Lost in Translation. Reading Performance and Math Performance of Second-Generation Immigrant Children in Italy

Mariagrazia Cavallo[1] and Giuseppe Russo[2]

[1]University of Bristol

[2]University of Salerno, CSEF, GLO

Abstract

We study the effect of language proficiency on Math achievement for 10-year-old second-generation immigrant children in Italian primary schools. Proficiency in the host country's language is the prerequisite to acquire any other skill. However, using an instrumental variable strategy that exploits heterogeneity in birth dates and variation in linguistic distances, we find that these children face a trade-off between learning Italian and learning Math. According to results in the linguistic literature, we show that the trade-off occurs when proficiency falls below the threshold ensuring a sufficient command of Italian. We also develop a model of skill production providing a theoretical foundation for our results.

JEL classification: I21, I24, Z13.

Keywords: Second-Generation Immigrants, Language Proficiency, Math Performance, Linguistic Distance, Threshold Effects.

1 Introduction

The adaptation of immigrants to a receiving society is an intergenerational process that involves many socioeconomic dimensions (Constant et al., 2009; Constant and Zimmermann, 2008). Throughout this progression, however, the main prerequisite for integration in *all* dimensions is language acquisition, suggesting that language is the most important form of human capital for immigrants (Chiswick and Miller, 2015; Ginsburgh and Weber, 2020).[1] In the case of education, language acquisition is not only an outcome but also a prerequisite for the acquisition of further skills (Isphording et al., 2016). Possibly the best example of this essential role as a prerequisite is in learning Math. While the view that language proficiency is weakly or not at all related to Math might be common, it is now widely accepted that Math achievement largely depends on oral communication, and it cannot be considered a non-verbal subject (Wilkinson, 2019).

The objective of this paper is to identify the causal effect of language acquisition on Math achievement for 10-year-old second-generation immigrant children.[2]

Since an immigrant background hinders school performance (see Table 1; Abatemarco et al., 2022; OCDE, 2015), language skills are even more critical to the education of these children. For second-generation children, in particular when their parents speak their mother tongue at home, Italian can be considered as a vehicular language. Italian language is the vehicle through which students learn and apply Math, and it is used to test Math skills.[3] As in other classes, the ability to interact in the classroom, ask questions, and express doubts is crucial to being able to benefit from Math classes. This is supported by recent findings indicating that verbal skills impact the acquisition of Math skills (Aucejo and James, 2021). Thus, it is essential to identify this effect in order to understand the formation of the second generation's human capital, and

[1] The literature provides extensive evidence that language proficiency affects various outcomes, including labor market earnings (Bleakley and Chin, 2004; Brell et al., 2020; Chiswick and Miller, 2010; Dustmann and Fabbri, 2003), occupational sorting and the choice of college major (Bacolod and Rangel, 2017), health and health insurance coverage (Clarke and Isphording, 2017; Dillender, 2017), and fertility decisions (Aoki and Santiago, 2018), among various other socioeconomic outcomes (Guven and Islam, 2015).

[2] To simplify the notation, this paper uses the term "second generation"to refer to "second-generation immigrant(s)". We define the second generation as children born in Italy to parents who are *both* non-Italian. In contrast, we define natives as children born in Italy with parents who are both Italian.

[3] In many countries, special Math textbooks that use simplified terminology are made available to immigrant children. In Italy, an example is Arici and Maniotti (2010).

thereby prevent early educational gaps from persisting well into adulthood (Almond et al., 2018).

However, estimating the extent to which language proficiency determines Math achievement is subject to fundamental identification challenges. Math and language scores are both driven by unobservable variables, like ability and motivation. In the case of immigrant children, other unobservable mechanisms —such as family self-selection along dimensions that are relevant for school performance— may be involved, further complicating causal estimations.

The difficulty of identification may be one reason why the literature remains narrow, consisting of two main contributions; namely, Isphording et al. (2016) and Aparicio-Fenoll (2018). Both papers use instrumental variables for language proficiency. The former exploits the interaction between age at arrival and a continuous measure of linguistic distance. The latter uses the interaction between age at arrival with a dummy variable for non English-speaking country of origin. Notably, these instruments cannot be used for second-generation immigrants, who were born in the host country. To overcome this issue, we use an instrumental variable (IV) strategy that exploits heterogeneity in birthdates and variation in the linguistic origins of the second generation. We construct an instrument that combines these two sources of variation, namely, the interaction between the linguistic distance from Italian and the child's age. Age captures the length of exposure to Italian language and society. Linguistic distance captures the difficulty of adaptation with respect to the language *spoken at home*, which our data includes. The coefficient of the interaction between these variables measures how the effect of linguistic distance on the score in Italian changes as exposure to Italian changes.[4] Using only the interaction term as our instrument allows us to control for both age and linguistic distance themselves, avoiding confounding direct effects on the Math score. This approach is analogous to the strategies of Bleakley and Chin (2004), Isphording et al. (2016), and (albeit within a different framework) Karadja and Prawitz (2019).[5] Notably, a series of placebo outcome tests, a zero-first-stage test and a Hansen J test for overidentification suggest no evidence of endogeneity

[4]Alternatively, it measures how the effect of exposure to Italian on the score in Italian changes as linguistic distance changes.

[5]Bleakley and Chin (2004) use the interaction between age at arrival and non-English origin. Karadja and Prawitz (2019) use the interaction between frost shocks and the distance to the nearest emigration port as an IV for emigration rates.

associated with our instrument.

Results in the literature are mixed. On the one hand, Isphording et al. (2016) confirm that language proficiency positively affects Math performance. Using PISA data on boys aged 15-16, they find that a one-standard-deviation increase in reading performance improves performance in Math by 0.57 standard deviations. On the other hand, Aparicio-Fenoll (2018) studies a sample of 1,529 immigrant children aged 6-12 in the United States and finds no effect of language proficiency on Math proficiency.[6]

We increase the disagreement in the literature, since we find that proficiency in Italian wields a *negative* effect on Math performance. According to our estimates, a one-standard-deviation increase in an individual's Italian score decreases, on average, the score in Math by 0.335 standard deviations, which is equivalent to about 5.7% of the gap in Math with native peers.[7] Thus, our results point to a trade-off between learning Italian and learning Math.

However, it is hard to accept this variety of results as proof that fluency in the host country language is useless, or even detrimental, in helping immigrant children to understand Math classes. Rather, it is conceivable that the puzzling evidence emerging from the research indicates non-linearities in the link between language proficiency and the ability to learn Math.

We believe that the linguistic literature offers a solution to this puzzle. In this field, it is well-known that there exist minimum level thresholds to be crossed for proficiency gains in language acquisition to be effective (Cummins and Gulutsan, 1974; Toukomaa and Skutnabb-Kangas, 1977). For instance, it is commonly assumed that a sufficient proficiency is required *to gain access to the curriculum*.[8] This corresponds to the universal practice of establishing *minimum* language requirements for foreign students, *even at the graduate level*.[9]

Hence, research on the effect of language proficiency on the acquisition of Math and other skills lacks a connection with the linguistic literature. To fill this gap, we develop

[6] See section 4.1 for a more detailed discussion of Aparicio-Fenoll (2018) and Isphording et al. (2016).

[7] Our results are not directly comparable to the above-mentioned papers, because they focus, respectively, on first-generation students aged 15-16 and a mix of first and second-generation children aged 6–12.

[8] See Cummins (2000) for a survey of the issues related to what is known as either the "threshold hypothesis"or the "Cummins hypothesis".

[9] For instance, Erasmus students are usually required to have a B2 English level certificate on the Common European Framework of Reference (CEFR). In the case of graduate admissions, a C1 level is even more common.

a simple theoretical model of skill production taking into account a sufficiency threshold in language acquisition. We provide a structure which unifies the different findings described above under a unique framework, proving that they do *not* contradict the idea that language proficiency is a prerequisite for the acquisition of Math skills. Our model is based on two building blocks: 1) language proficiency is a prerequisite to learning Math; 2) however, before the benefits of language proficiency can be reaped, a threshold must be crossed. The first building block uses language as an *intermediate input* for producing Math skills. The second building block embodies the threshold via a Stone-Geary-like production function of Math skills. We solve the problem of a child who obtains utility from school grades and must allocate her study time between Italian and Math. There are two types of children: those who are inside the threshold and those who are on or above it. In the first case, at equilibrium, further increases to the Italian score imply decreases to the Math score (i.e., a trade-off). In the second case, this trade-off can be avoided, with children who know Italian very well proving capable of simultaneously improving their grade in Italian and Math.

The main implication of our theoretical model is therefore that the effect of proficiency in Italian on the score in Math can be quite different for children inside and above the sufficiency threshold. We develop our empirical analysis accordingly, by considering these two groups separately. Our estimates suggest that the negative effect observed for the whole sample is led by children below the sufficiency threshold.[10] For above-sufficiency children, our instrument is weak: in all likelihood, once a sufficient command of Italian is achieved, further exposure to the language does not significantly increase the *academic* score in Italian. In addition, since children speaking a Romance language are 68.4% of those above sufficiency, and 56.2% of those below sufficiency, it is easy to hypothesize that a Romance linguistic background makes it easier to achieve sufficiency in Italian. Thus, we also carry out an indirect check of the threshold hypothesis by performing a heterogeneity analysis between children from Romance and non-Romance linguistic backgrounds. While we observe a trade-off for non-Romance

[10]On the scale that evaluates proficiency on a range from Level 1 (lowest) to 5 (highest), we use Proficiency Level 3, which is commonly considered indicative of a sufficient command of the language. Level 3 is defined by scores in the range of $(95\%; 110\%]$ of the average obtained by native students. The Italian National Institute for the Evaluation of the Education System complies with this definition (see the National Educational Criteria (*Indicazioni Nazionali e Linee Guida*) stated in the ministerial decree *D.M. n. 254-2012*, and INVALSI, 2018).

children, our instrument is weak for Romance ones. This outcome reflects what we observe for below- and above-sufficiency children.

We contribute to the literature in many respects. First, we show that marginal investments in education are insufficient to help the most disadvantaged children achieve a sufficient command of the language. This leads them to underperform in other fundamental subjects such as Math. Second, we shed further light on the mechanisms that govern the relationship between language proficiency and the acquisition of mathematical ability. Third, we propose a simple IV strategy that can be used to analyze the effect of linguistic proficiency on the Math performance of *second-generation* immigrant children.

Our findings are particularly worrying. Second-generation children, demonstrating poor linguistic performance, continue to struggle to become proficient in language at the age of 10, and can do so only by performing more poorly in other subjects. However, the structure of the curricula in terms of subjects taught and examinations given depends upon the assumption that children master Italian by the end of primary school. This could further amplify the educational disadvantage experienced by the second generation, which fosters social exclusion and inequality. Thus, in the absence of appropriate interventions, the future integration of the second generation seems to be jeopardized by the time they are 10 years old. This may cast some doubts on the capacity of the Italian school system to promote equality of opportunity. [11]

Generally speaking, our findings indicate that destination countries should endeavor to make linguistic integration effective during the very first years of a second generation child's education, where "effective" means that all children should *cross the sufficiency threshold*. Note that the existence of a threshold renders marginal actions inappropriate because marginal improvements in proficiency are hardly sufficient for children who lag behind. It is also likely that having a home environment where the opportunities to practice Italian are limited increases the difficulties faced by these children. Thus, to the extent that being more fluent incentivizes the use of Italian at home, investing in the linguistic integration of the first generation could benefit the second generation as well.

[11] Children below the sufficiency threshold accounted for 61.06% of the second generation in the school year 2014–15, 57.32% in the school year 2015–16, 61.44% in the school year 2016–17, 55.04% in the school year 2017–18, and 59.64% in the school year 2018–19.

The rest of this paper is organized as follows. Section 2 describes our data. Section 3 introduces our empirical strategy and discusses the validity of our instrument. Section 4 presents our IV findings. Section 5 provides a theoretical framework for interpreting our results and the underlying skills production mechanisms. Section 6 reports further empirical evidence in line with the predictions of our theoretical model, and Section 7 concludes the paper.

2 Data

This paper uses data for proficiency in Italian and Math from the standardized test administered by the Italian National Institute for the Evaluation of the Education System (INVALSI). We consider the entire population of second-generation immigrant students at the end of primary school, namely, in the fifth grade.[12] We exploit a repeated cross-section for the school years 2014–15, 2015–16, 2016–17, 2017–18, and 2018–19 that involves observing 136,613 second-generation immigrant students.[13] The INVALSI tests are standardized, anonymous, and marked externally. In principle, the results of these tests should only be used for assessing basic requirements and monitoring regional differences, without contributing to students' final grades. However, in practice, matters differ substantially. Upon their introduction in 2009, the tests were intended as a self-assessment device for internal use; that is, they were not supposed to be used to create any sort of school ranking. Yet, better performing schools quickly decided to disclose their results, to signal the quality of their teaching. Thereafter, since 2011, the Ministry of Education has maintained a website (*Scuola in Chiaro*) providing information about INVALSI results alongside general information about schools. This has made school principals and teachers alike concerned about the impact of bad results on their reputations and careers. In fact, some proposals have been advanced that advocate rewarding teachers whose students have performed

[12] See footnote 4 for the definition of second generation.

[13] Although the INVALSI tests were introduced in 2009, the question that identifies the linguistic origin of pupils was only introduced in 2014. Note that we exclude second-generation immigrant children who speak languages not included in the list. In order to avoid errors in data, we also exclude all children who are at least two years younger or older than 10. In Italy, it is not possible to start school before 5, and all children *must* finish primary school by the age of 12.

better.[14] As such, many teachers now use INVALSI results to assign grades, prompting the emergence of a market for test preparation textbooks. The ever-increasing importance of the tests has brought about attempts at cheating that several studies have documented (Angrist et al., 2017; Bertoni et al., 2013; Lucifora and Tonello, 2015, 2020; Paccagnella and Sestito, 2014; Pereda-Fernández, 2019). Accordingly, scores are adjusted to account for cheating.[15]

The test is structured in two main sections, which assess proficiency in Italian and proficiency in Math. The final scores range from 0 to 100 points and express the percentage of correct answers. The Math section includes questions about geometry, mappings, and data analysis. The Italian test focuses on reading comprehension and the grammatical and lexical structure of sentences. This language proficiency test is much more reliable than self-reported assessments, and avoids issues associated with non-classical measurement errors.

Additional information is derived from the “Student’s Questionnaire”, a questionnaire administered to fifth-grade students on the same date as the Math test. Information regarding parental education, parental occupational status, and home education resources is consolidated in a synthetic index of economic, social, and cultural status (the ESCS index) that is constructed and provided by INVALSI.[16] Other useful information derives from questions about the students themselves, their family, and their attitudes toward their classes and the test. In particular, pupils are asked whether they speak Italian or another language at home.[17] The distance of these languages from Italian can be computed in different ways, based on differences in vocabulary, phonetics, grammar and syntax (see Ginsburgh and Weber, 2020). Following Isphording et al.

[14] On May 5, 2014, the Minister of Education, Stefania Giannini, claimed, “we have to work in order to make evaluation the key to radical reform of the teacher’s status [...] and reward those who work harder.” Generally speaking, the INVALSI tests represent the most controversial innovation in Italian schools, and their role remains heavily contested. See (Agnelli, 2014; Dell’Anna, 2021) for useful surveys.

[15] Cheating describes a broad concept that denotes any attempt to alter results by either students and teachers (Jacob and Levitt, 2003). INVALSI has adopted a fuzzy clustering technique to adjust final test scores and improve the accuracy of outcomes. The algorithm considers four main indicators to detect cheating, namely, the share of correct answers, the share of missing answers, and the variability and homogeneity of response patterns (Lucifora and Tonello, 2020; Quintano et al., 2009).

[16] This index is a composite score, normalized to obtain zero-mean and unit-variance, computed from a principal component analysis that accounts for three indicators: parental education (PARED), parental occupational status (HISEI), and educational resources available at home (HOMEPOS). The latter category considers the availability of a desk and a quiet place to study, the number of books available at home, and the availability of a computer, an internet connection, and an encyclopedia. A similar index is also provided in the Program for International Student Assessment (PISA).

[17] Namely, Albanian, Arabic, Chinese, Croatian, French, Greek, Hindi, English, Ladin, Portuguese, Romanian, Slovenian, Spanish, German, or another language.

(2016) and Clarke and Isphording (2017), we use continuous measures rather than dummies to better control for the heterogeneity of linguistic origins.

We use four measures of linguistic distance from Italian (henceforth "linguistic distance"). Our preferred measure is provided by the Automated Similarity Judgment Program (ASJP) database, which is commonly used for linguistic analyses.[18] The ASJP distance is a lexicostatistical measure, based on the phonetic similarity between words. In addition to the ASJP distance, we use another measure of lexicostatistical distance provided by `elinguistics.net`.[19] For robustness, we also use two up-to-date measures of cladistic distance we obtained from the recent work by Cole et al. (2022). Cladistic distances account for various aspects of languages, including lexicon, but also syntax, phonology, and grammar. Following Laitin (2000), we compute the first measure of distance by counting the branches the two languages share on the language tree by Cole et al. (2022) before they break off from each other. Fearon (2003) notices that with this method earlier break-offs generate a higher dissimilarity than later break-offs. Consequently, he suggests using the square root of the Laitin measure to correct for this effect. Our second measure of cladistic distance is computed similarly.

A major advantage of our analysis is that linguistic distance applies to the language *actually* spoken at home. Although many authors have inferred this language from information about the country of origin, associating a country with a single language is not always possible because different languages can be spoken in the same country. This is the case for many North African countries, which are a major source of immigration to Italy.[20] Note that, even when such associations are possible, this does not

[18]See Wichmann, Holman, and Brown (eds.), 2018. The ASJP Database (version 18). This measure of linguistic distance is built by comparing the inner structure of 40 words in all the world's languages and provides a continuous measure that, in this context, ranges from a minimum distance of 58.77 (Romanian-Italian) to a maximum distance of 101.14 (Chinese-Italian). It compares the phonetic similarity between pairs of words with the same meaning in the two languages. This should capture the existence of common ancestries between languages that can affect the ease of learning Italian (Isphording and Otten, 2013).

[19]This measure focuses on the genetic proximity between languages, with an emphasis on their sound correspondence. Recent research in linguistics suggests that links between a language's sound system, the climate, and the geographic conditions of the place where it is spoken can exist (Everett et al., 2015, 2016). Consequently, this measure, which is also used by authors like Galloway and Gjefsen (2020), possibly controls for spatial and geographic factors. In our analysis, the results are indistinguishable with respect to the other measure of lexicostatistical distance, and are not reported for the sake of brevity. They are available upon request.

[20]For instance, immigrants from Morocco can speak Arabic, French, or Spanish; immigrants from Egypt speak Arabic, French, or English; immigrants from Tunisia speak Arabic or French, and immigrants from India speak Hindi or English. There are sizable immigrant populations from these countries in Italy: as of January 1, 2021, there are 428,947 immigrants from Morocco, 139,569 from Egypt, 97,407 from Tunisia, and 165,512 from India. Source: ISTAT (2022).

imply that the parents speak their mother tongue with their children in the destination country.

To investigate the validity of our IV strategy, Section 3 considers questions that are only asked in the Student's Questionnaire (2014–15), which includes questions about teacher attitudes, parental incentives, preferences for Math and Italian, socio-emotional skills, and whether the child has been bullied by peers. Some descriptive statistics for the main variables are summarized in Table 1.

Table 1: Descriptive Statistics. Natives vs. Second generation

Variable	**Natives**	**Second generation**	**Diff.**
Math	58.274	51.249	7.025***
	(0.050)	(0.181)	(0.188)
Italian	63.361	54.514	8.847***
	(0.044)	(0.160)	(0.166)
Female	0.495	0.496	-0.001
	(0.000)	(0.001)	(0.002)
Age in Months	129.392	130.354	-0.962***
	(0.008)	(0.027)	(0.028)
ESCS student	0.142	-0.501	0.643***
	(0.003)	(0.011)	(0.012)
Obs	1,704,750	136,613	

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$
Standard Errors in parenthesis clustered at the school-cohort level.

Natives perform better than the second generation in both Math and Italian. On average, the gap between the second generation and natives is 8.8 points in Italian and 7.0 in Math.

Observing the progression of the second generation's performance over time is of interest as it enables a rough assessment of the relative importance of linguistic distance compared to the general convergence patterns associated with age. In particular, if the effect of linguistic distance fades, the school successfully acts as an equalizer. To check this possibility, it would be worth observing the *initial* (namely, pre-school) linguistic distance gradient. In the absence of pre-school data, we exploit test scores administered in the second grade as a proxy for the initial gradient. Unfortunately, the second-grade data do not include the language spoken at home. We can retrieve this information by matching second- and fifth-grade students, but we lose about a third of the observations due to the absence of a link identifier. Additionally, the subsample of

Table 2: Second Grade

Variable	Math 2 (1)	Italian 2 (2)
Age in Months	0.533***	0.409***
	(0.018)	(0.019)
Distance	-0.030***	-0.044***
	(0.002)	(0.002)
Controls	✓	✓
Cohort FE	✓	✓
Observations	71,728	71,728

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$
Robust standard errors in parentheses
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index

matched students shows evidence of sample selection.[21] Thus, the comparison over time reported below can be considered an educated guess regarding the development from second to fifth grade. More precisely, to roughly assess the importance of linguistic distance and age over time, we estimate the following regression in the second and the fifth grade:

$$Score_{isct}^{it,mat} = \beta_0 + \beta_1 Dist_{isct} + \beta_2 Age_{isct} + \mathbf{X}\lambda + \vartheta_t + \xi_{isct} \tag{1}$$

where the dependent variable is the score in Italian (it) and Math (mat) of student i, in school s, class *c*, cohort *t*; *Dist* is the linguistic distance between the language spoken at home and Italian; *Age* is the student's age (measured in months), $\mathbf{X}$ is a vector of individual characteristics and socioeconomic controls that includes dummies for gender, previous enrollment in nursery school (age 0–3), previous enrollment in pre-school (age 3–6), and the ESCS index of socioeconomic status. ϑ_t are cohort fixed effects.

Including *Age* in the regression helps us to control for the effect of exposure to the Italian language. The results are reported in Tables 2 and 3.

The coefficient for *Dist* is negative and significant at the 1% level in all regressions. Somewhat unexpectedly, it is remarkably stable over time for both Math and Italian.

[21] In particular, a series of tests on the equality of means suggest that matched students have better socioeconomic conditions, are linguistically closer, older, and obtain better scores. This explains the differences between Table 1 and Table 4.

Table 3: Fifth Grade

Variable	Math 5 (1)	Italian 5 (2)
Age in Months	0.285***	0.231***
	(0.018)	(0.018)
Distance	-0.025***	-0.045***
	(0.002)	(0.002)
Controls	✓	✓
Cohort FE	✓	✓
Observations	71,728	71,728

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Robust standard errors in parentheses

Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index

Table 4: Descriptive Statistics. Natives vs. Second generation

Variable	**Natives**	**Second generation**	**Diff.**
Math2	58.557	52.003	6.555***
	(0.018)	(0.054)	(0.056)
Math5	57.903	52.038	5.865***
	(0.017)	(0.053)	(0.055)
Italian2	61.436	53.885	7.551***
	(0.018)	(0.055)	(0.059)
Italian5	63.967	57.795	6.172***
	(0.016)	(0.052)	(0.052)
Obs	1,231,410	132,798	

Standard Errors in parenthesis.

*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Apparently, schools are unsuccessful in reducing the effect of linguistic barriers. This aligns with the outcome of the IV analysis in Section 4. In contrast, the coefficient for *Age* is positive and 1% significant both in Italian and Math, although its magnitude shrinks from the second to the fifth grade. Overall, because the effect of *Dist* is constant and the effect of *Age* is positive, we can expect a slight increase in the score for both Math and Italian, which we indeed observe. The average score in Math changes from 52 (second grade) to 52.034 (fifth grade). In Italian, it changes from 53.885 (second grade) to 57.795 (fifth grade). The average Math gap compared to natives is 6.6 points in the second grade, and 5.866 points in the fifth grade.[22] The gap in Italian is 7.455 points in the second grade and 6.172 points in the fifth grade. Overall, these figures indicate, at most, some weak convergence (mainly in Italian) between natives and the second generation. While the reduced gap is essentially due to age, the effect of linguistic distance is seemingly unchanged.

3 Empirical Strategy

As a first approach to evaluating the relationship between Math proficiency and Italian proficiency, we estimate a standard regression:

$$Math_{isct} = \beta_0 + \beta_1 Italian_{isct} + \beta_2 Dist_{isct} + \beta_3 Age_{isct} + \mathbf{X}\lambda + \vartheta_{sct} + \xi_{isct} \tag{2}$$

where the dependent variable is the score in Math of student i, in school s, class c, cohort t; *Italian* is the score in Italian; *Dist* is the linguistic distance between the language spoken at home and Italian; *Age* is the student's age (measured in months); $\mathbf{X}$ is the vector of individual characteristics and socioeconomic controls used in equation (1). In some specifications, $\mathbf{X}$ also includes dummies for parents' area of origin; namely, EU, non-EU Europe, rest of the world. This helps to control for the heterogeneity of parents coming from countries where the push factor can be assumed to be more important than the pull factor. This check matters because the prevalence of push factors over pull factors (and vice versa) may induce a different selection of immigrants. To the extent that this selection concerns parental characteristics that affect Math

[22] Note that the difference in Math from second to fifth grade is fully accounted for by a worse performance by natives rather than by a better performance by the second generation.

learning in the second generation, discerning when one factor dominates the other is relevant.[23] Finally, ϑ_{sct} are school-by-class-by-cohort fixed effects, which allow us to control for unobserved differences within school-class-cohorts;[24] and ϵ_{isct} is the error term, clustered at the school-class-cohort level to allow for arbitrary correlation between children in each school-class-cohort.[25]

This specification suffers from an omitted variable bias because many crucial factors responsible for the results for both Italian and Math (such as abilities and motivation) cannot be observed. Therefore, to estimate the effect of linguistic proficiency on Math, we use an IV approach.

The IV method is common to Isphording et al. (2016) and Aparicio-Fenoll (2018), who also estimated the effect of immigrant language skills on Math results.[26] In accordance with Bleakley and Chin (2004, 2008, 2010), these authors rely on a quasi-experimental framework that involves comparing immigrants of different ages at arrival and different linguistic origins.[27] However, the features intrinsic to this construction make it impossible to use their IVs for the *second* generation, since the age-at-arrival effect cannot be exploited for children born in the destination country.[28]

[23]Our data do not include the parent's origin country; however,children are asked whether their parents come from the EU, non-EU Europe, or the rest of the world. Since immigration from non-European countries richer than Italy is in practice negligible (see figure 1 in Appendix A), this question approximately captures the preponderance of the push factor. In fact, Western and Eastern Europe roughly coincide with EU and non-EU European countries, the rich rest of the world is absent, the poor rest of the world includes Africa and the Middle East, where push factors are quite more serious. Regressions including these controls (available upon request) do not alter our results. Notice also that we cannot obtain further dummies combining —for instance— information on the language spoken at home with macro-areas, because this way we would exclude all the children who speak Italian at home, causing sample selection.

[24]This is equivalent to a within-group transformation, where we subtract the mean of the school-class-cohort from each variable in the model. We also estimate variations of this model using separate class or school and cohort fixed effects (namely, two-way fixed effects), assuming that cohort effects are not nested within school-classes, schools, or classes. We also estimate specifications with only cohort fixed effects to understand the overall effect (beyond the within-school and within-class effects).

[25]This level of clusterization, which accounts for the interaction effects (namely, attending the same school the same year), is equivalent to a one-way clustering at the school-class-cohort level. We estimate variations of this model with different levels of clusterization, such as at the school, language-wave, and school-cohort level. Finally, we estimate a three-way clustering model at the school, class and cohort level, which accounts for correlation along the three dimensions, namely, within schools and classes, across waves, and also within waves, across schools and classes. Our results are not affected by the type and level of clusterization and are robust across the different specifications.

[26]Isphording et al. (2016) use PISA test scores for reading and Math at the age of 15. Aparicio-Fenoll (2018) uses New Immigrant Survey (NIS) data concerning US immigrants, which include the results of various cognitive tests measuring reading skills, problem solving, and calculus ability. Proficiency in English was assessed by the interviewers using this scale: "very bad", "bad", "good", and "very good".

[27]Their instruments are the interaction between immigrant age at arrival and the linguistic distance from English, or a dummy for non-English speaking countries.

[28]Bleakley and Chin (2008) study the transmission of language skills from first- to second-generation immigrants, using an IV strategy where second-generation immigrant outcomes are related to *their parents*' ages at arrival.

Consequently, we develop an alternative IV for which, as we argue later, we propose the interaction between age and linguistic distance from Italian.
Accordingly, we estimate a two-stage least squares model in which the first stage is given by

$$Italian_{isct} = \alpha_0 + \alpha_1 Age_{isct} * Dist_{isct} + \alpha_2 Dist_{isct} + \alpha_3 Age_{isct} + \mathbf{X}\rho + \varphi_{sct} + \eta_{isct} \tag{3}$$

and the second stage is given by

$$Math_{isct} = \gamma_0 + \gamma_1 It\hat{al}ian_{isct} + \gamma_2 Dist_{isct} + \gamma_3 Age_{isct} + \mathbf{X}\delta + \theta_{sct} + \epsilon_{isct} \tag{4}$$

where $It\hat{al}ian_{isct}$ is the predicted score in Italian from the first stage.
Our instrument exploits the heterogeneity of birth dates and the variation in linguistic distance. Note that neither age nor linguistic distance alone constitutes a valid instrument.

For instance, although proficiency in Italian improves with longer exposure to the language, an "older" age gives the child more time to also learn Math, producing a violation of the exclusion restriction. Age could also capture either grade retention or various forms of parental help, or both of these. While grade retention in Italian primary school is forbidden,[29] we cannot rule out parental help, which we discuss in Appendix B. Finally, evidence exists for seasonality-in-fertility effects (Buckles and Hungerman, 2013), which we control for by adding month fixed effects in some specifications. All of these issues make age unsuitable as an IV.
Linguistic distance is also inappropriate, because different linguistic origins may induce a different selection in the migration decision (e.g., because a greater distance increases the cost of language acquisition or because of cultural differences). This kind of selection depends on unobservable characteristics that could be transmitted to the second generation. Accordingly, we use the *interaction* between age and linguistic distance as our instrument. An advantage of using this interaction as our instrument is that we can control for the direct effects of the two variables, as the literature often

[29] More precisely, grade retention *can* only be authorized if the child misses at least 25% of the classes. Retention based on school performance is *not* allowed. Primary school retention rates for the years 2015–2019 are, respectively, 0.00036, 0.00036, 0.0003, and 0.00025. Source: Italian Institute of Statistics (ISTAT), *Istruzione e Formazione -Scuole -Principali dati*, available on www.istat.it.

does (Andersson et al., 2022; Bleakley and Chin, 2004; Isphording et al., 2016; Karadja and Prawitz, 2019). In our case, this interaction captures exposure to Italian *across different languages*, or, alternatively, the effect of linguistic distance at different levels of exposure to Italian.

The key identifying assumption is that, conditional upon our controls, the interaction between age and linguistic distance directly impacts only the score for Italian. Formally, the exclusion restriction is:

$$E(\epsilon_{isct}|Age_{isct} * Dist_{isct}, Dist_{isct}, Age_{isct}, \mathbf{X}, \vartheta_{sct}) = 0. \tag{5}$$

Specifically, the interaction term $Age * Dist_{isct}$ has no *direct* effect on the score for Math. The relevance of the instrument is examined in Section 3.1. Meanwhile, the following section reviews different mechanisms that may undermine our identification. We discuss the conditions under which our identification would be threatened, and present several placebo outcome regressions that demonstrate that such conditions should not occur.

3.1 Identification issues

Threats to our identification arise if an unobservable pattern exists that renders the *interaction* between age and linguistic distance endogenous, violating the conditional independence assumption. We empirically verify the existence of such a pattern in a number of situations by checking correlations that would be significant if our instrument were endogenous.

The circumstances that may invalidate our instrument can be traced to the behavior of the second generation's parents and native peers. For example, younger children are more likely to be bullied (Ballatore et al., 2020), and children from a more distant linguistic background may be more likely to be discriminated against. However, it is an empirical question whether age-related discrimination and linguistic discrimination are linked. We use survey data from the Student's Questionnaire to assess whether this is the case. The questionnaire is administered in the fifth grade and asks about family circumstances and personal preferences.[30] Appendix B provides more detail

[30]Many questions change over time and can be found only in some waves. This could potentially produce a sample selection issue. For instance, groups that are more discriminated against might not

on the survey questions we use and the models we estimate. The results in Table 19 suggest that discrimination does not pose a threat to our identification strategy. Using Student's Questionnaire data, we perform similar checks for other behavioral responses on the part of immigrant students, their parents or their peers in Appendix B. Overall, the results suggest our instrument is valid.

The existence of a confounder that directly links the *interaction* of age and linguistic distance to the score in *Math*, apart from its indirect effect through the treatment, is a further threat to our identification. One possibility why such a confounder might arise is that the linguistic distance is correlated (via endogenous sorting) with parental characteristics that affect the child's math learning *at a given age*. In this respect, the skill distribution of immigrants can be relevant. For example, the linguistic distance may be correlated with the parents' socioeconomic status, which could affect how much children learn *Math* by a given age. In recent decades, and in many countries, we have actually observed a remarkable upward trend in the skills of immigrants (Ehrlich and Kim, 2015; Ehrlich and Pei, 2021). Apparently, this is not the case in Italy, where immigration is made up of low skilled individuals compared to the native population, as we can see from Figure 7 (see also Bratti and Conti, 2018; Brunello et al., 2020; Del Boca and Venturini, 2005; Signorotto, 2015; Venturini and Villosio, 2008).[31]

Still, socioeconomic status is a major example, but one can conceive other violations of the exclusion restriction. Endogenous sorting and transmission to the learning of Math at a given age may work through other unobservable channels. As a consequence, we further investigate the validity of our instrument.

First, since we have four measures of linguistic distance, we construct four IVs and compare their behavior. As we can see from Tables 33, 34, 35, and 36 they all give equivalent results.

Then, we perform a Hansen J test for overidentification, which does not reject the joint null hypothesis of both instruments being valid.[32] This outcome makes us more

---

be represented in the available waves. To check for this eventuality, we control that the composition in terms of the main observable parental characteristics is stable across waves (Table 37 and 38). We also report the reference wave for each question.

[31]This is in line with the low returns for foreign human capital in Italy (Dell'Aringa et al., 2015) and is confirmed *a contrario* by the vast literature on the brain drain from Italy initiated by Becker et al. (2004). For recent evidence of the ongoing brain drain, see Assirelli et al. (2019).

[32]The Hansen J-statistic is 2.106 (p-value 0.1467) when we consider the first and the second instrument, and 1.292 (p-value 0.2558) when we consider the first and the third instrument. As suggested by Roodman (2009) in a different context, a p-value could be worrying below 0.1 or above 0.25. (We do not consider

confident in the validity of the IV we use in the main analysis (namely, the one based on the ASJP measure).

Finally, in order to provide additional support to the instrument's validity, we also apply the zero-first-stage test (Van Kippersluis and Rietveld, 2018). This test requires a subsample of children for whom there is no effect of the instrument on Italian in the first stage, and no effect of the instrument on Math in the reduced form. We obtained this subsample by pinning down the children with a score in Italian above 80/100. The results are visible in Table 32, and, again, suggest that our IV is well-founded. Though the zero-first-stage test is not a formal test of the exclusion restriction, it gives even further support to the restriction's validity.

4 Results

In Table 7, we report the OLS estimation of Equation (2) on the whole sample of the second generation. We report results using different model specifications.[33] Specification (1) shows a simple model with clusters at the language-wave level. Specifications (2) and (3) show a cohort fixed effect model with (one-way) clusters at the school-class-cohort level, and three-way clusters at the school, class and cohort level. Specification (4) is a two-way fixed effect model in which we control for separate cohort and class fixed effects. Specification (5) is a class-by-cohort fixed effect model. Specification (6), which is our preferred model, is a school-by-class-by-cohort fixed effect model, that allows us to control for unobserved differences within school-class-cohorts.[34] Finally, the last specification (Specification (7)) is a school-by-class-by-cohort fixed effect model to which we also add month fixed effects to account for possible birth seasonality. Although these all provide close estimates, their interpretation differs slightly.

We observe a sizable positive relationship between the score for Italian and the score for Math. Increasing the score in Italian by one point increases the score in

the fourth measure since, as we explain in section 2, it is just the square root of the third measure).

[33] Additional results from alternative specifications are available upon request.

[34] See footnote 26.

Math by 0.620 points.[35] However, this relationship is biased by unobserved variables and can hardly capture a causal effect. Thus, we present below the 2SLS analysis.

4.1 IV Results

The first stage of our IV estimation (Equation 3) is reported in Table 9. Our instrument is the interaction between age (measured in months) and linguistic distance, which are both continuous variables. Consequently, the coefficient of their interaction measures how many units the slope of Italian on Age is predicted to change with a one-unit increase in linguistic distance. A greater linguistic distance would render exposure less effective. Accordingly, the negative correlation we find is expected.[36]

The coefficient of Age is positive and significant at the 1% level. Note that, because the first-stage regression includes the interaction *Age*Distance*, this coefficient captures the effect of age *only when linguistic distance is zero*, namely, in the case of the second generation who speak Italian at home. For non-Italian speakers, the effect of age is given by $\alpha_1 * (Distance) + \alpha_3$.[37] For instance, in the case of the Chinese second generation, which is the most linguistically distant group, it is -0.097 score points.[38] As expected, the effect of age (i.e., exposure) decreases as linguistic distance increases. The coefficient of linguistic distance is also positive and significant at the 1% level. Again, given we have the interaction *Age*Distance*, by itself, it only provides the effect of linguistic distance *when age is zero*. When we consider this effect conditioned on a more plausible age, it becomes negative, as expected.[39] For instance, at the age of 123 months (10.25 years), it is -0.040. At the age of 132 months (11 years), it is -0.067 score points. Thus, *the effect* of linguistic distance decreases as exposure increases. Finally, the weak identification test confirms that the instrument is relevant. The Kleibergen-Paap Wald rk F statistic varies from 16.816 to 122.562, depending on the specification. We report the second stage of the 2SLS estimation in Table 10. Here, we observe a *negative* and statistically significant relationship between the score for

[35] Equivalently, a one-standard-deviation increase in the Italian score implies, on average, a 0.603-standard-deviations increase in the Math score

[36] Alternatively, this coefficient can be interpreted as representing how many units the slope of Italian on linguistic distance is predicted to change with a one-unit increase in exposure to Italian. Similarly, the effect of linguistic distance decreases as exposure to Italian increases.

[37] See Jaccard and Turrisi (2003) for a useful review of interaction effects.

[38] The calculation is made by considering column (6).

[39] Note that the effect is given by $\alpha_1 * (Age) + \alpha_2$.

Italian and the score for Math. In our preferred specification, increasing the score in Italian by one point decreases the score in Math by 0.397 points. This is equivalent to 5.65% of the gap in Math between native Italian speakers and second-generation immigrants. This negative coefficient appears to be at odds with the role of language as a prerequisite.[40] However, in the literature it is evident that finding a positive effect is harder than expected. Aparicio-Fenoll (2018), studying a sample of 1,529 immigrant children aged 6-12 in the U.S., finds no effect at all of language proficiency on Math problems or calculations.[41] Isphording et al. (2016) find the expected positive effect: a one-standard-deviation increase in reading performance improves performance in Math by 0.57 standard deviations. However, they use PISA data on boys aged 15-16, who had many more years to learn the host country's language than the 10-year-old children in our sample. These students are more likely to reap the benefits of language proficiency. In addition, it is possible that many boys with poor school performance have already dropped out of school at the age of 15-16, exposing their results to possible selection bias —this is the case in Italy, where there are concerns about the high dropout rates. [42]

To sum up, the evidence for the prerequisite role of language is far from robust. Yet, the intuition is compelling. The alternative idea —that language proficiency is detrimental to the understanding of Math classes— looks even less plausible. Rather, it is useful to look at the negative sign we have estimated in terms of a trade-off. Then, the question is whether there is any reason why children who improve their proficiency in Italian should sacrifice their performance in other subjects. We propose an answer to this question in the next section, where we introduce a theoretical model that accounts for the contrasting results in the literature.

[40]It is also interesting to note the difference between the OLS and IV estimates. The OLS estimator appears heavily *upward* biased. Isphording et al. (2016) and Aparicio-Fenoll (2018) identify an upward bias as well, though less considerable. In our case, the bias seems to be higher in absolute value than the negative value of the instrumented coefficient, explaining why the OLS coefficient is positive. This outcome aligns with the recent literature confirming the crucial importance of unobserved cognitive and non-cognitive abilities—and their interaction—for educational performance (Cunha et al., 2006; Almlund et al., 2011).

[41]A major difference with our approach is that Aparicio-Fenoll (2018) puts together 612 first-generation children and 917 second-generation children.

[42]According to a recent report (Eurostat, 2023), early leavers (defined as those who are not in education or training at the age of 18-24 and have completed *at most* a lower secondary education) in Italy are 12.7%. The EU average is 9.7%.

5 A theoretical foundation for the literature

In this section, we argue that the negative effect of language proficiency on Math performance we have estimated could mask a heterogeneous effect. More specifically, the IV estimates could be interpreted as a LATE, which would be informative of the subsample of individuals for whom the treatment status is affected by a change from z to z' (the *compliers*), for any pair (z, z') of values for $Age * Distance_{isct}$. What is the possible source of this heterogeneity? Established results in linguistics (Cummins, 2000; Cummins and Gulutsan, 1974; Toukomaa and Skutnabb-Kangas, 1977) suggest that there exists a sufficiency threshold that has to be crossed for reaping the gains of language proficiency. In fact, it is normally assumed that sufficient proficiency is required *to gain access to the curriculum*. However, if we accept that proficiency in Italian only helps to improve the score in Math beyond the sufficiency threshold, then what happens *inside* the threshold? We argue that there is only one possible answer; namely, if the child devotes more time to Italian, then she has to sacrifice her performance in Math, and vice versa. Our estimates pick up this trade-off.

Since we know no other contributions that connect the results in linguistics with the literature on skill production, before proceeding to further empirical analysis, we make our point clearer with the help of a theoretical model. While retaining the fundamental insights by Cunha et al. (2006), we show that a basic model of skill production that incorporates a threshold in language acquisition accounts for the mixed findings in the literature.

5.1 The production of skills

In its very essence, learning is a technology. As such, it can be described by a production function, which we call “skill production function”(henceforth SPF). What do we know about the features of this technology? In the literature, few papers try to model how children learn. Our natural starting point is the seminal contribution by Cunha and Heckman (2007), who point out two essential mechanisms at work in childhood: namely, the *self-productivity of skills* and the *dynamic complementarity of skills.* The former indicates that, in the educational process, the output of a stage is the input of the following stage. The latter points out that skills produced at one stage raise the productivity at subsequent stages.

In our framework, we adopt both these ideas but, in line with the linguistics literature and our findings, we also consider the existence of a *threshold* in language acquisition. As we have stressed, this threshold is taken for granted in a number of everyday circumstances. The standard practice of requiring language certifications for foreign students is a major example. In other words, universities put a *lower bound* on the language proficiency of foreign students before enrolling them in *any* course. The implicit assumption here is that this lower bound discriminates the students able to understand the classes from those who cannot.

In this framework, the dynamic complementarity arises just because language proficiency is a prerequisite for learning Math: we can think of the knowledge of Italian (measured by the test score) as an *intermediate input* for producing Math skills. Thus, we need two SPFs: one for Italian, and one for Math. In order to preserve the intuition, the SPF of Italian, as described below, is as simple as possible, and we assume that its output (the score in Italian, I) is a strictly increasing, strictly concave, and twice continuously differentiable function of the time spent studying Italian in the current period (L_{it}) plus the stock of knowledge of Italian accumulated in the past (I_0), which catches the self-productivity of skills. This additive term is relevant because it summarizes a child's past achievements.

Finally, to complete the description of the SPF, it is important to note that any assessment of a child's proficiency is inevitably noisy. Scores can be affected by various shocks, like the child's health on the day of the exam, or even luck. In order to account for this noise in the simplest possible way, we introduce an additive, uniformly distributed random shock $\rho \in [-a, a]$ with zero mean. Thus, the SPF of Italian is

$$I_{it} = f(L_{it} + I_0) + \rho. \tag{6}$$

Math skills (i.e., the score in Math) depend on the time spent studying Math (L_{mat}) plus the previous knowledge of Math (M_0), and on the intermediate input of Italian, for which we introduce a threshold level. A notable example of a technology that uses thresholds is the Stone-Geary production function. [43] In the Stone-Geary function, a

[43] Notably, Beattie and Aradhyula (2015) stress that, "it is hard to think of many production processes where one may reasonably expect positive output with input levels close to zero". In many processes,

positive output only appears after the input has crossed the threshold, creating a kink. For our purposes, however, we assume that some output (i.e., learning) exists even *below* the threshold. This happens because a child who finds it hard to understand the teacher can still learn something, or even make an effort to read the textbook. In other words, we have two cases: on/above the threshold the child learns Math benefiting from language proficiency (case A). Inside the threshold, she learns Math without any benefit from language proficiency (case B). The noise in the measure of skill is added, for simplicity and without loss of generality, through the same random shock ρ we use for the Italian SPF. This can be summarized as follows.

Let M be the score in Math. The Math SPF is given by

$$M = \begin{cases} g(L_{mat} + M_0, I_{it}) + \rho & \text{if } I_{it} \geq \bar{I} \qquad \text{(A)} \\ \\ h(L_{mat} + M_0) + \rho & \text{if } I_{it} < \bar{I} \qquad \text{(B)} \end{cases} \tag{7}$$

where $\bar{I}$ is the sufficiency threshold, $g(.)$ and $h(.)$ are two strictly increasing, strictly concave, and twice continuously differentiable functions.[44] We also assume $\frac{\partial g(.)}{\partial L_{mat}} > \frac{\partial h(.)}{\partial L_{mat}}$. In line with the dynamic complementarity, this assumption ensures that the marginal productivity of time spent studying Math is higher in Case (A).

5.2 Utility maximization

For our illustrative purposes, we assume that the child has a strictly increasing, strictly concave, twice continuously differentiable utility defined on her score in Math and Italian:

$$u = u(M, I) \tag{8}$$

We assume that the scores in Italian and Math are complements. In other words, even though a child may have different preferences for Italian and Math, she likes to achieve good scores in *both* subjects. We normalize the time endowment to unity, thus

threshold levels of the requisite inputs are the norm rather than the exception.

[44]For any $\ell_{mat} \in (0,1)$, the function $M(.,.)$ is left-continuous in $(\ell_{mat}, \bar{I})$: $\lim_{(L_{mat}, I_{it}) \longrightarrow (\ell_{mat}, \bar{I})^-} M(L_{mat}, I_{it}) = h(\ell_{mat})$.

the time constraint is

$$L_{it} + L_{mat} = 1, \tag{9}$$

where L_{it} and L_{mat} are, respectively, the time devoted to studying Italian and Math. A child maximizes the utility (8) subject to the time constraint (9) and to the SPFs (6) and (7).

Obviously, being in Case (A) or (B) on the Math SPF makes a crucial difference, and we have to take into account how the threshold affects the child's behavior. To this end, we can categorize the children as those who are definitively above the threshold (Case A) and those who aren't (Case B). The different initial conditions are due to the different stocks of accumulated knowledge, which are known to the children, who are aware of their previous grades and of their proficiency, having different stocks of accumulated knowledge.[45] Therefore, those in Case (B) realize the benefits of moving to Case (A). We discuss the optimization problem as follows.

5.2.1 Case A:

Consider Case (A) of the Math SPF (7). The child maximizes the (expected) utility (8), namely

$$\mathbb{E}[u] = \mathbb{E}_{\rho}[u[(f(.) + \rho), (g(.) + \rho)]] \tag{10}$$

subject to the time constraint (9), with respect to L_{it}. Through the Bolzano-Weierstrass theorem, we know that this problem has a solution. Since the utility is strictly concave and the constraint is linear, the solution is unique.[46] Thus, there exists a pair (L^*_{mat}, L^*_{it}) that gives, respectively, the optimal time devoted to studying Math and the optimal time devoted to studying Italian.

By substituting the equilibrium values L^*_{mat} and L^*_{it} into the SPFs, the equilibrium score in Math is

[45] To be precise, children in Case A have I_0 such that $f(0 + I_0) > (\bar{I} + a)$. In other words, even though they choose the corner solution $L^*_{it} = 0$ and receive the worst negative shock $(-a)$, they are still above the threshold. All children with a lower I_0 are in Case B.

[46] This optimization problem is standard, except for the random shock ρ in the SPFs. However, since ρ is uniformly distributed on the closed interval $[-a, a]$ and its variance is finite, the objective function (8) is still continuous. Since the time endowment is normalized to unity, the choice variable L_{it} is in the closed interval $[0, 1]$, and the constraints are compact. Thus, we can apply the Bolzano-Weierstrass theorem.

$$M^* = g(L^*_{mat} + M_0, I^*) + \rho. \tag{11}$$

Note that, at equilibrium (11), L^*_{mat} is a decreasing function of I^* through the time constraint (9).[47] Thus, we can write $M^* = g(L^*_{mat}(I^*) + M_0, I^*)$, and the effect of the score in Italian on the score in Math is given by the derivative

$$\frac{\partial M^*}{\partial I^*} = \frac{\partial g(L^*_{mat} + M_0, I^*)}{\partial I^*} + \frac{\partial g(L^*_{mat} + M_0, I^*)}{\partial L^*_{mat}} \frac{dL^*_{mat}}{dI^*}. \tag{13}$$

In equation (13), the first term on the right-hand side is the marginal gain in the Math score of increasing the language proficiency. This effect is positive, because proficiency helps to understand the classes. Since, however, increasing the Italian score requires sacrificing some time spent on Math, the second term on the right-hand side is negative and measures the ensuing marginal loss. The net effect is non-negative if

$$\frac{\partial g(L^*_{mat} + M_0, I^*)}{\partial I^*} \geq -\frac{\partial g(L^*_{mat} + M_0, I^*)}{\partial L^*_{mat}} \frac{dL^*_{mat}}{dI^*}. \tag{14}$$

When condition (14) holds, increasing the score in Italian does *not* generate a trade-off with Math. When it holds with strict inequality, increasing the score in Italian causes an increase in the Math score. In this case, the child collects large gains from language acquisition and dynamic complementarity.

5.2.2 Case B:

The children inside the threshold are aware of the benefits of language acquisition, but, since their achievement is *uncertain*, they cannot be sure of moving across $\bar{I}$. Nonetheless, they realize that their chances increase with the time spent studying Italian. Let us denote the probability of crossing the threshold with $p(L_{it}) \in (0, 1)$.[48] Thus, they end up in Case (A) with probability $p(L_{it})$, and in Case (B) with probability

[47] To formally prove that this term is negative note that, by the time constraint,

$$\frac{dL^*_{mat}}{dI^*} = -\frac{df^{-1}(I^*)}{d(I^*)} < 0. \tag{12}$$

[48] We assume $\frac{dp}{dL_{it}} > 0$ and finite.

$(1 - p(L_{it}))$. Finally, they maximize

$$\mathbb{E}[u] = p(L_{it})\mathbb{E}_{\rho}[u(g((.) + \rho), (f(.) + \rho))] + (1 - p(L_{it}))\mathbb{E}_{\rho}[u((h(.) + \rho), (f(.) + \rho))] \tag{15}$$

subject to the time constraint (9), with respect to L_{it}.
In this problem, the conditions to apply the Bolzano-Weierstrass theorem still hold. Again, since the utility (15) is strictly concave and the constraint is linear, the problem has a unique solution (L^*_{mat}, L^*_{it}). Ex post, either the child succeeds in crossing the threshold or she does not. If she succeeds, she moves to Case (A) of the Math SPF. If she does not succeed, she stays in Case (B), thus $M^* = h(L^*_{mat}) + \rho$. At the equilibrium, the effect of the score in Italian on the score in Math is given by the derivative

$$\frac{\partial M^*}{\partial I^*} = \frac{\partial h(.)}{\partial L^*_{mat}} \frac{\partial L^*_{mat}}{\partial I^*}, \tag{16}$$

which is unambiguously negative through the time constraint (9).
Derivative (16) shows that, when the child is below the sufficiency threshold in Italian, she *always* faces a trade-off between Italian and Math. This is the case of the estimates in Table 10. More generally, derivatives (13) and (16) show the possibility of a trade-off between Italian and Math even though language proficiency is a prerequisite to understand Math. Thus, they provide a theoretical foundation to the mixed results in the literature. In particular, results in Isphording et al. (2016) and Aparicio-Fenoll (2018), correspond to condition (14), respectively, with strict inequality and equality. In the next section, we test the prediction of equation (16) by splitting our sample between children inside and above the threshold. This check adds to our empirical results, providing further evidence for the crucial role of the sufficiency threshold for second-generation children.

6 Additional evidence and heterogeneity

The theoretical analysis we have proposed predicts that children inside the sufficiency threshold who improve their language proficiency will not benefit from this in Math, and instead face a trade-off between language proficiency and Math proficiency. This prediction can easily be tested, once the sufficiency threshold is properly characterized.

In line with widely adopted criteria, INVALSI considers Proficiency Level 3 to represent "sufficient" knowledge of Italian; as such, we compute the threshold accordingly.[49]

In Table 11, we report the first-stage estimation for the second-generation children whose performance in Italian is below Proficiency Level 3. Note that we split the sample by considering students' *predicted* proficiency—based on the predetermined covariates—rather than their actual proficiency. This avoids any bias introduced by splitting the sample on the dependent variable of the first stage. Again, we may note that the effect of exposure to Italian decreases as linguistic distance increases, and the effect of linguistic distance decreases as exposure increases.

The weak identification test indicates that the instrument is relevant. The F statistic varies between 15.778 and 99.348, depending on the specification. Second-stage estimates are presented in Table 12. The effect of the Italian score on the Math score is negative and statistically significant at the 1% level. Increasing the score in Italian by one point decreases the score in Math by 0.452 points, depending on the specification. All else being equal, the impact of Italian on Math is equivalent to 6.43% of the gap between natives and the second generation. This suggests that the trade-off we bring to light is driven by the subsample of children below the sufficiency threshold (i.e., Proficiency Level 3).

The same estimation strategy for the children who perform above the threshold does not work as well. In Tables 13 and 14, we report the first and second stages for children above the sufficiency threshold. The instrument is weak, suggesting that exposure to Italian *per se* does not explain the performance *above sufficiency*. That is, once sufficiency is attained, further exposure to Italian could be of limited relevance to the *academic* score. This is consistent with the nature of our instrument, which in practice measures *passive* learning of Italian (roughly speaking, exposure to the language[50] weighted for the linguistic distance). Intuitively, passive learning helps to achieve a sufficient language proficiency, but looks less effective for achieving high *academic* scores. Consequently, we cannot use our instrument to estimate the causal effect of the Italian score on the Math score for students above Proficiency Level 3.

Not surprisingly, we observe that 68.4% of children above sufficiency speak a Romance

[49] Proficiency Level 3 is defined as a score belonging to a range of $(95 - 110\%)$ of the average score obtained by natives. We use the lower bound, 95%, as our threshold.

[50] The children in our sample were all born in Italy.

language. Below sufficiency, this share drops to 56.24%. It is natural to hypothesize that children with a Romance linguistic background may find it easier to achieve sufficiency in Italian. Thus, as a further check, we split our sample between Romance and non-Romance-language speakers. [51] Interestingly, the former outperform the latter in both Math and Italian. On average, the gap between non-Romance- and Romance-language speakers is 6.29 points in Italian and 3.24 points in Math (Table 5).

Table 5: Descriptive Statistics. Speakers of Romance languages vs. Speakers of non-Romance languages

Variable	Romance	Non-Romance	Diff.
Math	52.498	49.257	3.241***
	(0.105)	(0.134)	(0.170)
Italian	56.994	50.706	6.288***
	(0.096)	(0.122)	(0.156)
Female	0.500	0.491	0.008***
	(0.002)	(0.002)	(0.003)
Age in Months	130.166	130.424	-0.258***
	(0.017)	(0.022)	(0.028)
ESCS student	-0.412	-0.639	0.227***
	(0.004)	(0.006)	(0.007)
Obs	83,351	52,730	

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$
Standard Errors in parenthesis clustered at the school-cohort level.

We report our IV analysis for Romance and non-Romance language speakers in Tables 15, 16, 17, and 18. After instrumenting, we observe different trajectories for the two subsamples. In Table 15, we report the first stage for non-Romance-language speakers. The instrument is relevant. The F statistic varies between 12.305 and 87.743, depending on the specification. In Table 16, we report the second stage for non-Romance-language speakers. The effect of the score in Italian is negative and statistically significant at the 1% level. Increasing the score in Italian by one point decreases the score in Math by 0.692 points in our preferred specification. All else being equal, the impact of Italian on Math is equivalent to 9.85% of the gap between natives and the second generation.

As for Romance speakers, we report the first stage in Table 17. The instrument is weak.

[51]The Romance languages included in our sample are French, Ladin, Portuguese, Romanian, and Spanish. The non-Romance languages included are Albanian, Arabic, Chinese, Croatian, Greek, Hindi, English, Slovenian, and German.

The Kleibergen-Paap Wald rk F statistic varies between 0.151 and 1.581, depending on the specification. In the second stage, (Table 18) the estimated coefficient of the score in Italian is now not significant.

Overall, these findings suggest that the trade-off between acquiring Mathematical skills at the cost of performing more poorly in Italian especially concerns non-Romance-language speaking children. Romance language speakers plausibly find it easier to learn Italian and, at the age of 10, do not seem subject to the trade-off.

7 Conclusions

In this study, we estimated the effect of language acquisition on the Math performance of second-generation children at the end of Italian primary school. This age deserves special attention, since, as the extant literature suggests, 1) early educational gaps may have lifetime consequences and, in any case, are hard to overcome in later years; 2) language proficiency is required to acquire other forms of human capital; 3) language proficiency is crucial for the social and economic integration of the second generation. The literature is ever evolving, and the results are so far ambiguous. While Isphording et al. (2016) find a positive effect of linguistic performances on Math outcomes for 15 year-old first-generation immigrants, Aparicio-Fenoll (2018) finds no evidence of such an effect on a mixed sample of first and second-generation children aged 6-12. We find that higher scores in Italian *reduce* the score in Math for *second-generation* immigrant children in Italy.

As we have shown in our model, this finding does not contradict the idea that language proficiency is a prerequisite for understanding Math classes taught in Italian. Rather, it emphasizes that the second generation is still struggling to learn Italian at the age of 10, and can do so only at the cost of reducing its performance in other subjects. In line with the linguistics literature, our model is based on the hypothesis that the benefits of language acquisition are effective only once a child crosses a threshold of sufficiency, which we identify according to widely adopted criteria. In practice, we confirm that the trade-off between Italian and Math is driven by children below the threshold. These children account for 59% of our sample. Thus, it seems that the Italian school system is failing its objective of reducing the intergenerational transmission of inequality: the majority of the second generation is already being left behind. These children cannot benefit from the complementarity between language proficiency and other forms of human capital as their native peers do. They appear doomed to poor educational performances and, consequently, poor labor market outcomes.[52] In the long run, this penalization may easily transform immigrant communities into permanently disadvantaged minorities, fostering dangerous social stratification.

Overall, our findings have profound policy implications. They suggest that primary education should consider linguistic integration as a priority, and strive to lead children with poor linguistic backgrounds to the sufficiency threshold. This means that *marginal* interventions would hardly be effective for those who are lagging behind, and that only large-scale investments in education would be able to foster intergenerational integration.

In general, we confirm that policies aimed at linguistic integration should be of the greatest importance: in the short run they improve economic opportunities for newcomers, and in the long run they reduce intergenerational inequality. Unlike policies that take place later in life, achieving linguistic integration in primary school would be a simpler solution and produce permanent effects.

[52] Notice also that disadvantages tend to grow over time, due to mechanisms like the dynamic complementarity and the self-productivity of skills outlined by Cunha and Heckman (2007), which we described in section 5.

Acknowledgements

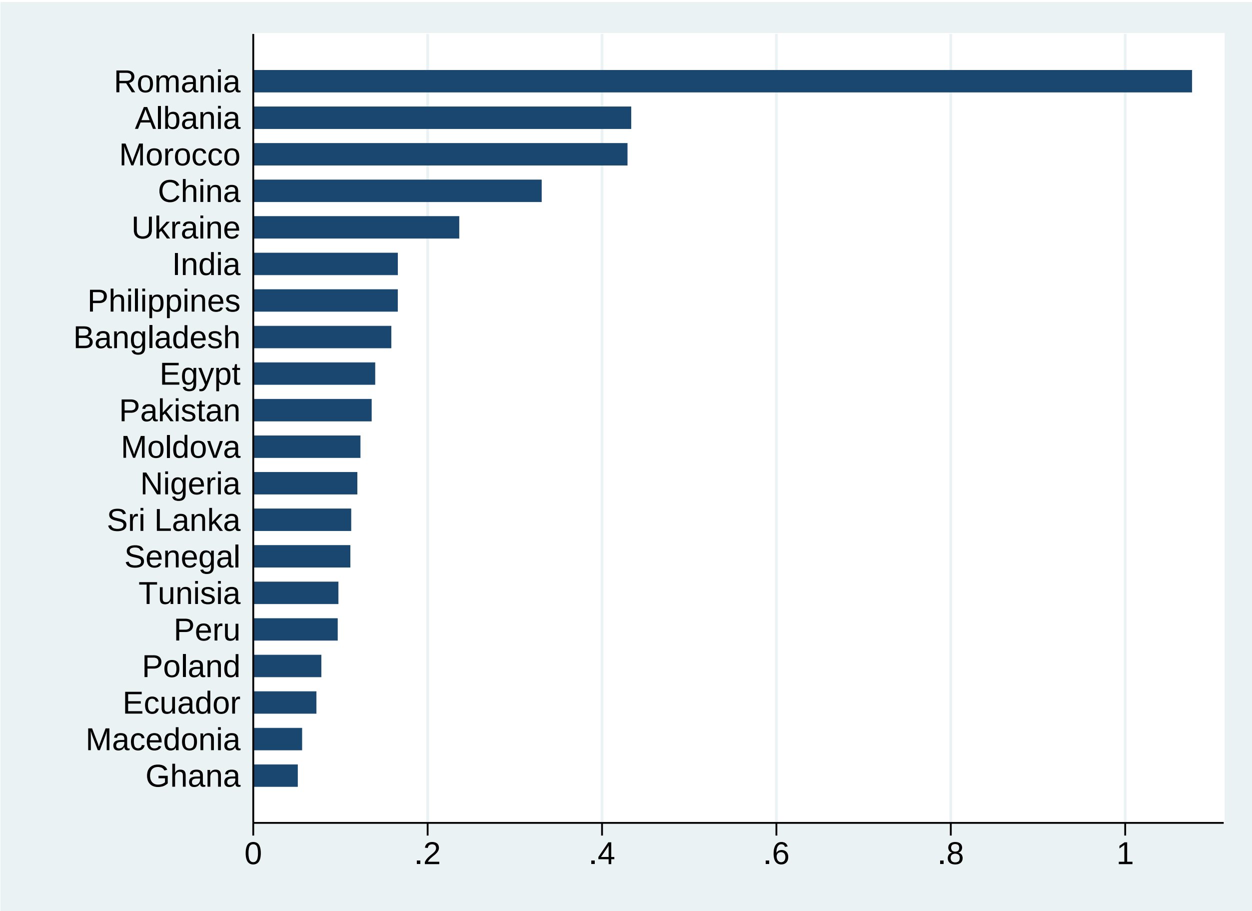
This paper uses confidential data kindly provided by the statistical archives of the Italian Institute for the Evaluation of the Educational System (INVALSI). We thank Patrizia Giannantoni for her support. The views expressed in this paper are solely those of the authors and do not necessarily reflect those of INVALSI.

We acknowledge the Editor, the Editor-in-Chief and two anonymous referees, who helped us to greatly improve our manuscript. We are grateful to Antonio Abatemarco, Michel Beine, Gaetano Bloise, Alberto Bucci, Simon Burgess, Vincenzo Carrieri, Antonio Cosma, Fausto Galli, Ellen Greaves, Tullio Jappelli, Sekou Keita, Immacolata Marino, Jeremy McCauley, Annamaria Menichini, Giuseppe Migali, Tommaso Oliviero, Vincenzo Platino, Francesco Principe, Annalisa Scognamiglio, Hans H. Sievertsen, Francesca Toscano, Christine Valente and seminar participants at LEM (University of Lille), the University of Bristol, the University of Trieste; we also thank participants to the 11th International Conference on Economics of Global Interactions (University of Bari), the 42nd ASSET Meeting (University of Marseille), the 35th ESPE Annual Conference (University of Cosenza), the 37th AIEL National Conference of Labour Economics (University of Salerno). This study was funded by the Italian Ministry of Research and University (PRIN 2017) grant no. 2017KHR4MB. The usual disclaimers apply.

A Immigration to Italy

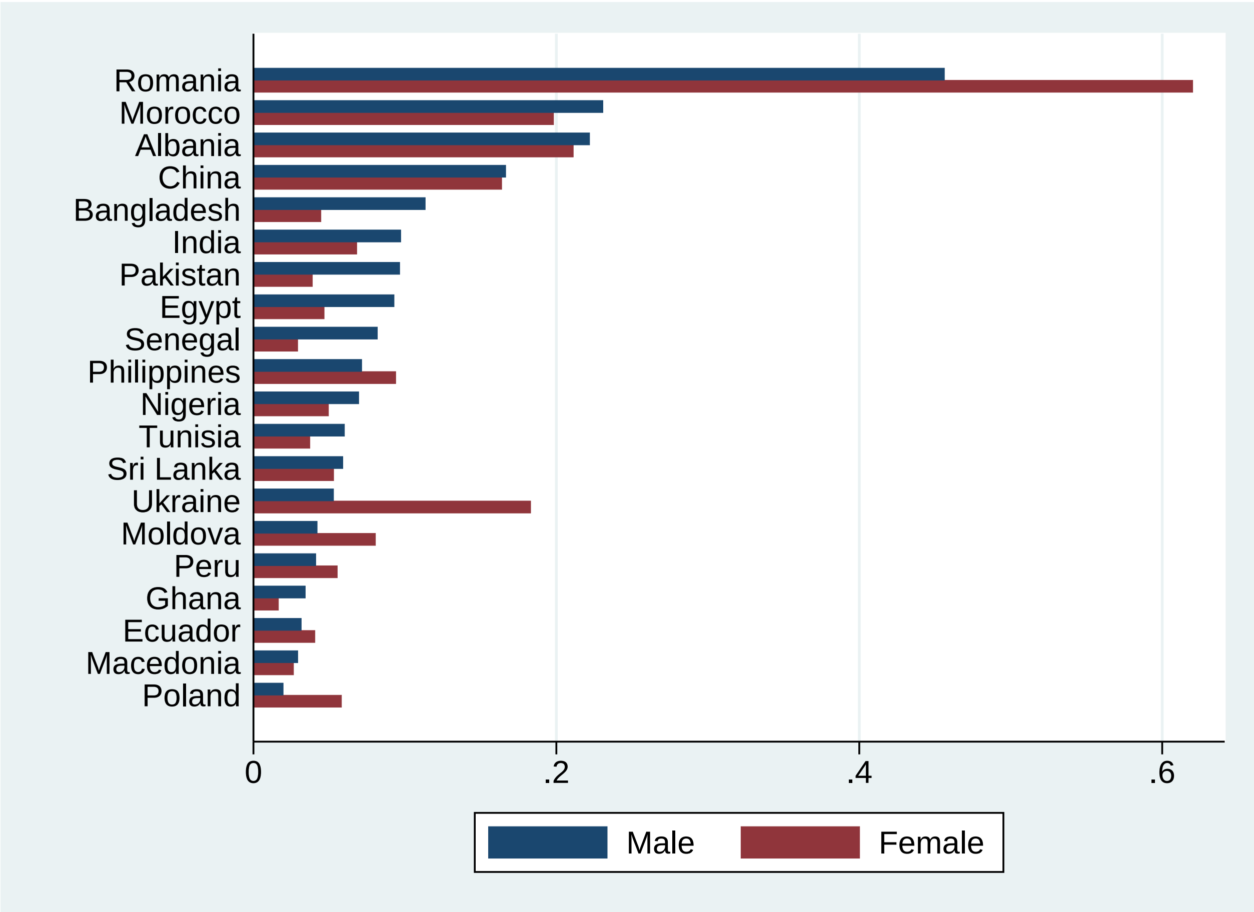
In Italy as of January 1, 2021, we observe 5,171,894 registered immigrants from 187 countries. The main nationalities are given in Figure 1. Romanians are by far the most substantial group, followed by Albanians, Moroccans, Chinese, Ukrainian, Indians, and Filipinos. Together, these groups account for half of the immigrant population.

Figure 1: Immigrant population in Italy by country of origin



Note: Immigrant population in Italy (in millions). Top 20 countries of origin on January 1^{st}, 2021.
Source: ISTAT (2022).

Figure 2: Immigrant population in Italy by gender and country



Note: Immigrant population by gender in Italy (in millions). Top 20 countries of origin on January 1^{st}, 2021.
Source: ISTAT (2022).

As Figure 2 shows, gender is fairly balanced except in the case of certain nationalities (Romania, Ukraine, and Poland) that frequently account for caregivers, among which women are the majority.

The structure of the immigrant population mirrors the history of immigration to Italy. Mass immigration is a recent phenomenon in Italy, which was a sending country until the 1970s. The first significant inflows date to the fall of communism in Eastern Europe. On August 8, 1991, news media reported the landing of a cargo ship that had been rerouted from Albania to the port of Bari; the ship's 20,000 immigrants soon fled into the city. Authorities tried to confine them to a stadium to proceed with repatriation, causing riots. This event marked the onset of mass immigration. Thereafter, inflows of Albanians continued regularly, and they are now the second largest community in Italy.

They are mainly employed in low-skill jobs in the construction and catering sectors MLPS (2021a).

The Romanian revolution of 1989 fostered an inflow of Romanian citizens that continued steadily and increased substantially after January 2007, when Romania joined the EU. Today, Romanian women are mostly employed in health care, tourism, and catering. Romanian men work mostly in the construction sector, although 50,000 businesses have been registered by members of this community Caritas (2020). The stories of other Eastern European communities (including Poland, Ukraine, and Moldova) are similar, with flows beginning in the 1990s, and continuing steadily since. Immigrants from these countries supply labor to the health care and construction sectors.

Immigration from Africa has a different history. African immigration was already established in the 1980s, with most immigrants working in agriculture. Interestingly, early communities of Tunisian workers settled in Sicily in the 1960s, where they were needed in the fishing industry. Over time, jobs shifted from agriculture to industry, and the majority is now employed in low-skill, manual industrial jobs in Northern Italy (MLPS, 2021e,f). However, the number of African immigrants working in agriculture is underestimated due to unregistered seasonal workers.

Meanwhile, Asia has become the most recent region of origin. The population of Chinese immigrants, who numbered about 13,000 in 1991, is now larger than 300,000. They are particularly involved in trade, import-export, and commercial sectors. Chinese immigrants demonstrate the best results in terms of employment rate among non-EU communities, reaching 68.7% in 2020 compared to an average of 56.6% for non-EU immigrants. Their unemployment rate was 3%, compared to 13% for non-EU immigrants. This difference is due to the participation of women in the labor market (MLPS, 2021b). In contrast, Indian immigrants are mainly employed in agriculture, particularly in sheep-farming (MLPS, 2021d).

Filipinos represent an interesting exception among Asians: their first inflows date back to the 1960s, typified by small numbers of domestic workers (mainly women) entering Italy with the help of religious authorities (because the Philippines is a Catholic country). The Filipino community is now in its third generation, and individuals perform well in the labor market, with many having acquired Italian nationality (MLPS, 2021c).

This brief survey of immigration to Italy provides insight into the magnitude of the phe-

Table 6: Residence permits (2020)

	Number of Permits	%
Family reunion	1,657,591	48.2%
Work	1,430,506	41.6%
Asylum	194,799	5.7%
Study	52,004	1.5%
Religious discrimination	27,558	0.8%
Custody and integration of minors	17,869	0.5%
Other	57,780	1.7%
Total	3,438,707	100

Source: Caritas (2020).

nomenon, as well as the main nationalities involved. However, we are also interested in knowing possible features of immigrants that could affect second-generation children. For instance, economic immigrants can differ substantially from other categories of immigrants in many respects (as is the case for refugees). Hence, we handle this issue by considering residence permits. The table below indicates that family reunion represents the main way for individuals to enter Italy, followed by work permits. Family reunion is often used to facilitate entry to immigrants who are searching for a job. Meanwhile, entries related to refugee status (Asylum, Religious discrimination, Custody and integration of minors) only account for 7% of the total. In addition, the entry "Custody and integration of minors" depicts unaccompanied *first-generation* children under 18. Thus, we can conclude that most second-generation immigrants are the offspring of economic immigrants, and that other categories are of little relevance.

A.1 Rules for immigrant students in Italy

The criteria for the integration of immigrant children in Italian schools are reported in the Guidelines for the Reception and Integration of Foreign Students (*Linee guida per l'accoglienza e l'integrazione degli studenti stranieri*), issued in 2014 by the Ministry of Education (hereafter Guidelines). Notably, this document is very general and concerns schools of all levels. It assures equal treatment for natives and immigrants and provides some recommendations for the integration of foreign students. First, immigrants should be uniformly distributed across schools, fostering heterogeneity in the classroom, rather than concentrating children of the same nationality or religion.

Thus, a 30% cap on the proportion of foreign students in a given class is recommended; however, this can be increased or reduced given reasonable cause.

Enrollment of first-generation immigrants is possible immediately upon their arrival, avoiding the deadlines for residents in Italy.[53] The standard procedure is to enroll these children in the grade corresponding to their age, assuming that they have received education in their country of origin. However, the Board of Teachers can enroll children in the previous grade if they consider their linguistic disadvantage as too severe.

Regarding exams, the law requires that they be the same for all students, unless certified disabilities exist. For the final exam at the end of primary school, immigrant children may receive assistance in their mother tongue if one of their teachers speaks it. The Guidelines also foster cooperation between families and schools with the help of *cultural mediators*, who keep parents informed about their child's performance and the procedures necessary for their child's attendance at the school (including enrollment, deadlines, and documents). The cultural mediators are widely used to facilitate the interaction between schools and immigrant families. However, *they have no teaching role*; instead, they are supposed to enable communication between schools and families, especially when the latter do not speak Italian.

An interesting section of the Guidelines concerns teaching Italian as a second language. However, it is specified that these recommendations concern *secondary* schools, where—according to the document—it is most urgently needed. The Guidelines provide a three-step procedure for teaching Italian as a second language: a) *Learning to communicate*, which requires the student to learn 2000 basic words and basic grammar, and develop basic reading and listening skills; b) *Learning to study*, in which the student should strengthen their knowledge and fluency with the help of multilingual glossaries, simplified textbooks, and pathways to the development of further reading and writing skills; c) *Learning with classmates*, in which students have a command of the language sufficient to attend standard classes with their native peers under the supervision of a teacher, who should facilitate communication.

For Step a), the Guidelines suggest 8–10 hours of teaching per week for 3–4 months. Steps 2) and 3) can be less intensive. Teaching in small groups is recommended. However, the actual organization of these activities is left to schools, without specifying

[53] Note that irregular immigrants retain the right to education.

any sources of funding required to conduct these special teaching programs.

A 2012 law provides a more concrete approach to the integration of immigrant children, introducing the possibility of adopting customized study plans *(piani di studio personalizzati)* for children with limited proficiency in Italian. These plans may be adopted by the schools after a linguistic assessment of the child and involve replacing the most linguistically demanding subjects with easier subjects, or simply reducing the educational objectives the student must meet to proceed to the next level. There also some Math textbooks available for children with a limited command of Italian (see, for instance, Arici and Maniotti, 2010). However, in most cases, the measures we have reported are mainly a declaration of intent; broadly speaking, Italy still lacks a comprehensive approach to the linguistic integration of minorities.

B Threats to Identification

B.0.1 Discrimination.

Native peers could discriminate against second-generation children in social interactions, leading to a failure of exposure to Italian. This section uses reported bullying as a proxy for discrimination. Although discrimination against peers who speak different languages is plausible, as we have argued, this is is insufficient to invalidate our identification. To understand our reasoning, consider first that, as has been documented, younger children are more likely to be bullied (Ballatore et al., 2020). This produces endogeneity in one dimension. It is also true that children from a distant linguistic background (e.g., demonstrating a foreign accent) are more likely to be discriminated against. This produces endogeneity in the other dimension. For our instrument to be endogenous, a pattern linking age-related discrimination to linguistic discrimination is required. Thus, discrimination should be correlated to the *interaction* between age and linguistic distance.

To perform this check, we use the following questions regarding bullying (Student's Questionnaire 2014–15):

- *During this school year, how often have you been teased by other students at school?*
- *During this school year, how often have you been insulted by other students at school?*
- *During this school year, how often have you been isolated or excluded from other students at school?*
- *During this school year, how often have you been beaten by other students at school?*

and we estimate the following linear probability model (LPM):[54]

$$y_{isct,\,t=2015} = \beta_0 + \beta_1 Dist_{isct,\,t=2015} + \beta_2 Age_{isct,\,t=2015} + \beta_3 Age_{isct,\,t=2015} * Dist_{isct,\,t=2015} + \\ + \mathbf{X}\lambda + \vartheta_{sc} + \epsilon_{isct,\,t=2015} \quad (17)$$

where for $y_{isct,\,t=2015}$ we consider, in turn, the dummy for whether the child is teased, insulted, isolated, or beaten by other children at school. For all these outcomes, the coefficient of the interaction (*Age*Distance*) is not significant (Table 19).

Discrimination by teachers represents another potential source of endogeneity.

To investigate this issue, we estimate Equation 17 by considering various measures of teacher attitudes as dependent variables. In particular, we consider the following questions (Student's Questionnaire 2014-15):

- *In my class, when we have an issue, we are listened to*;
- *In my class, we are listened to attentively when we interact during the lesson*;
- *In my class, we are encouraged to ask questions during the lesson*.

In these cases, the coefficient of the interaction (*Age*Distance*) is, again, not significant (Table 28).

B.0.2 Self-segregation.

Immigrant households might try to reduce their exposure to Italian culture to preserve their traditional norms and customs. We can roughly check for this issue by considering the correlation between interaction *Age*Distance* and the probability that second-generation children isolate their peers. In particular, to estimate Equation 17, we use the following question (Student's Questionnaire 2014–15):

- *During this school year, how often have you isolated or excluded other students at school?*

Table 19 shows no significant correlation between the probability of isolating or excluding peers and the interaction *Age*Distance*. This aligns with the literature on

[54] In alternative specifications, available upon request, we have also estimated a probit and an ordered probit.

endogenous segregation, which suggests that the (choice of) separation from the receiving society may occur at *any* level of linguistic distance (Battu et al., 2007; Bisin et al., 2011, 2016; Constant et al., 2009; De Martì Beltran and Zenou, 2017). Self-segregation may even occur among minorities speaking *the same* language as the majority. For instance, it has been documented that some African-American students in the US may be ambivalent about performing well at school because doing so might be considered "acting white"(Austen-Smith and Fryer, 2005; Fryer Jr. and Torelli, 2010).

B.0.3 Enrichment activities.

Households can support their children in the form of help with homework, tutoring, reading, and other activities. Although one can understand the rationale for parental behavior in response to age and linguistic distance alone, it is unlikely that families adjust the intensity of these activities due to a further (unobserved) non-separable effect linking age and linguistic distance. The existing evidence shows that parental support depends on (parental) education and income (which we control for) rather than, for instance, linguistic distance (Aguiar and Hurst, 2007; Doepke and Zilibotti, 2019; Duncan and Murnane, 2011).

However, as a further check, we examine the correlation between *Age*Distance* and several variables that should predict parental support: the number of books at home, the availability of a quiet place to study, computers, encyclopedias, internet connections, and students having their own room (Student's Questionnaire 2014–15; 2015–16; 2016–17; 2017–18; 2018–19). We estimate the following LPM:

$$y_{isct} = \beta_0 + \beta_1 Dist_{isct} + \beta_2 Age_{isct} + \beta_3 Age_{isct} * Dist_{isct} + \mathbf{X}\lambda + \vartheta_{sct} + \epsilon_{isct} \quad (18)$$

where y_{isct} are dummies for any of these outcomes. These variables are mostly unrelated to the interaction *Age*Distance* in most cases (Table 20).[55] By comparison, parental employment and education are far more likely to be related to *Age*Distance* (Tables 22, 23, 25, 26). This seems to suggest that if parental assistance plays some role, the effect is likely to be captured by our observable variables.

[55] For this table, questions derive from the Student's Questionnaire 2014–15; 2015–16; 2016–17; 2018–19.

B.0.4 Preferences.

Another matter for consideration is the possibility that, due to their lack of language proficiency, children with an immigrant background prefer Math or make more effort in the subject.

To investigate this issue, we consider the following questions (Student's Questionnaire 2014–15):

- *How much do you agree with the following statement? I like studying Math*;
- *How much do you agree with the following statement? I like studying Italian*.

We subsequently estimate the following LPM:

$$MathPref_{isct,\,t=2015} = \beta_0 + \beta_1 Dist_{isct,\,t=2015} + \beta_2 Age_{isct,\,t=2015} + \beta_3 Age_{isct,\,t=2015} * Dist_{isct,\,t=2015} + \\ + \mathbf{X}\lambda + \vartheta_{sc} + \epsilon_{isct,\,t=2015} \quad (19)$$

where $Math\ Preferences_{isct,\,t=2015}$ is a dummy representing whether the student likes studying Math.

We find that preferences for Math are mostly unrelated to the interaction Age*Distance (Table 29). We similarly investigate the role of student preference for Italian, revealing no relation with Age*Distance. These results plausibly confirm that our findings are not driven by preferences.

B.0.5 Socio-emotional skills

Another possible confounding factor concerns the specific impact of socio-emotional skills on test scores for the second generation. Because more mature children should feel more confident, we can conjecture a negative relation between anxiety and age. It is possible that more linguistically distant children feel more anxious when taking a test written in Italian. To investigate this issue, we consider the following statements (Student's Questionnaire 2014–15; 2015–16; 2016–17):

- *Even beforehand, I was worried about the test*;
- *I was so nervous that I was not able to find the answers*;

- *While I was answering, I had the feeling that I was wrong*;
- *While I was answering, I felt calm*.

We subsequently estimate Equation (17) by considering these outcomes.

In Table 31, we can observe that the interaction *Age*Distance* is unrelated to these features of text anxiety.

B.0.6 Enrollment manipulation.

In many cases, parents can manipulate school enrollment.[56] Parents with a linguistic background other than Italian could delay the enrollment of their children to give them more time to learn the language. However, the opposite also holds true. That is, parents could be willing to expose their children to Italian schools as soon as possible, thus hastening their enrollment. These possibilities indicate the potential for a relationship between age and linguistic distance.[57] To investigate this issue, we estimate the following LPM:

$$y_{isct} = \beta_0 + \beta_1 Dist_{isct} + \mathbf{X}\lambda + \vartheta_{sct} + \epsilon_{isct} \tag{20}$$

where y_{isct} is a dummy equal to one if the child is enrolled late. The correlation between linguistic distance and the probability of postponing the enrollment is significant but its magnitude is almost negligible (0.000267) (Table 30). Considering the range of linguistic distance–––from 58.77 (Romanian) to 101.14 (Chinese)–––it can be observed that the probability of late enrollment increases by 1.13% in the shift from the closest to the farthest language. As such, linguistic distance alone appears unimportant for enrollment decisions.

[56] In Italy, the school year begins in September. Although enrollment is compulsory for all children who turn 6 before December 31, parents can enroll younger children provided that they turn 6 by April 30 of the following year. Consequently, children born between January 1 and April 30 can either be enrolled when they are still 5 or be enrolled one year later.

[57] Note that age is unaffected by grade retention, which is forbidden in Italian primary schools (see footnote 27).

References

Abatemarco, Antonio, Mariagrazia Cavallo, Immacolata Marino, and Giuseppe Russo (2022). “Age Effects in Primary Education: A Double Disadvantage for Second-Generation Immigrants”. *SSRN Discussion Paper* 4199896.

Agnelli Giovanni, Fondazione (2014). *La valutazione della scuola: A che cosa serve e perché è necessaria all'Italia*. Editori Laterza, Bari.

Aguiar, Mark and Erik Hurst (2007). “Measuring Trends in Leisure: The Allocation of Time Over Five Decades”. *The Quarterly Journal of Economics* 122.3, pp. 969–1006.

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz (2011). “Personality psychology and economics”. In E. Hanushek, S. Machin, and L. Woessmann (eds.) Handbook of the Economics of Education(.4), pp. 1–181.

Almond, Douglas, Janet Currie, and Valentina Duque (2018). “Childhood Circumstances and Adult Outcomes: Act II”. *Journal of Economic Literature* 56.4, pp. 1360–1446.

Andersson, David, Mounir Karadja, and Erik Prawitz (2022). “Mass Migration and Technological Change”. *Journal of the European Economic Association* 20.5, pp. 1859–1896.

Angrist, Joshua D., Erich Battistin, and Daniela Vuri (2017). “In a Small Moment: Class Size and Moral Hazard in the Italian Mezzogiorno”. *American Economic Journal: Applied Economics* 9.4, pp. 216–49.

Aoki, Yu and Lualhati Santiago (2018). “Speak better, do better? Education and health of migrants in the UK”. *Labour Economics* 52, pp. 1 –17.

Aparicio-Fenoll, Ainhoa (2018). “English proficiency and mathematics test scores of immigrant children in the US”. *Economics of Education Review* 64, pp. 102–113.

Arici, Maria and Paola Maniotti (2010). *Studiare matematica e scienze in italiano L2. Unità di apprendimento per alunni stranieri della scuola primaria*. Erickson.

Assirelli, Giulia, Carlo Barone, and Ettore Recchi (2019). ““You better move on”: Determinants and labor market outcomes of graduate migration from Italy”. *International Migration Review* 53.1, pp. 4–25.

Aucejo, Esteban M. and Jonathan James (2021). “The Path to College Education: The Role of Math and Verbal Skills”. *Journal of Political Economy* 129.10, 2905–2946.

Austen-Smith, David and Roland G. Fryer (2005). "An Economic Analysis of Acting White". *The Quarterly Journal of Economics* 120.2, pp. 551–583.

Bacolod, Marigee and Marcos A. Rangel (2017). "Economic Assimilation and Skill Acquisition: Evidence From the Occupational Sorting of Childhood Immigrants". *Demography* 54.2, pp. 571–602.

Ballatore, Rosario Maria, Marco Paccagnella, and Marco Tonello (2020). "Bullied because younger than my mates? The effect of age rank on victimization at school". *Labour Economics* 62, p. 101772.

Battu, Harminder, McDonald Mwale, and Zenou Yves. (2007). "Oppositional Identities and the Labor Market". *Journal of Population Economics* 20.3, pp. 643–667.

Beattie, Bruce R. and Satheesh Aradhyula (2015). "A Note on Threshold Factor Level(s) and Stone-Geary Technology". *Journal of Agricultural and Applied Economics* 47.4, pp. 482–493.

Becker, Sascha O, Andrea Ichino, and Giovanni Peri (2004). "How Large is the Brain Drain from Italy?" *Giornale degli economisti e annali di economia*, pp. 1–32.

Bertoni, Marco, Giorgio Brunello, and Lorenzo Rocco (2013). "When the cat is near, the mice won't play: The effect of external examiners in Italian schools". *Journal of Public Economics* 104.C, pp. 65–77.

Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, and Yves Zenou (2011). "Formation and Persistence of Oppositional Identities". *European Economic Review* 55.8, pp. 1046–1071.

— (2016). "Bend It Like Beckham: Ethnic Identity and Integration". *European Economic Review* 90.C, pp. 146–164.

Bleakley, Hoyt and Aimee Chin (2004). "Language Skills and Earnings: Evidence from Childhood Immigrants". *The Review of Economics and Statistics* 86.2, pp. 481–496.

— (2008). "What Holds Back the Second Generation?: The Intergenerational Transmission of Language Human Capital Among Immigrants". *Journal of Human Resources* 43.2, pp. 267–298.

— (2010). "Age at Arrival, English Proficiency, and Social Assimilation Among U.S. Immigrants". *American Economic Journal: Applied Economics* 2.1, pp. 165–192.

Bratti, Massimiliano and Chiara Conti (2018). "The effect of immigration on innovation in Italy". *Regional Studies* 52.7, pp. 934–947.

Brell, Courtney, Christian Dustmann, and Ian Preston (2020). "The Labor Market Integration of Refugee Migrants in High-Income Countries". *Journal of Economic Perspectives* 34.1, pp. 94–121.

Brunello, Giorgio, Elisabetta Lodigiani, and Lorenzo Rocco (2020). "Does low skilled immigration increase the education of natives? Evidence from Italian provinces". *Labour Economics* 63, p. 101794.

Buckles, Kasey S. and Daniel M. Hungerman (2013). "Season of Birth and Later Outcomes: Old Questions, New Answers". *Review of Economics and Statistics* 95.3, pp. 711–724.

Caritas, e Fondazione Migrantes (2020). *XXIX Rapporto Immigrazione*.

Chiswick, Barry R. and Paul W. Miller (2010). "Occupational language requirements and the value of English in the US labor market". *Journal of Population Economics* 23.1, pp. 353–372.

— (2015). "International Migration and the Economics of Language". Handbook of the Economics of International Migration, 1.5, pp. 211–269.

Clarke, Andrew and Ingo Isphording (2017). "Language Barriers and Immigrant Health". *Health Economics* 26.6, pp. 765–778.

Cole, Theodor, Erika Siebert-Cole, Diego Medan, Andrey Podgurenko, Menghan Zhang, Eleonora Selvi, Anna Sokolova, Vladimir Godin, Joanna Sato la-Staśkowiak, Sharif Alghazo, Panagiotis Krimpas, Ciler Hatipoglu, Seongha Rhee, Aphiwit Liang-Itsara, and Hiromi Yamamura (2022). "The International "Language Phylogeny Poster" (LangPP) Project: Teaching Tools for a First Overview on the Evolution of Languages".

Constant, Amelie, Liliya Gataullina, and Klaus Zimmermann (2009). "Ethnosizing Immigrants". *Journal of Economic Behavior & Organization* 69.3, pp. 274–287.

Constant, Amelie F. and Klaus F. Zimmermann (2008). "Measuring Ethnic Identity and Its Impact on Economic Behavior". *Journal of the European Economic Association* 6.2-3, pp. 424–433.

Cummins, James (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Multilingual Matters LTD.

Cummins, James and Metro Gulutsan (1974). "Some effects of bilingualism on cognitive functioning". *Bilingualism, biculturalism and education* 129.07-072, :–36.

Cunha, Flavio and James J. Heckman (2007). “The Technology of Skill Formation”. *American Economic Review* 97.2, pp. 31–47.

Cunha, Flavio, James J. Heckman, and Lance Lochner (2006). “Interpreting the Evidence on Life Cycle Skill Formation”. *Handbook of the Economics of Education. In E. Hanushek, S. Machin, and L. Woessmann (eds.)* 1.12, pp. 697–812.

De Martì Beltran, Joan and Yves. Zenou (2017). “Segregation in Friendship Networks”. *Scandinavian Journal of Economics* 119.8, pp. 656–708.

Del Boca, Daniela and Alessandra Venturini (2005). *Italian migration*. In European migration: What do we know, Oxford University Press, pp. 303–336.

Dell’Anna, Silvia (2021). *Modelli di valutazione di un sistema scolastico inclusivo*. Franco Angeli Open Access.

Dell’Aringa, Carlo, Claudio Lucifora, and Laura Pagani (2015). “Earnings differentials between immigrants and natives: the role of occupational attainment”. *IZA Journal of Migration* 4, pp. 1–18.

Dillender, Marcus (2017). “English Skills and the Health Insurance Coverage of Immigrants”. *American Journal of Health Economics* 3.3, pp. 312–345.

Doepke, Matthias and Fabrizio Zilibotti (2019). “Love, Money, and Parenting: How Economics Explains the Way We Raise Our Kids”. *Princeton University Press*.

Duncan, Greg J. and Richard. Murnane (2011). “Restoring Opportunity: The Crisis of Inequality and the Challenge for American Education”. *Russell Sage, NY*.

Dustmann, Christian and Francesca Fabbri (2003). “Language proficiency and labour market performance of immigrants in the UK”. *The Economic Journal* 113, pp. 695–717.

Ehrlich, Isaac and Jinyoung Kim (2015). “Immigration, human capital formation, and endogenous economic growth”. *Journal of Human Capital* 9.4, pp. 518–563.

Ehrlich, Isaac and Yun Pei (2021). “Endogenous immigration, human and physical capital formation, and the immigration surplus”. *Journal of Human Capital* 15.1, pp. 34–85.

Eurostat (2023). *Early Leavers from Education and Training, in Education and Training in the EU —Facts and Figures*. Online publication: ISSN 2443-8219.

Everett, Caleb, Damián E. Blasí, and Seán G. Roberts (2015). "Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots". *Proceedings of the National Academy of Sciences* 112.5, pp. 1322–1327.

— (2016). "Language evolution and climate: the case of desiccation and tone". *Journal of Language Evolution* 1.1, pp. 33–46.

Fearon, James D (2003). "Ethnic and cultural diversity by country". *Journal of economic growth*, pp. 195–222.

Fryer Jr., Roland G. and Paul Torelli (2010). "An Empirical Analysis of "Acting White"". *Journal of Public Economics* 94.5-6, pp. 380–396.

Galloway, Taryn Ann and Hege Marie Gjefsen (2020). "Assimilation of immigrants: Does earlier school exposure matter?" *Economics of Education Review* 76, p. 101976.

Ginsburgh, Victor and Shlomo Weber (2020). "The economics of language". *Journal of Economic Literature* 58.2, pp. 348–404.

Guven, Cahit and Asadul Islam (2015). "Age at Migration, Language Proficiency, and Socioeconomic Outcomes: Evidence From Australia". *Demography* 52, pp. 513–542.

INVALSI (2018). *Quadro di Riferimento delle Prove INVALSI di Italiano del 30.08.2018*. INVALSI documentation.

Isphording, Ingo and Sebastian Otten (2013). "The Costs of Babylon—Linguistic Distance in Applied Economics". *Review of International Economics* 21.2, pp. 354–369.

Isphording, Ingo E., Marc Piopiunik, and Nuria Rodriguez-Planas (2016). "Speaking in Numbers: The Effect of Reading Performance on Math Performance among Immigrants". *Economics Letters* 139, pp. 52–56.

ISTAT (2022). "Population and Households, Foreigners and Immigrants, Resident Foreigners on 1st January -Citizenship". *I.Stat, your direct access to the Italian Statistics.* Available at http://dati.istat.it/?lang=en.

Jaccard, James and Robert Turrisi (2003). *Interaction Effects in Multiple Regression, (*$2^{nd}ed$*)*. 07-072. Sage University Papers Series on Quantitative Applications in the Social Sciences.

Jacob, Brian A. and Steven D. Levitt (2003). "Rotten apples: an investigation of the prevalence and predictors of teacher cheating". *Quarterly Journal of Economics* 118.3, pp. 843–877.

Karadja, Mounir and Erik Prawitz (2019). "Exit, Voice and Political Change: Evidence from Swedish Mass Migration to the United States". *Journal of Political Economy* 127.4, pp. 1864–1925.

Laitin, David D (2000). "What is a language community?" *American Journal of political science*, pp. 142–155.

Lucifora, Claudio and Marco Tonello (2015). "Cheating and social interactions. Evidence from a randomized experiment in a national evaluation program". *Journal of Economic Behavior & Organization* 115.C, pp. 45–66.

— (2020). "Monitoring and sanctioning cheating at school: What works? Evidence from a national evaluation program". *Journal of Human Capital* 14.4, pp. 584–616.

MLPS (2021a). *La comunità Albanese in Italia*. Ministero del Lavoro e delle Politiche Sociali.

— (2021b). *La comunità Cinese in Italia*. Ministero del Lavoro e delle Politiche Sociali.

— (2021c). *La comunità Filippina in Italia*. Ministero del Lavoro e delle Politiche Sociali.

— (2021d). *La comunità Indiana in Italia*. Ministero del Lavoro e delle Politiche Sociali.

— (2021e). *La comunità Marocchina in Italia*. Ministero del Lavoro e delle Politiche Sociali.

— (2021f). *La comunità Senegalese in Italia*. Ministero del Lavoro e delle Politiche Sociali.

OCDE, OECD (2015). *Immigrant students at school: Easing the journey towards integration*. Organisation for Economic Co-operation and Development.

Paccagnella, Marco and Paolo Sestito (2014). "School cheating and social capital". *Education Economics* 22.4, pp. 367–388.

Pereda-Fernández, Santiago (2019). "Teachers and Cheaters: Just an Anagram?" *Journal of Human Capital* 13.4, pp. 635–669.

Quintano, Claudio, Rosalia Castellano, and Sergio Longobardi (2009). "A fuzzy clustering approach to improve the accuracy of Italian student data : an experimental procedure to correct the impact of outliers on assessment test scores". *Statistica & Applicazioni. Milano : Vita e Pensiero* VII.2, pp. 149–171.

Roodman, David (2009). "How to do xtabond2: An introduction to difference and system GMM in Stata". *The stata journal* 9.1, pp. 86–136.

Signorotto, Cora (2015). *Essays on international migration*. PhD dissertation, University of Milan, pp. 105–150.

Toukomaa, Pertti and Tove Skutnabb-Kangas (1977). *The intensive teaching of the mother tongue to migrant children at pre-school age*. Department of sociology, University of Tampere.

Van Kippersluis, Hans and Cornelius A Rietveld (2018). “Beyond plausibly exogenous”. *The Econometrics Journal* 21.3, pp. 316–331.

Venturini, Alessandra and Claudia Villosio (2008). “Labour-market assimilation of foreign workers in Italy”. *Oxford Review of Economic Policy* 24.3, pp. 517–541.

Wilkinson, Louise C. (2019). “Learning language and mathematics: A perspective from Linguistics and Education”. *Linguistics and Education* 49, pp. 86–95.

Table 7: OLS

Variable	Math (1)	Math (2)	Math (3)	Math (4)	Math (5)	Math (6)	Math (7)
Italian	0.609***	0.642***	0.642***	0.626***	0.620***	0.620***	0.619***
	(0.009)	(0.003)	(0.013)	(0.003)	(0.004)	(0.004)	(0.004)
Age in Months	0.214***	0.187***	0.187***	0.186***	0.194***	0.194***	0.187***
	(0.052)	(0.012)	(0.014)	(0.013)	(0.015)	(0.014)	(0.018)
Distance	0.008	0.008***	0.008***	0.007***	0.007***	0.007***	0.007***
	(0.020)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
Obs	136,019	102,298	102,298	96,654	78,655	78,655	78,655

Notes. This table reports the OLS estimates of the relation between Math and Italian.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 8: Reduced Form

Variable	Math (1)	Math (2)	Math (3)	Math (4)	Math (5)	Math (6)	Math (7)
Age in Months	0.157***	0.159***	0.159***	0.169***	0.166***	0.166***	0.001
	(0.028)	(0.022)	(0.032)	(0.024)	(0.027)	(0.027)	(0.030)
Distance	-0.143**	-0.149***	-0.149**	-0.130***	-0.174***	-0.174***	-0.231***
	(0.071)	(0.041)	(0.053)	(0.044)	(0.050)	(0.050)	(0.050)
Age*Distance	0.001*	0.001***	0.001*	0.001**	0.001***	0.001***	0.002***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
Obs	136,019	102,298	102,298	96,654	78,655	78,655	78,655

Notes. This table reports the reduced form estimates.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 9: IV First Stage (All)

Variable	Italian (1)	Italian (2)	Italian (3)	Italian (4)	Italian (5)	Italian (6)	Italian (7)
Age*Distance	-0.004***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.002***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age in Months	0.185***	0.213***	0.213***	0.209***	0.206***	0.206***	-0.029
	(0.027)	(0.021)	(0.021)	(0.023)	(0.027)	(0.027)	(0.031)
Distance	0.438***	0.398***	0.398***	0.365***	0.329***	0.329***	0.246***
	(0.116)	(0.041)	(0.043)	(0.045)	(0.052)	(0.051)	(0.051)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	16.816	122.562	105.980	84.554	52.697	54.964	33.471
Obs	136,019	102,298	102,298	96,654	78,655	78,655	78,655

Notes. This table reports the first stage of the 2SLS estimates.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 10: IV Second Stage (All)

Variable	Math (1)	Math (2)	Math (3)	Math (4)	Math (5)	Math (6)	Math (7)
Italian	-0.238**	-0.278**	-0.278**	-0.262**	-0.397**	-0.397**	-0.708***
	(0.120)	(0.108)	(0.109)	(0.126)	(0.172)	(0.171)	(0.263)
Age in Months	0.202***	0.218***	0.218***	0.224***	0.248***	0.248***	-0.019
	(0.020)	(0.018)	(0.019)	(0.020)	(0.025)	(0.025)	(0.053)
Distance	-0.039*	-0.038***	-0.038***	-0.034***	-0.043***	-0.043***	-0.057***
	(0.020)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.013)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	16.816	122.562	105.980	84.554	52.697	54.964	33.471
Obs	136,019	102,298	102,298	96,654	78,655	78,655	78,655

Notes. This table reports the second stage of the 2SLS estimates.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 11: IV First Stage (Below Sufficiency)

Variable	Italian (1)	Italian (2)	Italian (3)	Italian (4)	Italian (5)	Italian (6)	Italian (7)
Age*Distance	-0.004***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.002***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age in Months	0.177***	0.199***	0.199***	0.189***	0.186***	0.186***	-0.047
	(0.027)	(0.023)	(0.023)	(0.026)	(0.030)	(0.030)	(0.034)
Distance	0.431***	0.376***	0.376***	0.332***	0.307***	0.307***	0.224***
	(0.118)	(0.043)	(0.045)	(0.048)	(0.056)	(0.055)	(0.055)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	15.778	99.348	87.541	62.701	40.831	42.023	24.838
Obs	123,952	92,965	92,965	87,245	69,773	69,773	69,773

Notes. This table reports the first stage of the 2SLS estimates on the sub-sample of children below the sufficiency threshold.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 12: IV Second Stage (Below)

Variable	Math (1)	Math (2)	Math (3)	Math (4)	Math (5)	Math (6)	Math (7)
Italian	-0.271**	-0.321***	-0.321***	-0.329**	-0.452**	-0.452**	-0.798**
	(0.132)	(0.123)	(0.124)	(0.153)	(0.200)	(0.201)	(0.320)
Age in Months	0.196***	0.214***	0.214***	0.219***	0.237***	0.237***	-0.045
	(0.022)	(0.018)	(0.020)	(0.022)	(0.027)	(0.026)	(0.066)
Distance	-0.044**	-0.041***	-0.041***	-0.039***	-0.046***	-0.046***	-0.063***
	(0.020)	(0.007)	(0.007)	(0.007)	(0.010)	(0.010)	(0.016)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	15.778	99.348	87.541	62.701	40.831	42.023	24.838
Obs	123,952	92,965	92,965	87,245	69,773	69,773	69,773

Notes. This table reports the second stage of the 2SLS estimates on the sub-sample of children below the sufficiency threshold.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 13: IV First Stage (Above Sufficiency)

Variable	Italian (1)	Italian (2)	Italian (3)	Italian (4)	Italian (5)	Italian (6)	Italian (7)
Age*Distance	-0.002	-0.004*	-0.004**	-0.001	0.002	0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.004)	(0.007)	(0.007)	(0.007)
Age in Months	0.221***	0.285***	0.285***	0.285***	0.139	0.139	-0.126
	(0.042)	(0.049)	(0.049)	(0.085)	(0.141)	(0.135)	(0.165)
Distance	0.205	0.533*	0.533*	0.180	-0.202	-0.202	-0.088
	(0.300)	(0.295)	(0.289)	(0.489)	(0.870)	(0.860)	(0.893)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	0.520	3.736	3.916	0.155	0.072	0.074	0.017
Obs	12,067	9,333	9,333	4,459	1,854	1,854	1,854

Notes. This table reports the first stage of the 2SLS estimates on the sub-sample of children above the sufficiency threshold.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 14: IV Second Stage (Above)

Variable	Math (1)	Math (2)	Math (3)	Math (4)	Math (5)	Math (6)	Math (7)
Italian	-0.803	-0.170	-0.170	-0.690	4.577	4.577	7.865
	(2.390)	(0.593)	(0.581)	(4.222)	(15.122)	(14.947)	(55.786)
Age in Months	0.387	0.262	0.262*	0.496	-0.524	-0.524	0.822
	(0.528)	(0.162)	(0.159)	(1.180)	(2.318)	(2.273)	(6.950)
Distance	-0.063*	-0.030	-0.030	-0.029	-0.122	-0.122	-0.214
	(0.034)	(0.025)	(0.025)	(0.061)	(0.482)	(0.476)	(1.658)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	0.520	3.736	3.916	0.155	0.072	0.074	0.017
Obs	12,067	9,333	9,333	4,459	1,854	1,854	1,854

Notes. This table reports the second stage of the 2SLS estimates on the sub-sample of children above the sufficiency threshold.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 15: IV First Stage (Non Romance)

Variable	Italian (1)	Italian (2)	Italian (3)	Italian (4)	Italian (5)	Italian (6)	Italian (7)
Age*Distance	-0.060***	-0.061***	-0.061***	-0.055***	-0.050***	-0.050***	-0.035***
	(0.012)	(0.007)	(0.007)	(0.008)	(0.010)	(0.010)	(0.010)
Age in Months	5.698***	5.855***	5.855***	5.261***	4.789***	4.789***	3.186***
	(1.109)	(0.636)	(0.675)	(0.796)	(1.005)	(0.989)	(0.991)
Distance	7.277***	7.516***	7.516***	6.741***	6.152***	6.152***	4.265***
	(1.455)	(0.850)	(0.903)	(1.064)	(1.342)	(1.318)	(1.317)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	26.785	87.743	77.375	44.590	23.536	24.355	12.305
Obs	52,692	39,405	39,405	33,672	21,597	21,597	21,597

Notes. This table reports the first stage of the 2SLS estimates on the sub-sample of Non Romance children.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 16: IV Second Stage (Non Romance)

Variable	Math (1)	Math (2)	Math (3)	Math (4)	Math (5)	Math (6)	Math (7)
Italian	-0.501***	-0.465***	-0.465***	-0.399**	-0.692**	-0.692**	-1.201**
	(0.192)	(0.139)	(0.146)	(0.190)	(0.306)	(0.298)	(0.551)
Age in Months	0.130***	0.174***	0.174***	0.217***	0.154***	0.154***	-0.219
	(0.035)	(0.032)	(0.033)	(0.036)	(0.058)	(0.058)	(0.176)
Distance	0.162	0.259***	0.259***	0.296***	0.262*	0.262**	0.084
	(0.315)	(0.075)	(0.081)	(0.091)	(0.136)	(0.133)	(0.221)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	26.785	87.743	77.375	44.590	23.536	24.355	12.305
Obs	52,692	39,405	39,405	33,672	21,597	21,597	21,597

Notes. This table reports the second stage of the 2SLS estimates on the sub-sample of Non Romance children.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 17: IV First Stage (Romance)

Variable	Italian (1)	Italian (2)	Italian (3)	Italian (4)	Italian (5)	Italian (6)	Italian (7)
Age*Distance	-0.001	-0.001	-0.001	-0.000	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age in Months	0.154***	0.180***	0.180***	0.171***	0.176***	0.176***	-0.036
	(0.021)	(0.021)	(0.021)	(0.024)	(0.031)	(0.031)	(0.038)
Distance	0.122	0.069	0.069	0.023	0.113	0.113	0.139
	(0.097)	(0.086)	(0.085)	(0.096)	(0.130)	(0.128)	(0.127)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	1.581	0.831	0.864	0.151	1.018	1.043	1.488
Obs	83,327	62,893	62,893	56,566	39,939	39,939	39,939

Notes. This table reports the first stage of the 2SLS estimates on the sub-sample of Romance children.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 18: IV Second Stage (Romance)

Variable	Math (1)	Math (2)	Math (3)	Math (4)	Math (5)	Math (6)	Math (7)
Italian	0.028	-0.041	-0.041	-1.756	-0.870	-0.870	-0.607
	(0.816)	(1.185)	(1.159)	(6.526)	(1.686)	(1.669)	(1.207)
Age in Months	0.164	0.176	0.176	0.471	0.322	0.322	-0.054
	(0.113)	(0.203)	(0.198)	(1.090)	(0.275)	(0.272)	(0.083)
Distance	-0.011	-0.012	-0.012	-0.043	-0.033	-0.033	-0.026
	(0.031)	(0.012)	(0.012)	(0.095)	(0.030)	(0.030)	(0.020)
Controls	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✗	✓	✓	✓	✗	✗	✗
Class FE	✗	✗	✗	✓	✗	✗	✗
Class-Cohort FE	✗	✗	✗	✗	✓	✗	✗
School-Class-Cohort FE	✗	✗	✗	✗	✗	✓	✓
Month FE	✗	✗	✗	✗	✗	✗	✓
First Stage F statistic	1.581	0.831	0.864	0.151	1.018	1.043	1.488
Obs	83,327	62,893	62,893	56,566	39,939	39,939	39,939

Notes. This table reports the second stage of the 2SLS estimates on the sub-sample of Romance children.
Model (1): standard errors clustered at the language-wave level.
Model (2), (6), and (7): standard errors clustered at the school-class-cohort level.
Model (3): standard errors multi-way-clustered at the school, class and cohort level.
Model (4): standard errors clustered at the school level.
Model (5): standard errors clustered at the school-cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 19: Discrimination, Socialization and Self-Segregation

Variable	Teased (1)	Insulted (2)	Isolated (3)	Beaten (4)	Self-Seg (5)
Age*Distance	-0.022	0.003	-0.011	0.020	0.013
	(0.034)	(0.030)	(0.028)	(0.020)	(0.020)
Age in Months	-0.001	-0.002	-0.000	-0.001	0.000
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Distance	3.415	-0.108	1.738	-2.403	-1.608
	(4.384)	(3.954)	(3.632)	(2.568)	(2.649)
Controls	✓	✓	✓	✓	✓
School-Class FE	✓	✓	✓	✓	✓
Obs	15,172	15,147	15,176	15,200	15,186

Notes. This table reports the LPM estimates of the relation between a set of outcomes describing discrimination, socialization and self-segregation, and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 20: Home Possessions

Variable	Quiet Place (1)	Computer (2)	Desk (3)	Encyclopedias (4)	Internet (5)	Room (6)
Age*Distance	-0.007	-0.011	-0.010	-0.008	-0.035***	0.006
	(0.013)	(0.017)	(0.013)	(0.017)	(0.013)	(0.017)
Age in Months	0.002	0.003**	0.000	0.001	0.001*	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Distance	0.633	1.040	1.099	0.407	4.059**	-1.200
	(1.744)	(2.172)	(1.642)	(2.253)	(1.678)	(2.251)
Controls	✓	✓	✓	✓	✓	✓
School-Class FE	✓	✓	✓	✓	✓	✓
Obs	79,673	79,575	79,811	79,279	79,566	79,579

Notes. This table reports the LPM estimates of the relation between a set of outcomes describing Home Possessions, and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 21: Home Possessions: Books

Variable	Books (0-10) (1)	Books (11-25) (2)	Books (26-100) (3)	Books (101-200) (4)	Books (> 200) (5)
Age*Distance	0.019	-0.005	-0.002	-0.008	-0.004
	(0.015)	(0.017)	(0.016)	(0.010)	(0.006)
Age in Months	0.000	0.001	-0.001	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Distance	-1.735	0.793	-0.343	0.898	0.388
	(1.918)	(2.204)	(2.021)	(1.259)	(0.828)
Controls	✓	✓	✓	✓	✓
School-Class FE	✓	✓	✓	✓	✓
Obs	79,615	79,615	79,615	79,615	79,615

Notes. This table reports the LPM estimates of the relation between the number of books at home, and the instrumental variable. Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 22: Father Education

Variable	ISCED-1 (1)	ISCED-2 (2)	ISCED-3 (3)	ISCED-4 (4)	ISCED-5 (5)	ISCED-6 (6)
Age*Distance	0.015	0.017	0.005	-0.021	-0.008	-0.008
	(0.010)	(0.017)	(0.011)	(0.015)	(0.007)	(0.009)
Age in Months	0.000	0.002	0.000	-0.001	0.000	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Distance	-1.559	-1.543	-0.913	2.059	0.983	0.973
	(1.244)	(2.175)	(1.415)	(1.980)	(0.866)	(1.154)
Controls	✓	✓	✓	✓	✓	✓
School-Class FE	✓	✓	✓	✓	✓	✓
Obs	90,994	90,994	90,994	90,994	90,994	90,994

Notes. This table reports the LPM estimates of the relation between father education and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 23: Father Occupation

Variable	HISEI-1 (1)	HISEI-2 (2)	HISEI-3 (3)	HISEI-4 (4)	HISEI-5 (5)
Age*Distance	-0.014	0.001	-0.002	0.007	-0.001
	(0.009)	(0.003)	(0.002)	(0.005)	(0.006)
Age in Months	0.001	0.000	0.000	-0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Distance	1.924	-0.129	0.261	-0.888	0.106
	(1.204)	(0.375)	(0.275)	(0.596)	(0.756)
Controls	✓	✓	✓	✓	✓
School-Class FE	✓	✓	✓	✓	✓
Obs	97,522	97,522	97,522	97,522	97,522

Notes. This table reports the LPM estimates of the relation between father occupation and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 24: Father Occupation (continued)

Variable	HISEI-6 (1)	HISEI-7 (2)	HISEI-8 (3)	HISEI-9 (4)	HISEI-10 (5)
Age*Distance	0.029**	0.006	-0.029*	0.001	0.002
	(0.012)	(0.005)	(0.015)	(0.002)	(0.007)
Age in Months	0.000	-0.001**	-0.000	-0.000	0.000
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
Distance	-3.317**	-0.946	3.412*	-0.109	-0.314
	(1.596)	(0.684)	(1.978)	(0.313)	(0.871)
Controls	✓	✓	✓	✓	✓
School-Class FE	✓	✓	✓	✓	✓
Obs	97,522	97,522	97,522	97,522	97,522

Notes. This table reports the LPM estimates of the relation between father occupation and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 25: Mother Education

Variable	ISCED-1 (1)	ISCED-2 (2)	ISCED-3 (3)	ISCED-4 (4)	ISCED-5 (5)	ISCED-6 (6)
Age*Distance	0.005	0.020	0.009	-0.001	-0.009	-0.025***
	(0.010)	(0.017)	(0.009)	(0.015)	(0.007)	(0.010)
Age in Months	0.001	0.002*	-0.001	-0.002**	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Distance	-0.264	-1.888	-1.403	-0.675	1.122	3.108**
	(1.256)	(2.154)	(1.232)	(1.943)	(0.871)	(1.249)
Controls	✓	✓	✓	✓	✓	✓
School-Class FE	✓	✓	✓	✓	✓	✓
Obs	91,845	91,845	91,845	91,845	91,845	91,845

Notes. This table reports the LPM estimates of the relation between mother education and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 26: Mother Occupation

Variable	HISEI-1 (1)	HISEI-2 (2)	HISEI-3 (3)	HISEI-4 (4)	HISEI-5 (5)
Age*Distance	-0.009	-0.052***	0.000	0.004	-0.002
	(0.008)	(0.015)	(0.001)	(0.003)	(0.005)
Age in Months	0.000	0.001	-0.000	-0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Distance	1.107	7.432***	-0.051	-0.532	0.230
	(1.072)	(1.933)	(0.178)	(0.379)	(0.617)
Controls	✓	✓	✓	✓	✓
School-Class FE	✓	✓	✓	✓	✓
Obs	97,522	97,522	97,522	97,522	97,522

Notes. This table reports the LPM estimates of the relation between mother occupation and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 27: Mother Occupation (continued)

Variable	HISEI-6 (1)	HISEI-7 (2)	HISEI-8 (3)	HISEI-9 (4)	HISEI-10 (5)
Age*Distance	0.047***	0.003	0.007	0.001	-0.000
	(0.008)	(0.005)	(0.013)	(0.001)	(0.006)
Age in Months	-0.001	-0.001	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Distance	-5.820***	-0.539	-1.633	-0.116	0.022
	(1.049)	(0.675)	(1.710)	(0.169)	(0.744)
Controls	✓	✓	✓	✓	✓
School-Class FE	✓	✓	✓	✓	✓
Obs	97,522	97,522	97,522	97,522	97,522

Notes. This table reports the LPM estimates of the relation between mother occupation and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 28: Teacher Attitude

Variable	Being heard (1)	Pay attention (2)	Encouraged (3)
Age*Distance	-0.045	-0.062	-0.035
	(0.034)	(0.039)	(0.042)
Age in Months	0.003	0.004	0.002
	(0.002)	(0.003)	(0.003)
Distance	5.611	7.823	4.284
	(4.464)	(5.060)	(5.514)
Controls	✓	✓	✓
School-Class FE	✓	✓	✓
Obs	15,095	15,076	15,094

Notes. This table reports the LPM estimates of the relation between a series of outcomes indicating the teacher attitude and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 29: Preferences

Variable	Math Pref (1)	Italian Pref (2)
Age*Distance	0.055*	-0.025
	(0.031)	(0.030)
Age in Months	-0.003	0.002
	(0.002)	(0.002)
Distance	-6.926*	3.216
	(3.976)	(3.959)
Controls	✓	✓
School-Class FE	✓	✓
Obs	20,341	20,262

Notes. This table reports the LPM estimates of the relation between preferences for Math and Italian and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 30: Postponement

Variable	Postponement (1)
Distance	0.267*** (0.029)
Controls	✓
School-Class FE	✓
Observations	102,298

Notes. This table reports the LPM estimates of the relation between enrollment postponement and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 31: Socio-emotional skills

Variable	Worried (1)	Nervous (2)	Wrong (3)	Calm (4)
Age*Distance	0.001	0.011	-0.010	-0.033
	(0.021)	(0.017)	(0.022)	(0.021)
Age in Months	-0.001	-0.000	-0.002	0.003*
	(0.001)	(0.001)	(0.002)	(0.001)
Distance	0.096	-0.998	1.745	3.986
	(2.701)	(2.278)	(2.801)	(2.768)
Controls	✓	✓	✓	✓
Cohort FE	✗	✓	✗	✓
School FE	✗	✓	✗	✗
School-Class-Cohort FE	✓	✗	✓	✗
Obs	55,337	55,311	55,161	55,130

Notes. This table reports the LPM estimates of the relation between a series of outcomes indicating socio-emotional skills and the instrumental variable.
The distance variable has been re-scaled by a factor of 1000 to improve readability. Thus, in this table, it ranges from a minimum distance of 0.05877 (Romanian-Italian) to a maximum distance of 0.10114 (Chinese-Italian).
Standard errors are clustered at the school and class level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 32: Zero-first-stage Test

Variable	First-Stage Italian	Second-Stage Math	Reduced Form Math
Age*Distance	-0.000		0.001
	(0.000)		(0.001)
Age in Months	0.041***	0.303	0.095***
	(0.014)	(0.238)	(0.020)
Distance	0.029	-0.016	-0.160
	(0.031)	(0.017)	(0.084)
Italian		-5.037	
		(7.024)	
Controls	✓	✓	✓
Cohort FE	✓	✓	✓
Obs	9,366	9,366	9,366

Notes. This table reports the first stage, the second stage of the 2SLS estimates, and the reduced form estimation for the zero-first stage group. In the models standard errors are multi-way-clustered at the school, class and cohort level.
Controls: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 33: OLS

Variable	Math (1)	Math (2)	Math (3)	Math (4)
Italian	0.642***	0.645***	0.648***	0.644***
	(0.013)	(0.013)	(0.013)	(0.013)
Age in Months	0.187***	0.184***	0.176***	0.183***
	(0.014)	(0.014)	(0.013)	(0.014)
Distance	0.008***	0.021***	3.990***	2.093***
	(0.002)	(0.003)	(0.393)	(0.252)
Controls	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓
Obs	102,298	102,298	102,298	102,298

Notes. This table reports the OLS estimates of the relation between Math and Italian, using different measures of linguistic distance.
Model (1): ASJP.
Model (2): ethnolinguistic.net
Model (3): cladistic distance computed from Cole et al. (2022), according to Laitin (2000).
Model (4): square root of the previous measure (Fearon, 2003).
Standard errors are multi-way-clustered at the school, class and cohort level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 34: Reduced Form

Variable	Math (1)	Math (2)	Math (3)	Math (4)
Age*Distance	0.001*	0.002**	0.247***	0.172**
	(0.000)	(0.000)	(0.046)	(0.045)
Age in Months	0.159***	0.147***	0.133***	0.139***
	(0.032)	(0.029)	(0.024)	(0.028)
Distance	-0.149**	-0.238**	-34.065***	-24.565**
	(0.053)	(0.057)	(5.534)	(5.542)
Controls	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓
Obs	102,298	102,298	102,298	102,298

Notes. This table reports the OLS estimates of the relation between Math and Italian, using different measures of linguistic distance.
Model (1): ASJP.
Model (2): ethnolinguistic.net
Model (3): cladistic distance computed from Cole et al. (2022), according to Laitin (2000).
Model (4): square root of the previous measure (Fearon, 2003).
Standard errors are multi-way-clustered at the school, class and cohort level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

Table 35: IV First Stage (All)

Variable	Italian (1)	Italian (2)	Italian (3)	Italian (4)
Age*Distance	-0.003***	-0.005***	-0.568***	-0.457***
	(0.000)	(0.000)	(0.049)	(0.042)
Age in Months	0.213***	0.220***	0.225***	0.229***
	(0.021)	(0.020)	(0.020)	(0.021)
Distance	0.398***	0.540***	65.338***	53.092***
	(0.043)	(0.053)	(6.416)	(5.394)
Controls	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓
First Stage F statistic	105.980	129.035	131.924	120.692
Observations	102,298	102,298	102,298	102,298

Notes. This table reports the OLS estimates of the relation between Math and Italian, using different measures of linguistic distance.
Model (1): ASJP.
Model (2): ethnolinguistic.net
Model (3): cladistic distance computed from Cole et al. (2022), according to Laitin (2000).
Model (4): square root of the previous measure (Fearon, 2003).
Standard errors are multi-way-clustered at the school, class and cohort level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
The Hansen J-statistic for the overidentification test between the first and the second instrument is 2.106, with a p-value of 0.1467. The Hansen J-statistic for the overidentification test between the first and the second instrument is 1.292, with a p-value of 0.2558.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; * significant at $p < 0.1$

Table 36: IV Second Stage (All)

Variable	Math (1)	Math (2)	Math (3)	Math (4)
Italian	-0.278**	-0.349***	-0.436***	-0.377***
	(0.109)	(0.103)	(0.107)	(0.108)
Age in Months	0.218***	0.224***	0.231***	0.225***
	(0.019)	(0.020)	(0.021)	(0.020)
Distance	-0.038***	-0.050***	-5.587***	-4.566***
	(0.006)	(0.007)	(0.978)	(0.723)
Controls	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓
Obs	102,298	102,298	102,298	102,298

Notes. This table reports the second stage of the 2SLS etimates of the relation between Math and Italian, using different measures of linguistic distance.
Model (1): ASJP.
Model (2): ethnolinguistic.net
Model (3): cladistic distance computed from Cole et al. (2022), according to Laitin (2000).
Model (4): square root of the previous measure (Fearon, 2003).
Standard errors are multi-way-clustered at the school, class and cohort level. *Controls*: gender, previous enrollment in nursery school, previous enrollment in pre-school, and ESCS index.
*** significant at $p < 0.01$; ** significant at $p < 0.05$; *significant at $p < 0.1$

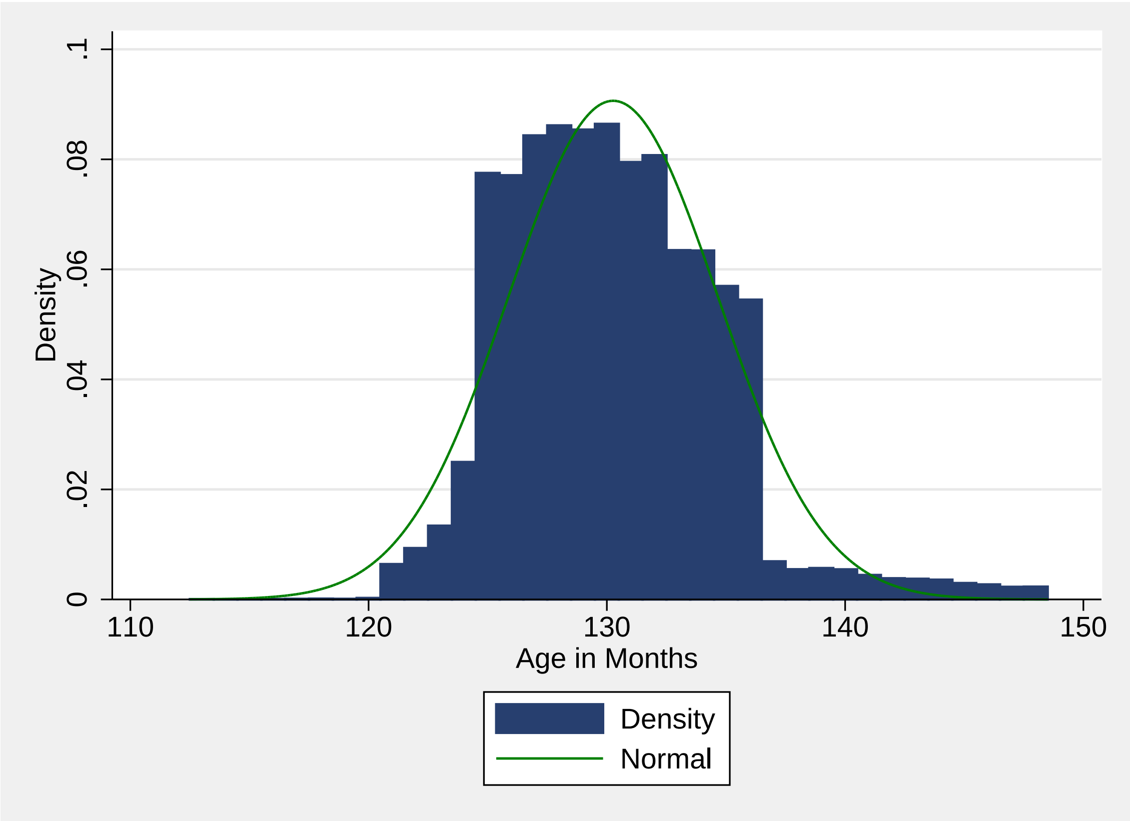
Table 37: Descriptives by wave (Mother characteristics)

	2015		2016		2017		2018		2019	
VARIABLES	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ESCS student	-0.466	0.802	-0.432	0.795	-0.553	0.839	-0.491	0.817	-0.538	0.853
Geographic Origin										
Mother from UE	0.184	0.387	0.211	0.408	0.220	0.414	0.223	0.416	0.221	0.415
Mother from Europe	0.273	0.445	0.249	0.432	0.249	0.432	0.251	0.434	0.254	0.435
Mother from Outside	0.543	0.498	0.540	0.498	0.531	0.499	0.526	0.499	0.526	0.499
Education										
Mother ISCED-1	0.058	0.233	0.059	0.235	0.055	0.228	0.057	0.231	0.055	0.227
Mother ISCED-2	0.300	0.458	0.302	0.459	0.301	0.459	0.315	0.464	0.309	0.462
Mother ISCED-3	0.071	0.257	0.069	0.254	0.075	0.264	0.074	0.261	0.078	0.269
Mother ISCED-4	0.269	0.443	0.276	0.447	0.268	0.443	0.266	0.442	0.256	0.437
Mother ISCED-5	0.033	0.179	0.033	0.179	0.035	0.184	0.034	0.181	0.033	0.178
Mother ISCED-6	0.081	0.272	0.082	0.274	0.077	0.267	0.087	0.282	0.084	0.277
Employment Status										
Mother HISEI-1	0.068	0.252	0.069	0.253	0.067	0.250	0.068	0.252	0.069	0.253
Mother HISEI-2	0.443	0.497	0.455	0.498	0.446	0.497	0.452	0.498	0.439	0.496
Mother HISEI-3	0.002	0.043	0.002	0.043	0.001	0.036	0.002	0.041	0.002	0.044
Mother HISEI-4	0.007	0.082	0.006	0.079	0.006	0.079	0.007	0.081	0.007	0.081
Mother HISEI-5	0.024	0.153	0.025	0.156	0.022	0.147	0.023	0.150	0.023	0.149
Mother HISEI-6	0.051	0.220	0.051	0.220	0.049	0.216	0.050	0.219	0.049	0.216
Mother HISEI-7	0.031	0.172	0.031	0.174	0.031	0.174	0.032	0.177	0.031	0.174
Mother HISEI-8	0.222	0.416	0.216	0.411	0.219	0.414	0.226	0.418	0.222	0.415
Mother HISEI-9	0.002	0.048	0.001	0.036	0.001	0.035	0.001	0.038	0.001	0.035

Table 38: Descriptives by wave (Father characteristics)

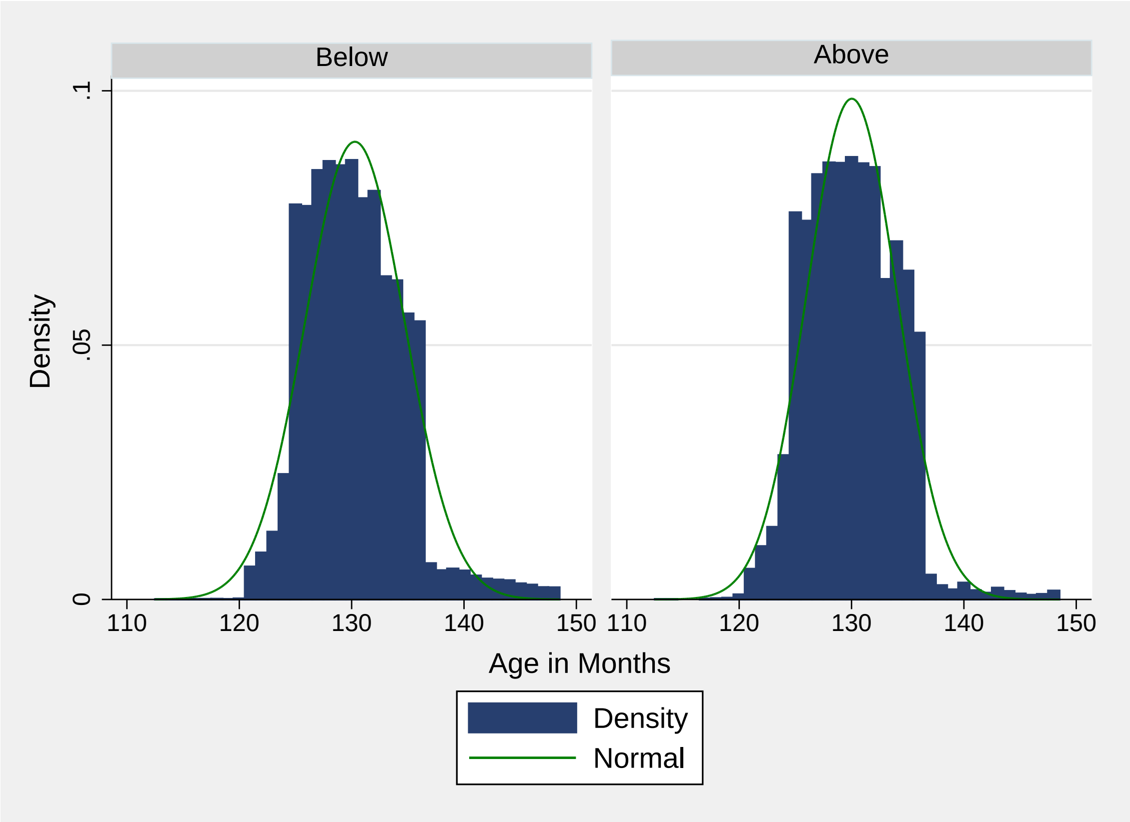
	2015		2016		2017		2018		2019	
VARIABLES	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
ESCS student	-0.466	0.802	-0.432	0.795	-0.553	0.839	-0.491	0.817	-0.538	0.853
Geographic Origin										
Father from UE	0.180	0.384	0.206	0.404	0.213	0.410	0.217	0.412	0.218	0.413
Father from Europe	0.273	0.445	0.249	0.432	0.251	0.434	0.253	0.435	0.254	0.435
Father from Outside	0.547	0.498	0.545	0.498	0.535	0.499	0.530	0.499	0.528	0.499
Education										
Father ISCED-1	0.055	0.228	0.057	0.232	0.0535	0.225	0.052	0.222	0.053	0.225
Father ISCED-2	0.284	0.451	0.293	0.455	0.296	0.456	0.313	0.464	0.308	0.462
Father ISCED-3	0.096	0.295	0.093	0.290	0.097	0.296	0.095	0.293	0.103	0.303
Father ISCED-4	0.273	0.446	0.274	0.446	0.268	0.443	0.269	0.443	0.253	0.435
Father ISCED-5	0.031	0.174	0.031	0.174	0.030	0.172	0.032	0.175	0.027	0.162
Father ISCED-6	0.063	0.243	0.066	0.248	0.060	0.238	0.064	0.245	0.061	0.240
Employment Status										
Father HISEI-1	0.092	0.289	0.084	0.277	0.082	0.275	0.076	0.264	0.070	0.255
Father HISEI-2	0.009	0.097	0.007	0.084	0.008	0.089	0.007	0.084	0.007	0.083
Father HISEI-3	0.004	0.065	0.004	0.065	0.003	0.057	0.003	0.059	0.003	0.054
Father HISEI-4	0.020	0.140	0.019	0.136	0.019	0.137	0.020	0.140	0.020	0.138
Father HISEI-5	0.037	0.190	0.036	0.186	0.034	0.180	0.035	0.183	0.034	0.181
Father HISEI-6	0.157	0.364	0.157	0.364	0.152	0.359	0.155	0.361	0.149	0.356
Father HISEI-7	0.029	0.167	0.031	0.174	0.031	0.172	0.029	0.168	0.028	0.164
Father HISEI-8	0.480	0.500	0.496	0.500	0.496	0.500	0.516	0.500	0.509	0.500
Father HISEI-9	0.005	0.073	0.005	0.067	0.005	0.072	0.005	0.071	0.006	0.076

Figure 3: Histogram (All)



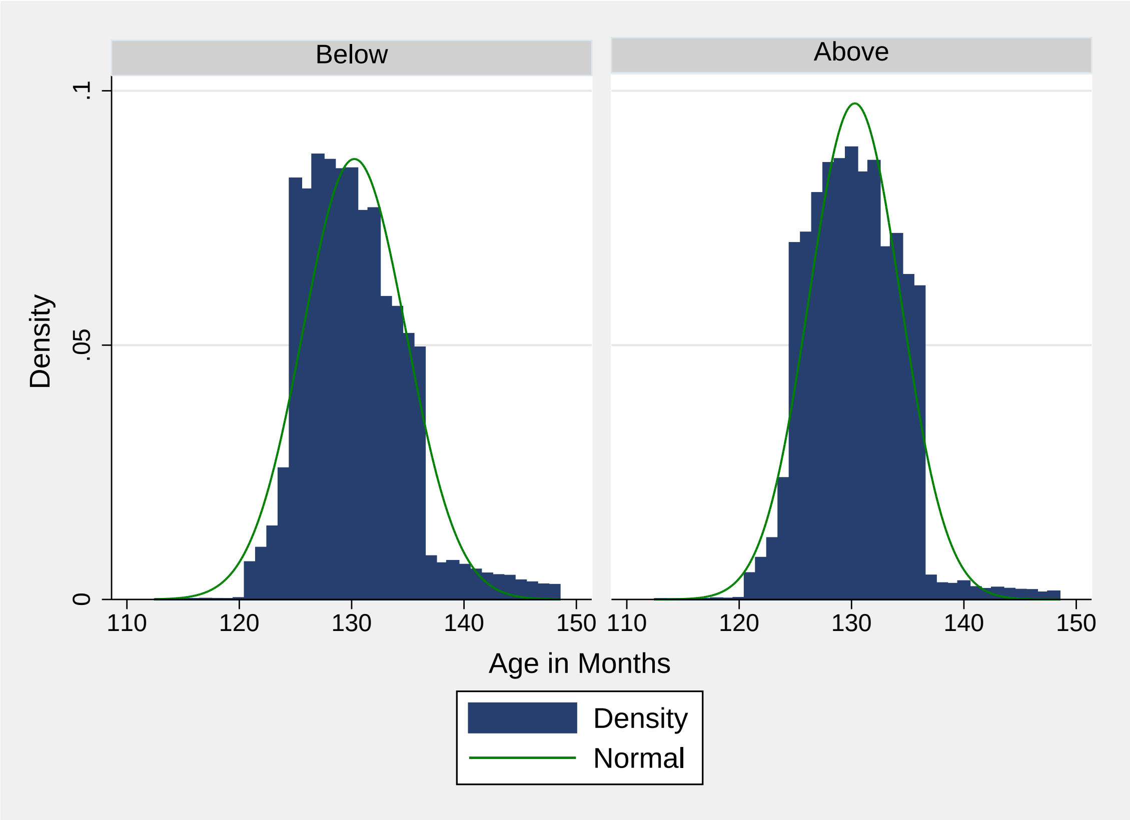
Notes: This histogram shows the variation in age (calculated as age in months) for the entire sample of second-generation immigrant children. Normal density is plotted in green.

Figure 4: Histogram (Below vs. Above the (Predicted) Proficiency Threshold)



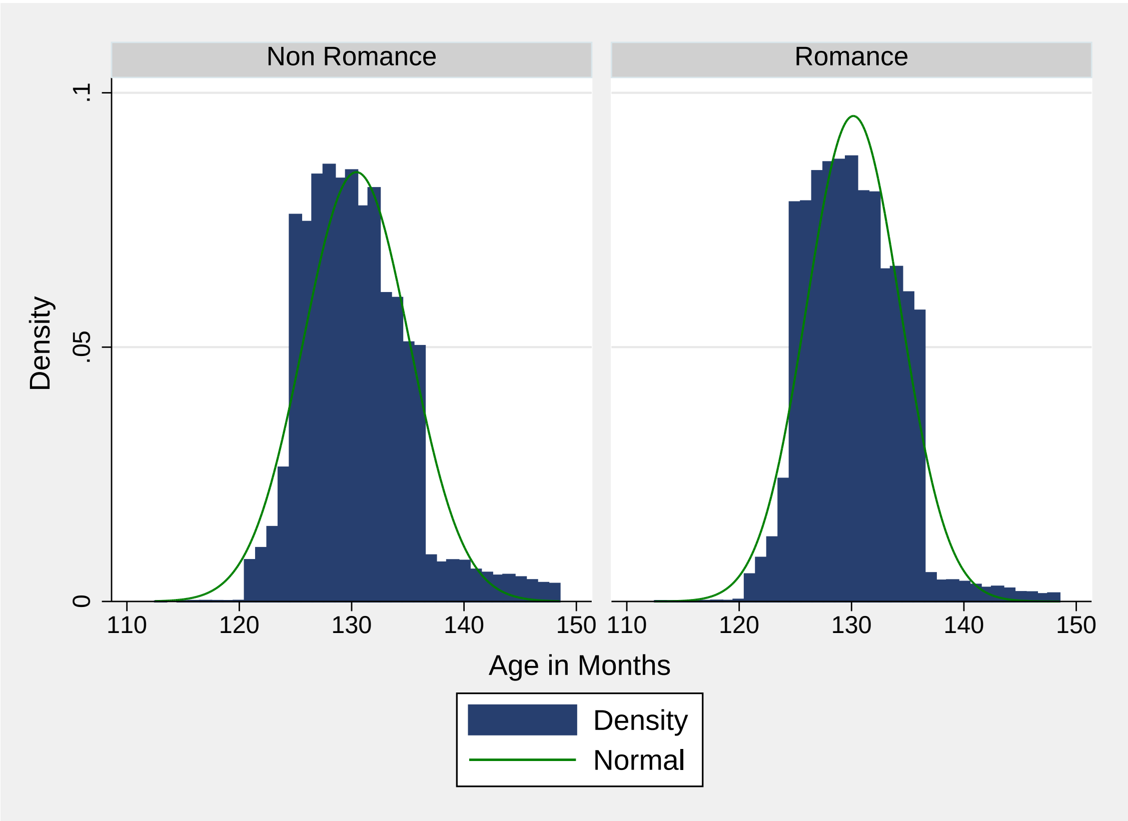
Notes: This histogram shows the variation in age (calculated as age in months) for the subsamples of second-generation immigrant children below and above the (predicted) proficiency threshold. Normal density is plotted in green.

Figure 5: Histogram (Below vs. Above the (Realized) Proficiency Threshold)



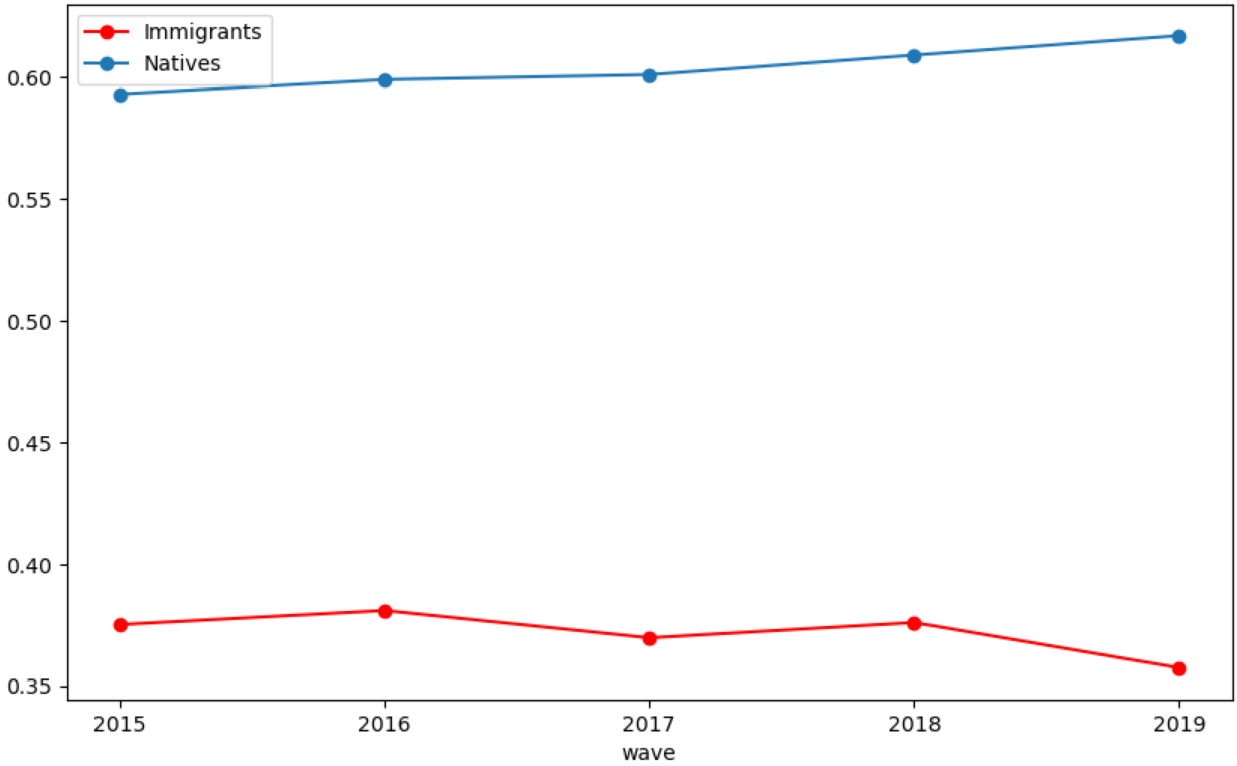
Notes: This histogram shows the variation in age (calculated as age in months) for the subsamples of second-generation immigrant children above and below the (realized) proficiency threshold. Normal density is plotted in green.

Figure 6: Histogram (Romance vs. Non-Romance)



Notes: This histogram shows the variation in age (calculated as age in months) for the subsamples of Romance and Non-Romance second-generation immigrant children. Normal density is plotted in green.

Figure 7: Skill composition of the migrant population vs. total domestic population (weighted averages)



Notes: the weighted average (by population shares) of the skill composition of migrants is measured as the share of all skilled migrants (those with at least upper secondary education) in the total migrant population in our sample (Source: *INVALSI*). The weighted average (by population shares) of the skill composition of natives is measured as the share of the population aged 25-64 years with at least upper secondary education (Source: *Eurostat*).