alba, R., Sloan, J., & Sperling, J. (2011). The integration imperative: The children of low-status immigrants in the schools of wealthy societies. *Annual Review of Sociology*, 37, 395-415.

Argentin, G., & Triventi, M. (2015). The North-South Divide in School Grading Standards: New Evidence from National Assessments of the Italian Student Population. *Italian Journal of Sociology of Education*, *7*(2), 157-185.

Argentin, G., & Triventi, M. (2016). Come mi giudichi? Analisi delle pratiche e degli standard di attribuzione dei voti agli studenti nelle scuole italiane. In P. Falzetti P. (Ed.), *Concorso di idee per la ricerca*. Cleup, Padova, pp. 103-138.

Azzolini, D. (2011). A new form of educational inequality? What we know and what we still do not know about the immigrant-native gap in Italian schools. *Italian Journal of Sociology of Education*, *3*(1).

Azzolini, D., & Barone, C. (2013). Do they progress or do they lag behind? Educational attainment of immigrants' children in Italy: The role played by generational status, country of origin and social class. *Research in Social Stratification and Mobility*, *31*, 82-96.

Azzolini, D., Schnell, P., & Palmer, J. R. (2012). Educational achievement gaps between immigrant and native students in two "new" immigration countries: Italy and Spain in comparison. *The Annals of the American Academy of Political and Social Science*, *643*(1), 46-77.

Barban, N., & White, M. J. (2011). Immigrants' children's transition to secondary school in Italy. *International Migration Review*, *45*(3), 702-726.

Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied psychology*, *46*(1), 5-34.

Bertozzi, R. (2016). Transizioni e scelte formative. Opportunita` per gli allievi stranieri nelle diverse aree territoriali. In: Santagati M and Ongini V (eds) *Alunni con Cittadinanza Non Italiana. La Scuola Multiculturale nei Contesti Locali*. Milan: Fondazione ISMU, pp.81–108.

Blau, D. & Kahn, L.M. (2003) Understanding international differences in the gender pay gap. *Journal of Labor Economics* 21(1): 106–44.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436-455.

Bonesrønning, H. (1999). The variation in teachers' grading practices: causes and consequences. *Economics of Education Review*, *18*(1), 89-106.

Bonizzoni, P., Romito, M., & Cavallo, C. (2016). Teachers' guidance, family participation and track choice: the educational disadvantage of immigrant students in Italy. *British Journal of Sociology of Education*, *37*(5), 702-720.

Borghans, L. & Lebedinski, L. (2014). Do teachers grade immigrants like they grade native children?. IWAEE paper.

Botelho, F., Madeira, R. A., & Rangel, M. A. (2015). Racial discrimination in grading: evidence from Brazil. *American Economic Journal: Applied Economics*, *7*(4), 37-52.

Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in western society*.

Burgess, S., & Greaves, E. (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics*, *31*(3), 535-576.

Chang, D. F., & Sue, S. (2003). The effects of race and problem type on teachers' assessments of student behavior. *Journal of Cconsulting and Clinical Psychology*, *71*(2), 235.

Checchi, D. (2010). *Uguaglianza delle opportunità nella scuola secondaria italiana* (Vol. 25). FGA Working Paper.

Colombo, A., & Sciortino, G. (2004). Italian Immigration: The Origins, Nature and Evolution of Italy's Migratory Systems. *Journal of Modern Italian Studies*, 9, 49-70.

Contini, D., & Azzolini, D. (2016). Performance and decisions: immigrant–native gaps in educational transitions in Italy. *Journal of Applied Statistics*, *43*(1), 98-114.

Contini, D., & Triventi, M. (2016). Between formal openness and stratification in secondary education: Implications for social inequalities in Italy. In Blossfeld, H.-P., Buchholz, S., Skopek, J., and Triventi, M. (Eds.), *Models of Secondary Education and Social Inequality – An International Comparison*. Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 305–322.

Costrell, R. M. (1994). A simple model of educational standards. *The American Economic Review*, 956-971.

Dardanoni, V., Modica, S., & Pennisi, A. (2009). Grading across schools. *The BE Journal of Economic Analysis & Policy*, *9*(1).

Darity Jr, W., Guilkey, D., & Winfrey, W. (1995). Ethnicity, race, and earnings. *Economics Letters*, *47*(3-4), 401-408.

Dalla Zuanna, G., Farina, P., & Strozza, S. (2009). *Nuovi italiani. I giovani immigrati cambieranno il nostro Paese?*. Bologna: Bologna.

Di Bartolomeo, A. (2011). Explaining the gap in educational achievement between second-generation immigrants and natives: The Italian case. *Journal of Modern Italian Studies*, 16, 437–449.

Downey, D.B. (2001). Number of Siblings and Intellectual Development. The Resource Dilution Explanation. *American Psychologist* 56(6/7), 497-504.

Dustmann, C., Frattini, T., & Lanzara, G. (2012). Educational achievement of second-generation immigrants: an international comparison. *Economic Policy*, 27(69), 143-185.

Epstein, J.L., & Dauber, S.L. (1991). School programs and teacher practices of parent involvement in inner-city elementary and middle schools. *The Elementary School Journal*, 91, 289–305.

Eurydice (2009). National Testing of Pupils in Europe: Objectives, Organisation and Use of Results. Brussels: Education, Audiovisual and Culture Executive Agency.

Eurydice (2014). The Italian Education System. Firenze: INDIRE.

Gambetta, D. (1987). *Were they pushed or did they jump?:Individual decision mechanisms in education*. Cambridge University Press.

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*(1), 59-68.

Guryan, J., & Charles, K. K. (2013). Taste-based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots. *The Economic Journal*, *123*(572), F417-F432.

Heath, A.F. & Brinbaum, Y. (2014). *Unequal Attainments. Ethnic Educational Inequalities in ten western countries.* Oxford: Oxford University Press.

Heath, A. F., Rothon, C., & Kilpi, E. (2008). The second generation in Western Europe: Education, unemployment, and occupational attainment. *Annual Review of Sociology*, *34*, 211-235.

Hinnerich, B. T., Höglin, E., & Johannesson, M. (2011). Are boys discriminated in Swedish high schools?. *Economics of Education review*, *30*(4), 682-690.

Hinnerich, B. T., Höglin, E., & Johannesson, M. (2015). Discrimination against students with foreign backgrounds: evidence from grading in Swedish public high schools. *Education Economics*, *23*(6), 660-676.

INVALSI (2012). *Rapporto nazionale sulla rilevazione degli apprendimenti 2011-12.* Roma: Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e Formazione.

INVALSI (2017). *Rilevazione nazionale degli apprendimenti 2016-2017: Rapporto risultati*. Roma: Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e Formazione.

James, C. (2004). *Stereotyping and its consequence for racial minority youth*. Toronto: Ontario Human Rights Commission. Retrieved on 2017-11-06 from http://www.ohrc.on.ca/it/book/export/html/8977.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, *8*(4), 453-479.

Johnes, G. (2004). Standards and grade inflation. *International handbook on the economics of education*, 462.

Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology*, *29*(1), 417-442.

Kiss, D. (2013). Are immigrants and girls graded worse? Results of a matching approach. *Education Economics*, 21(5), 447-463.

Landale, N. S., Thomas, K. J. A., & Hook, J. V. (2011). The Living Arrangements of Children of Immigrants. The Future of Children / Center for the Future of Children, the David and Lucile Packard Foundation, 21(1), 43–70.

Lareau, A. (1987). Social class differences in family-school relationships: The importance of cultural capital. *Sociology of education*, 73-85.

Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics*, *92*(10-11), 2083-2105.

Lindahl, E. (2007). Comparing teachers' assessments and national test results: evidence from Sweden. Working Paper n. 2007:24, IFAU-Institute for Labour Market Policy Evaluation.

Lindahl, E. (2016). Are teacher assessments biased? Evidence from Sweden. *Education economics*, *24*(2), 224-238.

Lord, F.M. (2012). *Applications of item response theory to practical testing problems*. Routledge.

Lüdemann, E. and· Schwerdt, G. (2013). Migration background and educational tracking. Is there a double disadvantage for second-generation immigrants? *Journal of Population Economics*, 26:455–481.

Malouff, J. M. (2008). Bias in grading. *College Teaching*, 56, 191–192. doi: 10.3200/CTCH.56.3.191-192.

Malouff, J. M., Emmerton, A. J., & Schutte, N. S. (2013). The risk of a halo bias as a reason to keep students anonymous during grading. *Teaching of Psychology,* 40, 233–237. doi: 10.1177/0098628313487425.

Marks, G. N. (2005). Accounting for immigrant non-immigrant differences in reading and mathematics in twenty countries. *Ethnic and racial studies*, *28*(5), 925-946.

Minello, A., & Dalla Zuanna, G. (2014). Children of immigrants in the Italian school system: what kind of assimilation?. *Statistica applicata: Italian journal of applied statistics*, *23*, 193-213.

Miur. (2012). Gli alunni stranieri nel sistema scolastico italiano a.s. 2011/2012. MIUR: Ufficio di Statistica.

Miur. (2017). Gli alunni stranieri nel sistema scolastico italiano a.s. 2015/2016. MIUR: Ufficio di Statistica.

Miur-Ismu. (2013). Alunni con cittadinanza non italiana. L'eterogeneità dei percorsi scolastici. Rapporto nazionale A.s. 2012/2013.

Mussino, E., & Strozza, S. (2012). The delayed school progress of the children of immigrants in lower-secondary education in Italy. *Journal of Ethnic and Migration Studies*, *38*(1), 41-57.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693-709.

OECD. (2012). *Grade Expectations: How Marks and Education Policies Shape Students' Ambitions*. Paris: OECD Publishing.

OECD (2013). *Education at a Glance 2013. OECD Indicators*. Paris: OECD Publishing.

OECD (2014). *TALIS 2013 - Results An International Perspective on Teaching and Learning*. Paris: OECD Publishing.

Resh, N., & Sabbagh, C. (2014). Sense of justice in school and civic attitudes. *Social Psychology of Education, 17*(1), 51-72.

Schnepf, S. (2007). Immigrants' educational disadvantage: an examination across ten countries and three surveys. *Journal of Population Economics*, 20 (3), 527-545.

Sprietsma, M. (2013). Discrimination in grading: experimental evidence from primary school teachers. *Empirical Economics*, 1-16.

Teachman, J. D. (1987). Family background, educational resources, and educational attainment. *American Sociological Review*, 548-557.

Tierney, R. D., Simon, M., & Charland, J. (2011). Being fair: Teachers' interpretations of principles for standards-based grading. *The Educational Forum*, 75(3), 210-227.

Thomson, M., & Crul, M. (2007). The second generation in Europe and the United States: How is the transatlantic debate relevant for further research on the European second generation?. *Journal of Ethnic and Migration Studies*, 33(7), 1025-1041.

Van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal, 47*(2), 497-527.

van Ewijk, R. (2011). Same work, lower grade? Student ethnicity and teachers' subjective assessments. *Economics of Education Review*, 30(5), 1045-1058.

Figures

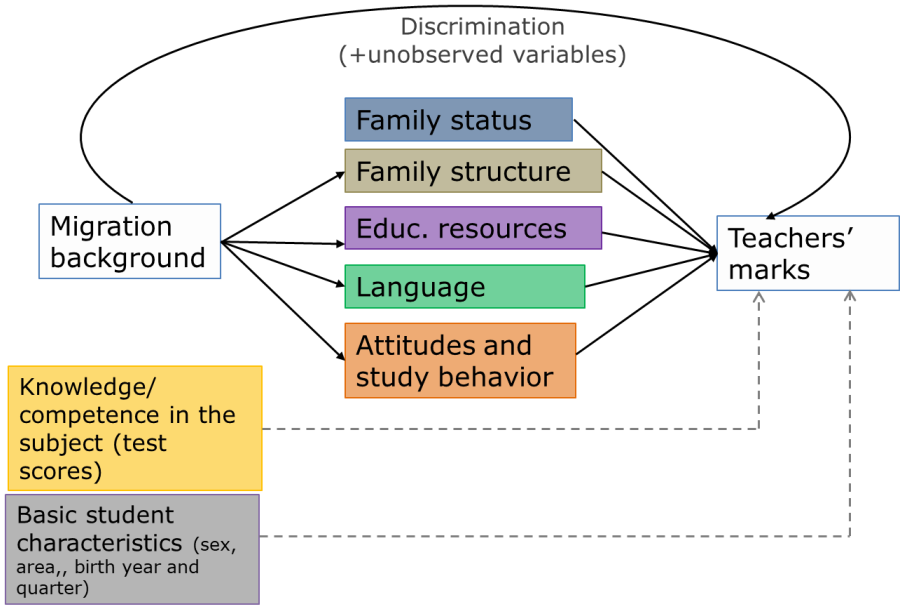

Fig. 1 – Conceptual model adopted in this study to analyze the under−evaluation of children of immigrants.

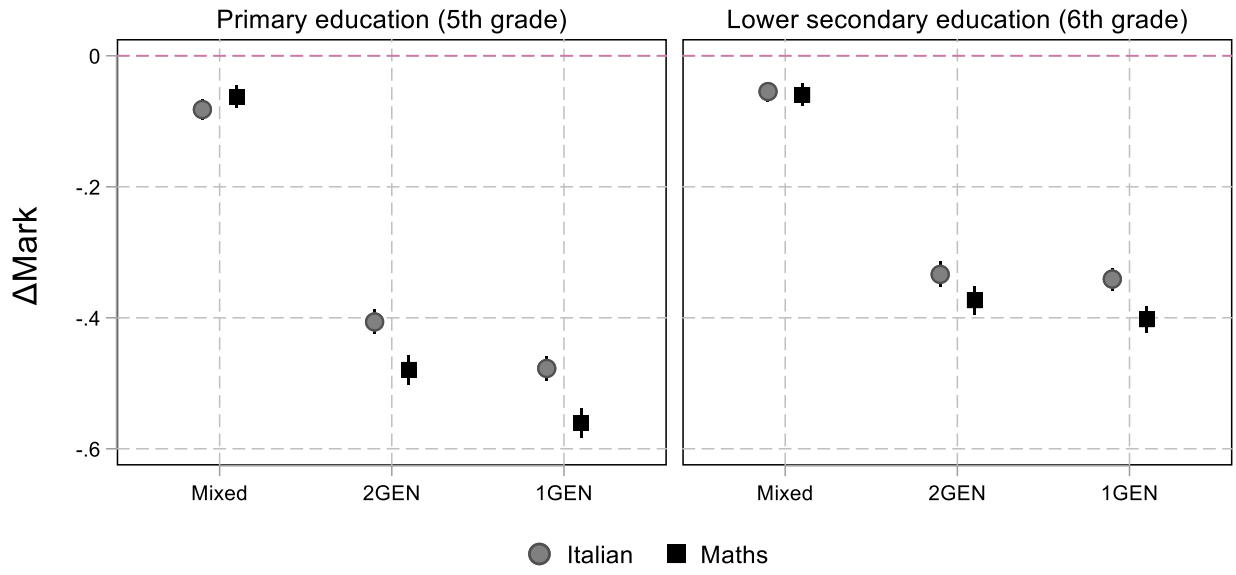

Fig. 2 − OLS linear regression estimates of the gap in grading between children of immigrants and natives in Italian (circles) and mathematics (squares), across school grades.

Note: the models control for performance in subject−specific standardized test scores, gender, geographical area, year and quarter of birth, dummy for a random sample of school in which an external observer was present when the test was administered.

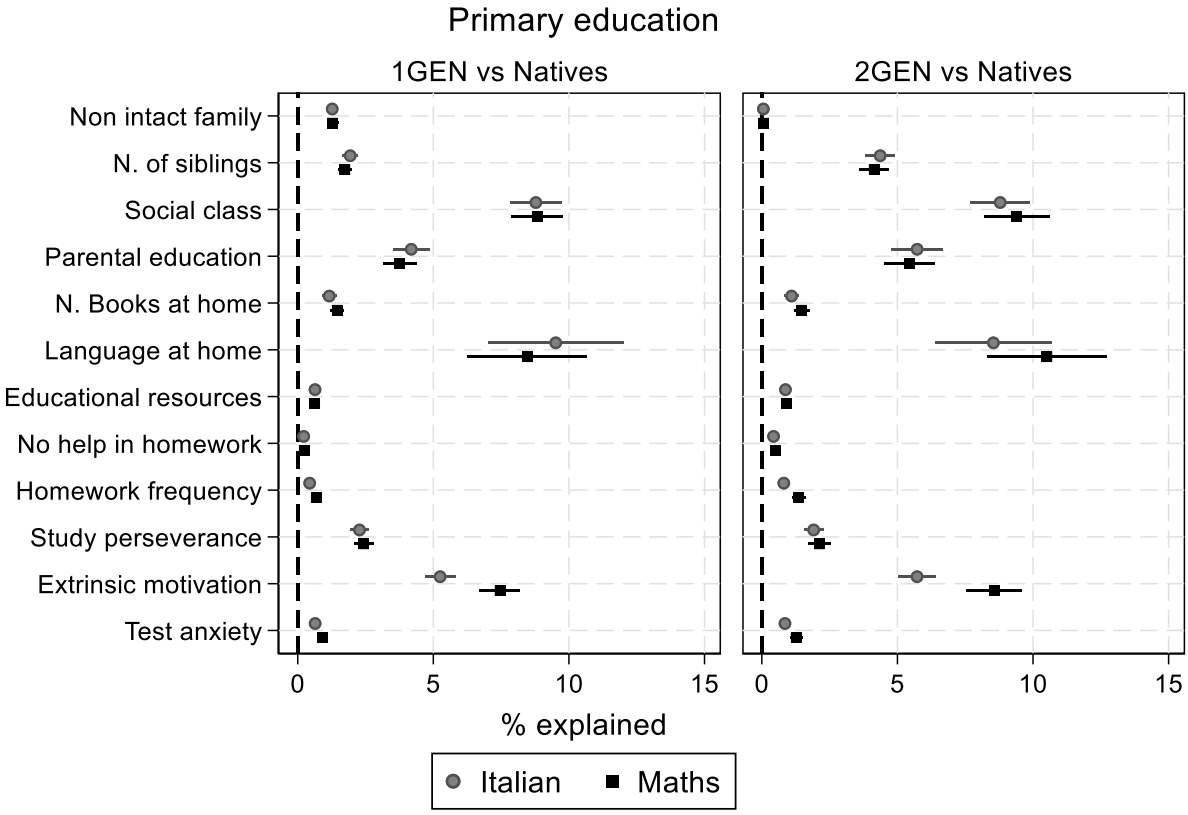

Fig. 3 – Detailed Oaxaca−Blinder decomposition analysis: percentage of the under−evaluation of children of immigrants (first generation, second generation with both parents born abroad) accounted for by a set of students' characteristics in the 5th grade (primary education).

Note: the models control for performance in subject−specific standardized test scores, gender, geographical area, year and quarter of birth, dummy for a random sample of school in which an external observer was present when the test was administered.

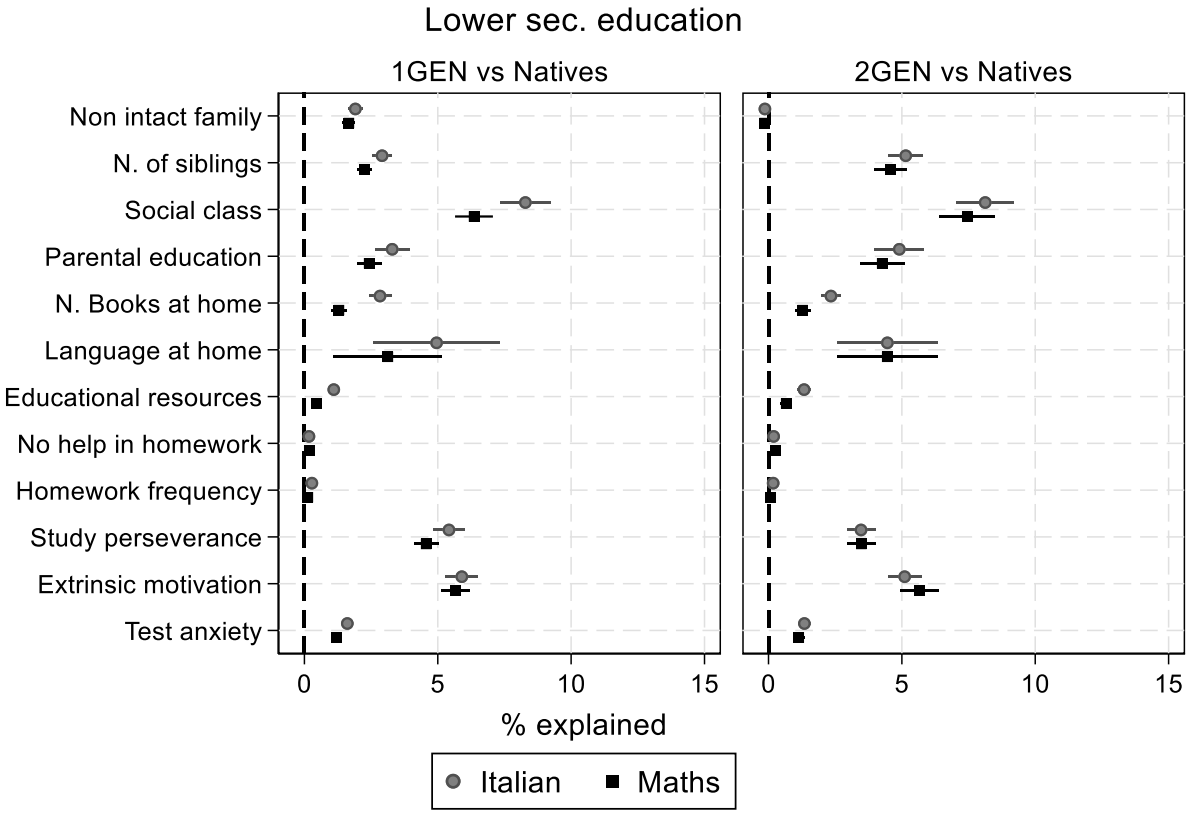

Fig. 4 – Detailed Oaxaca−Blinder decomposition analysis: percentage of the under−evaluation of children of immigrants (first generation, second generation with both parents born abroad) accounted for by a set of students' characteristics in the 6th grade (lower secondary education).

Note: the models control for performance in subject−specific standardized test scores, gender, geographical area, year and quarter of birth, dummy for a random sample of school in which an external observer was present when the test was administered.

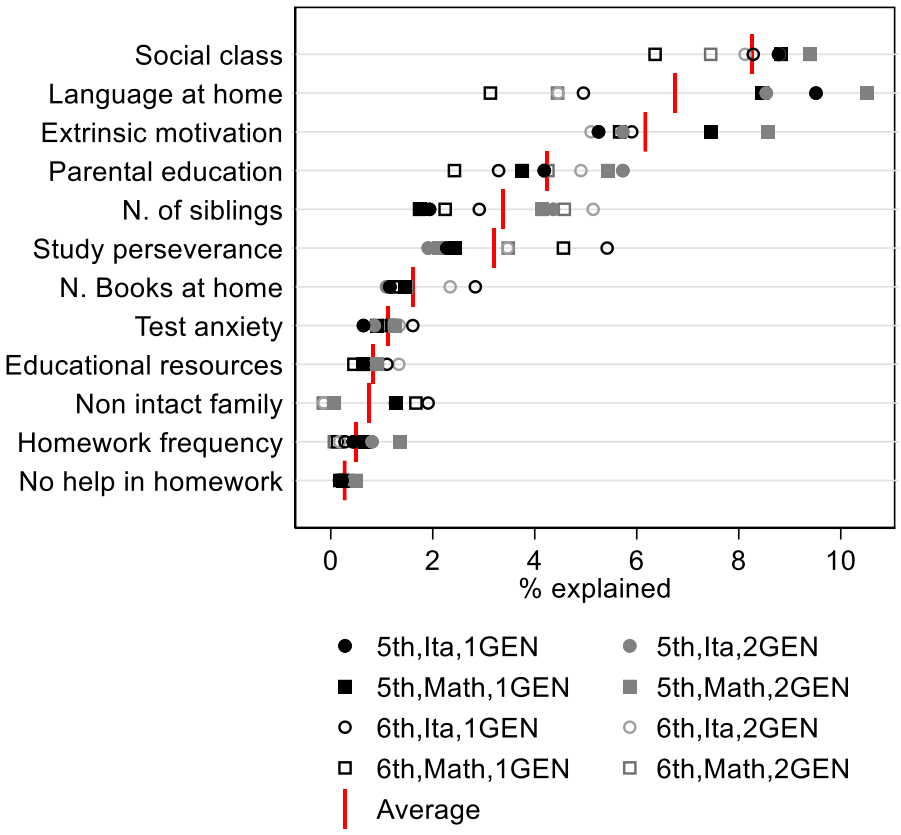

Fig. 5 – Summary of the contribution of each factor in explaining the under-evaluation of children of immigrants across different outcomes and educational levels (presented in Fig. 3 and 4).

Note: full dots refer to 5th grade, hollow dots refer to 6th grades; circles refer to Italian, squares to Mathematics; darker dots refer to first generation students, lighter dots refer to second generation students

Tables

Table 1 – Description of the variables used in the analysis

Variable	Definition
Independent variable	
Migration background	4−categories variable, based on the combination of mother, father and student's place of birth: (1) Natives (students born in Italy with both parents born in Italy); (2) Mixed (students born in Italy from mixed couples); (3) 2GEN (students born in Italy from foreign born parents); (4) 1GEN (foreign born students from mixed couples or foreign born parents). Note: Citizenship variable is considered to include students with missing information on these variables. In some analysis, the variable was recoded in 3 categories (natives, second generation, first generation).
Outcome variables	
Mark in Italian Mark in Mathematics	Average of written and oral mark in mid−term reporting card (February 2012); scale 0−10.
Stratifying variable	
Educational level	Two different educational levels: grade 5th (end of primary education) and grade 6th(beginning of lower secondary education).
Control variables	
Test score in Italian Test score in Mathematics	Scores obtained by students in the INVALSI standardized tests of Italian and Mathematics. The scores are standardized to have a mean of 200 and a standard deviation of 40. We use scores adjusted for potential cheating, divided into quintiles.
Geographical areas	Geographical areas, as defined by the National Institute of Statistics (ISTAT): (1) North West; (2) North East; (3) Centre; (4) South; (4) Isles.
Sex	Dummy variable indicating if the student is male (0) or female (1).
Year of birth	Student's year of birth centered on the modal category.
Quarter of birth	Student's quarter of birth.
Mediators	
Family of origin: Socio−cultural−economic status	
Parents' social class	Parents' social class based on the dominance criterion, considering father and mother's occupation; it has 5 categories: (1) service class; (2) white collars; (3) petty bourgeoisie; (4) working class; (5) other. Note: the residual category of 'other' includes unemployed, retired people, house workers or in any other condition.
Parents' years of education	Average of the years of education attained by the student's father and mother.

Number of books at home	Mean of the number of books at home: (1) 5; (2) 18; (3) 63; (4) 150; (5) 200. Note: this is the average number between the lower and the upper bound of the number of books in the original ordinal variable.
Family of origin: Structure	
Number of siblings	4−categories variable on the number of siblings (from 0 to 4, where 4 means at least four).
Non−intact family	Dummy variable indicating if the student is living with both parents (0) or not (1).
Language proficiency	Dummy variable indicating the language spoken at home: Italian or dialect (0); a foreign language (1).
Educational resources	
Material educational resources	Summary index of material educational resources, derived from a question asking the respondents if they have at home: a quiet place to study; a computer used to study; a desk to do the homework; encyclopedias (composed of books or CDs/DVDs); an internet connection; a room only for themselves.
Educational resources: lack of help at home	Dummy variable derived from a question asking the respondents who helps them the most when they need help in doing their homework: no need or someone helps (0); no help (1).
Student attitudes' and study behavior	
Homework frequency	How often the student to their homework. 5th grade: *medium-high* (at least 3 times a week) vs lower frequency; 6th grade: *medium-high* (at least one hour a day) vs lower.
Perseverance in studying	Additive index derived from a question asking the respondent about their study perseverance. To what extent do you agree with the following sentences (Not at all agree, Partially disagree, Partially agree, Completely agree): i) When an argument is difficult, I let it go; ii) When I study, I commit myself even if the topic is boring.
Extrinsic motivation	Additive index derived from the following question. To what extent do you agree with the following sentences (Very much agree, quite agree, disagree, disagree): i) I study to please my parents; ii) I study to please my teachers; iii) I study to receive presents; iv) I study to not make a bad impression with my classmates.
Anxiety in doing the test	Index derived from the items: *Thinking about the Italian and Mathematics' test you have done, how much do you agree with the following statements* (strongly disagree, somewhat disagree, somewhat agree, strongly agree): a. Already before starting, I was worried about the test; b. I was so nervous that I could not answer; c. While answering, I felt as if I was wrong; d. While answering, I felt calm; e. Mathematics questions were easier than the exercises we usually do; f. Grammar questions were more difficult than the exercises we usually do;

	g. Italian tests were simpler than the ones we read during the year.

Table 2 – Average marks in Italian and mathematics in primary school (5th grade) and lower secondary education (6th grade)

	Primary education (5th grade)		Lower secondary education (6th grade)	
	Italian	Maths	Italian	Maths
Natives	7.7	7.7	6.7	6.7
Mixed	7.6	7.7	6.7	6.7
2GEN	7.0	7.2	6.2	6.3
1GEN	6.8	7.0	5.9	6.0

Table 3 – Descriptive statistics of mediator variables by student migration background

Mediator	**Natives**	**Mixed**	**2GEN**	**1GEN**
	Mean/ Proportion	Mean/ Proportion	Mean/ Proportion	Mean/ Proportion
Non-intact family	0.124	0.158	0.125	0.173
N. siblings	1.181	1.206	1.527	1.360
Social class: Service class	0.286	0.277	0.152	0.148
Social class: White collars	0.281	0.238	0.062	0.056
Social class: Petty Bourgeoisie	0.167	0.190	0.176	0.150
Social class: Working class	0.206	0.227	0.507	0.504
Social class: Other	0.060	0.068	0.103	0.142
Parental years of education	11.4	11.6	10.9	11.0
N. books at home	76.7	77.6	53.9	48.9
Foreign language at home	0.014	0.082	0.504	0.664
Material educational resources (Z)	0.038	0.056	-0.222	-0.194
Educational resources: lack of help at home	0.482	0.458	0.395	0.411
Perseverance in studying (Z)	0.049	-0.012	-0.087	-0.171
Homework frequency: Medium-High	0.821	0.811	0.768	0.778
Extrinsic motivation (Z)	-0.067	-0.021	0.333	0.395
Anxiety in doing the test (Z)	-0.018	-0.04	0.131	0.151

Note: Z=standardized index

Table 4 –Oaxaca-Blinder decomposition analysis: percentage of the native-migrants gap in grading accounted for by all the observed explanatory factors (standard errors in parenthesis)

	Primary education (5h grade)		Lower secondary education (6th grade)	
	Mathematics	Italian	Mathematics	Italian
	% (SE)	% (SE)	% (SE)	% (SE)
1GEN vs Natives	37.9	36.3	29.3	38.7
	(2.0)	(2.0)	(1.6)	(2.1)
2GEN vs Natives				
	45.7	39.2	33.1	36.4
	(2.8)	(2.3)	(2.1)	(2.2)

Note: the models control for performance in subject−specific standardized test, gender, geographical area, year and quarter of birth, dummy for a random sample of school in which an external observer was present when the test was administered. The explanatory factors included to account for the gaps are the 'Mediators' listed in Table 1.

Appendix

Table A1 – Results from OLS regression and Classroom Fixed-Effects linear models (CFE) to analyze teachers' grading (*** p<.001)

	Mathematics		Italian	
	OLS	CFE	OLS	CFE
Primary education (5th grade)				
Migration background (ref.: Natives)				
Mixed	−0.048***	−0.064***	−0.070***	−0.072***
2GEN	−0.207***	−0.220***	−0.235***	−0.241***
1GEN	−0.316***	−0.309***	−0.339***	−0.334***
Lower secondary education (6th grade)				
Migration background (ref.: Natives)				
Mixed	−0.055***	−0.067***	−0.051***	−0.064***
2GEN	−0.216***	−0.214***	−0.194***	−0.183***
1GEN	−0.289***	−0.275***	−0.239***	−0.212***

Note: all the models include basic control variables and mediator variables described in Table 1.

Table A2 - Estimates presented in figures 3 and 4

	Primary education (5th grade)				Lower secondary educ. (6th grade)			
	Italian		Maths		Italian		Maths	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.
1GEN vs Natives								
Non intact family	1.3	0.1	1.3	0.1	1.9	0.1	1.7	0.1
N. of siblings	1.9	0.1	1.7	0.1	2.9	0.2	2.2	0.1
Social class	8.8	0.5	8.8	0.5	8.3	0.5	6.4	0.4
Parental education	4.2	0.3	3.8	0.3	3.3	0.3	2.4	0.2
N. Books at home	1.2	0.1	1.5	0.1	2.8	0.2	1.3	0.1
Language at home	9.5	1.3	8.5	1.1	5.0	1.2	3.1	1.0
Educational resources	0.6	0.1	0.6	0.1	1.1	0.1	0.5	0.1
No help in homework	0.2	0.1	0.2	0.1	0.2	0.0	0.2	0.0
Homework frequency	0.4	0.1	0.7	0.1	0.3	0.0	0.1	0.0
Study perseverance	2.3	0.2	2.4	0.2	5.4	0.3	4.6	0.2
Extrinsic motivation	5.3	0.3	7.5	0.4	5.9	0.3	5.7	0.3
Test anxiety	0.6	0.1	0.9	0.1	1.6	0.1	1.2	0.1
2GEN vs Natives								
Non intact family	0.1	0.1	0.1	0.1	-0.1	0.1	-0.1	0.1
N. of siblings	4.4	0.3	4.1	0.3	5.1	0.3	4.6	0.3
Social class	8.8	0.6	9.4	0.6	8.1	0.5	7.4	0.5
Parental education	5.7	0.5	5.4	0.5	4.9	0.5	4.3	0.4
N. Books at home	1.1	0.1	1.5	0.1	2.3	0.2	1.3	0.2
Language at home	8.5	1.1	10.5	1.1	4.5	1.0	4.4	1.0
Educational resources	0.9	0.1	0.9	0.1	1.3	0.1	0.7	0.1
No help in homework	0.4	0.1	0.5	0.1	0.2	0.0	0.3	0.1
Homework frequency	0.8	0.1	1.4	0.1	0.2	0.0	0.1	0.0
Study perseverance	1.9	0.2	2.1	0.2	3.5	0.3	3.5	0.3

Extrinsic motivation	5.7	0.4	8.6	0.5	5.1	0.3	5.7	0.4
Test anxiety	0.9	0.1	1.3	0.1	1.3	0.1	1.1	0.1