Length of stay in the host country and educational achievement of immigrant students: the Italian case

Journal:	*International Journal of Manpower*
Manuscript ID:	IJM-11-2013-0261.R2
Manuscript Type:	Research Paper
Keywords:	Immigrant students, educational attainments, age at immigration, Human capital

Length of stay in the host country and educational achievement of immigrant students: the Italian case

1. INTRODUCTION

"In order to close the achievement gap, institutional changes must be made at the school level, including changes in language teaching… Proficiency in the language of instruction is a major tool and precondition for learning."

OECD (2010)

Since the last decade of the 20th century European countries have experienced large waves of migration flows from both within the EU and from outside it. As a consequence, the integration of third-country nationals has been identified as a priority at European level. Migration-related issues are a central part of the Stockholm Programme, adopted by the EU Member State governments at the December 2009 European Council.

Among the proposed indicators of migrant integration, language skills and educational attainments are identified as crucial. Low skilled workers are more at-risk of poverty or social exclusion and young people with a migrant background are found to be at greater risk of dropping out of school and of exiting the education and training system without having obtained an upper secondary qualification. Data are striking: in 2008 regardless of gender, the share of early school leavers with a foreign background was four percentage points higher than that of their counterparts with native-born parents (Eurostat, 2011b). In general, the school performance gap between native and immigrant children is well documented for a number of industrialized countries and it is a real concern for policy makers since it also predicts a gap in labour-market performance and other long-term outcomes.

In this paper we use the language standardized test data provided by INVALSI, the Italian institute in charge of evaluating schools' performance, to analyse the gap of young immigrant children in Italy. In particular, we examine if this gap is significantly influenced by pupils' length of stay in Italy, their age at immigration and their area of origin. Italian data are most suitable for studying these issues. Together with Spain and Greece, Italy is a recent migration destination. As of 31st December 2008, in Italy the percentage of foreigners as a share of the total population was 6.5%, of whom 1.9% were citizens of (other) EU Member States and 4.6% were from non-EU countries. With respect to educational levels, according to PISA 2009 results, Italy has some of the largest native-immigrant school performance gaps among OECD countries.[1] In particular, this result

[1] "In 2009, the reading scores of immigrant students were lower than those of non-immigrant students in 23 out of 28 OECD countries with sufficient data. The performance gap reaches 99 score points in Mexico, more than 80 in Iceland and more than 72 in Italy. In Finland, Austria, Belgium, Sweden, Denmark and

holds for second generation students, even after adjusting for socio-economic background (OECD, 2012). Furthermore, the share of early leavers from education and training (aged 18-24) among the subgroup of foreign-born is 46%, compared to 19% observed for the total population.[2] Finally, among the EU countries, Italy has the lowest proportion of foreign citizens with tertiary education and a large one of low educational attainment level (Eurostat LFS, 2008). In sum, migrants in Italy have a lower level of income and are at increased risk of poverty and social exclusion. With its low educational attainments of both migrants and natives, Italy is among the group of countries most at risk. [3]

In this study we investigate the differences in standardized language test scores between immigrant and Italian children conditional on personal, family and school characteristics, distinguishing between first and second-generation immigrants. Although we do not have longitudinal data, for first-generation immigrants, we are able to study whether the age at immigration influences their academic achievements. In particular, unlike most previous studies in this field, our data allow us to compare the results obtained by children enrolled in different grades, more specifically, the second and fifth year of (primary) school, sixth and tenth year (secondary school).[4] Indeed, since the Italian INVALSI data sample at multiple ages, they permit a broader picture showing if the immigrant students impact of not speaking the language of instruction at home changes across education levels. Comparing the results across the different grades we investigate if the educational gap narrows at a different pace in the early or later years of a student's life. That is, we address the important question of whether the age at arrival and the length of stay in the host country matters for immigrants' educational achievements. Indeed, sociologists have long identified immigrants who arrived as young children, the so called 1.5 generation, as different from those who arrived after and observed that the outcomes of the two groups may significantly differ.[5]

France, the gap is 60 score points or more, the equivalent of over a year and a half of schooling", OECD (2012).

[2] At the European Union level, the share of foreign-born early school leavers aged 18-24 is higher than the share of early school leavers aged 18-24 from all population. The most marked differences in the shares are in Germany, Greece, Spain, France, Italy and Cyprus, Austria and Finland.

[3] In 2008, the at risk of poverty or social exclusion rate among foreign-born persons was highest in Greece (45%), followed by Belgium, Italy and France, the latter with percentage above 30%. See European Union (2011).

[4] We exclude data on the National exam performed by all Italian students at the end of year eight since they are not comparable with those used in this study.

[5] "…first-generation immigrants who arrive at a young age are often more similar to second-generation ones than to adults from the first-generation. Individuals who immigrate to a new country before or during their early teens bring with them characteristics from their home country but experience at least some of their formative years while in a new country". See Sweetman et al. (2014).

We focus on language tests because the lack of language skills plays a crucial role in immigrants' integration processes. Further, we also concentrate on the age of immigrant children and their length of stay in the host country since a large literature suggests that, although people can learn languages at any age, young children have an innate ability to learn the rules of new languages, and this ability tends to decrease by adulthood (Johnson and Newport, 1989).[6] In particular, the recent existing literature on the economics of the language, examine the impact of immigrants' host–country–language ability on long-run economic and social outcomes. For example, Bleakley and Chin (2004, 2010) estimate the causal impact of English-language ability on different outcomes, namely, earnings in the US and on marriage, fertility, and residential location choices. First, they find a significant positive effect of English-language skills on wages among childhood immigrants, where much of this effect appears to be mediated by years of schooling. Second their results also stress that English proficiency affects other outcomes: it raises the probabilities of being divorced, marrying a US native, having a more educated and higher earning spouse, having fewer children, and, for some groups, living outside of ethnic enclaves. Overall, these results indicate that language skills have an important role not only for labour market outcomes but also in the process of social assimilation.

Our analysis is also related to the growing literature that dates back to the end of the nineties and investigates the role of social distance and social capital on economic outcomes (Helliwell and Putnam, 1999). In general, note that social distance is a very broad concept which refers to the cognitive relationship between two cultures that co-exist within an individual, and it is influenced by many factors including the immigrant's length of residence. Glaeser (1999) argues that, together with education levels, one of the factors that appear to be crucial in creating social capital at the community level is ethnic and linguistic heterogeneity.[7] Moreover, according to linguistic scholars, social distance is one of the socio-cultural factors affecting the second language acquisition by immigrants which, in turn, is crucial for their integration in the host country.[8]

This study is structured in six different sections. The next section introduces the literature review, while the third the descriptive analysis. Section 4 discusses the main results and Section 5 the extensions and robustness checks. Conclusions are in section 6.

---

[6] Moreover, children who learn a language before adolescence are more likely than older learners to attain native-like pronunciation (Patkowski, 1990).

[7] Alesina and La Ferrara (2002) also document the positive effects of homogeneity on social participation across American states. They argue that schools are a primary area where social capital is developed.

[8] According to the Schumann's Acculturation Model, social distance explains the acquisition of second language and it is influenced by many factors such as the equality of native and immigrant groups, assimilation or integration, enclosure, cohesiveness, size, cultural congruence, attitude and length of residence. See Schumann (1976).

2. LITERATURE REVIEW

Among economist, only relatively few recent studies address the important question of whether the age at arrival and the length of stay in the host country matters for immigrants' educational achievements. Most early studies focused on traditional immigrant countries, such as the United States and Canada but among more recent studies we also find an increasing number of papers that exploit new immigrant receiving countries datasets. Among the former we include the seminal paper of Borjas (1995) who find that age at immigration could bias estimates of economic integration in cohort models, while subsequent studies by Schaafsma and Sweetman (2001) using Canadian data, and Chiswick and DebBurman (2004) and Gonzalez (2003) using American data, have explicitly addressed issues directly related to age at immigration and educational attainments.

More recently, Böhlmark (2008) exploits a panel of siblings graduating after nine years of schooling in Sweden from 1988 to 2003, to examine the impact of age at immigration on school performance among immigrant ninth-graders upon graduation. The use of siblings allows him to control for likely neighbourhood effects, which constitute a potential source of bias in his cross-sectional estimation analysis. Indeed, immigrant families can be hardly considered randomly assigned to cities and neighbourhoods, or children to schools, and he exploits the siblings strategy to control for any omitted variables capturing time-invariant, family-specific characteristics, assuming that older and younger siblings would have performed equally well in the absence of immigration. Unlike older studies that identify a critical age at seven, Böhlmark (2008) finds that the estimated critical age at immigration is about nine: children arriving in Sweden up to about the third year of school seem to catch up well with their peers who came before preschool age, and this result is stable for both boys and girls. Above the age of 9, however, he finds a strong negative impact on performance.[9] Second, similarly to natives, immigrant girls outperform immigrant boys and, the age-at-immigration performance profiles are similar in shape for children with different parental educational backgrounds. Moreover, he also finds significant differences by geographical origin: the estimated age-at-immigration performance profiles are steepest for Asian children and flattest for Western children.

Ohinata and Van Ours (2012) investigate the determinants of the observed differences in test scores by both first and second-generation immigrants and native Dutch children, conditional on personal and family characteristics and classroom environment. Their sample includes children aged 9 and 10 and, for the first-generation immigrants, they study whether the age at immigration

[9] Cahan et al. (2001) suggest that age 7 may represent a critical age. Similarly, using Norwegian data, Bratsberg et al. (2011) point to age seven as the critical age for non-OECD students.

influences the school outcomes. Their educational attainment outcomes include language, science and math test scores and they find that these are affected by both age at immigration and whether or not one of the parents is native Dutch. However, when exploring whether age at immigration has an effect on the educational attainment of the immigrant children, language results are omitted from the analysis due to shortage of data. Results from science and math test scores suggest that age at immigration is important: the later immigrant children arrived in the Netherlands the lower their educational achievement. Finally, they also find that second-generation immigrants do not have lower language scores compared to native Dutch children irrespective of the origin of their parents. [10]

Among the recent non-European studies we briefly mention two additional analyses. The first is Cortes (2006) who, using educational performance data of children in San Diego and Miami, finds that the gap in test scores between first-generation and second-generation immigrant children decreases the longer the former reside in the United States. Finally, similar to Böhlmark (2008), Corak (2012) analyses high-school dropout rates in Canada and finds that up to age 9 the chances of being a high-school dropout do not vary according to age, but children arriving after that age are more likely not to graduate from high school.

Mostly due to data limitations, cross country analyses represent the exception rather than the rule and none of them focuses specifically on the age at arrival. One exception is provided by Heath et al. (2012) who compare cross-country results based on PISA data and confirm the existence of a late-arrival penalty for immigrant students. However, results are heterogeneous across countries, with Italy, Belgium, Sweden and Ireland being the countries with the largest late-arrival penalties. Second, using TIMSS data, Sweetman (2010) finds that length of stay beyond the first year has no impact on standardized test scores results in a comparison of immigrant children in Australia, Canada and the United States.[11]

Finally, to our knowledge, no existing empirical research addresses the question of the age at arrival of immigrants and their educational outcomes for the Italian case. We will therefore investigate this issue in the following sections.

3. DATA AND DESCRIPTIVES

Our source of data is the database provided by the National Institute for the Evaluation of the Educational System of Instruction and Training (INVALSI henceforth), a government agency that

[10] In fact, the datasets employed in this study are the 2006 Progress in International Reading Literacy Study (PIRLS) and the 2007 Trends in International Mathematics and Science Study (TIMSS), and age at immigration information is only available in TIMSS and, therefore, this analysis is only possible for the science and math scores and not for the reading scores.

[11] See also Sweetman et al. (2014) for a recent survey.

carries out a yearly evaluation of students' achievement in both Mathematics and Language. Since the focus of the paper is on language skills of immigrant students we use the results on the overall language test (or Italian test), that covers the different domains of reading comprehension, knowledge of the language and grammar. In Section 5 we will also discuss some evidence using the results on the math test as a robustness check. Further, in order to better compare the results for students attending different years and grades, we use the normalized values of the language standardized test. The latter are the test scores for language expressed as percentage of right answers, and values range from 0 to 100. In general, INVALSI tests are similar to the PISA standardized tests since their aim is to measure how far students have acquired the knowledge and skills essential for full participation in the knowledge society. Unlike PISA tests, the INVALSI standardized tests are compulsory for all Italian schools, both public and private, and all students attending specific school grades.

In our analysis we use the 2010-11 school-year data for four stages of schooling: second and fifth grade of primary school, sixth grade of lower secondary school and tenth grade upper secondary school. We therefore use four different samples, each consisting of approximately 400 to 500 thousand individuals/students per year (see Table 1). The Italian school system starts at age six, with five years of primary school (grades 1 to 5) followed by three years of lower secondary school (grades 6 to 8). Upper secondary education starts at year 9 and lasts three to five years depending on the type of school chosen. INVALSI tests were introduced in the 2008-09 school year, but 10th grade students were administered these evaluation tests for the first time only in 2010-11. For this reason, 10th year data need to be interpreted with some caution since, for this specific year, the language test had been intentionally designed by INVALSI to be easier than normal. Finally, note that the tests made in primary schooling are easier than those of both lower and upper secondary school students and this is reflected in a higher level of correct answer for second and fifth graders.[12]

INVALSI also collects detailed information about a significant number of student's background and family characteristics. In particular, this data are collected through a separate 'Family Questionnaire' sent to each family before the test, a 'Student Questionnaire' filled by each student the first day of the test, and a general information part on the students that is compiled by school administrative staff. However, one important exception is represented by primary school children attending year two, for whom data on personal characteristics are not collected. Therefore, for this cohort we do not perform any regression analysis and only report some descriptive statistics on the proportion of immigrant students and their school outcomes depending on their age of arrival in

[12] See INVALSI (2011).

Italy and their place of birth since they provide some interesting hints.

Datasets enable us to distinguish between Italian and non-Italian students. It is important to note that this classification refers to a pure citizenship criterion and that, unlike other countries, in Italy this follows the *Ius sanguinis* rule according to which individuals' identity (and their citizenship) is determined by family and not territory.[13] Thus, following Tonello (2011), from now on we will use the terms native or Italian as synonyms, implying that a student born abroad is native/Italian if at least one of the parents is an Italian citizen. Conversely, for immigrant students we use a standard approach and separate first generation students, that is, students born abroad of foreign-born parents, from second generation students, that is, Italian-born children of foreign-born parents.

Table 1 introduces the main data on students enrolled in the Italian schools for all grades examined distinguishing natives from immigrant students. The overall percentage of immigrant students is broadly similar across the different grades and ranges from 9.7% in year 6 to 7.3% in the 10th year. As also stressed by Contini (2013), the observed lower share of students in upper secondary school may be the result of the immigrant students relative disadvantage: drop-out and non-continuation rates among immigrants are much higher than among natives, and a much larger percentage of children entering upper secondary education opt for leaving the school system especially during the first two years of upper secondary schooling. Furthermore, the distribution of first and second generation immigrant students reveals some interesting variation across the different grades. In particular, the percentage of second generation pupils is higher among younger children (2nd grade), while older students in grade 10 have a larger proportion of first generation immigrants (5.2% versus 1.6 of 2nd generation). This is a well-known phenomenon in the analysis of migration patterns. In countries with established migration histories, there is a larger proportion of second-generation students than first-generation students. Conversely, in countries like Italy, where immigration is a recent phenomenon, first-generation students are the majority. The 2011 Census data show that the Italian immigrant population is extremely young: the mean age average of the total immigrant population is about 31.[14] Thus, it is likely that the change in pattern between the two components of immigrant students is firstly observed in the earlier years of schooling.[15]

Table 2 presents the distribution of immigrant students in the traditional three Italian geographical

[13] The Italian citizenship rules follow the Roman law rule of the *Ius sanguinis* that states that citizenship is defined by the family of birth and not the country of birth. In other country studies where the *Ius soli* is applied, that it, is the right of anyone born in the territory of a state to nationality or citizenship we find that the native vs. non-native categorization follows different criteria with respect to the ones described here.

[14] Istat (2012).

[15] See OECD (2012).

areas: North, Centre and South. In Italy there exists a deep, persistent duality between the developed North-Centre and the less developed South and this may affect migration flows. Table 2 confirms the expected patterns. First of all, there are large variations in the geographic distribution of immigrant students, with the richer northern areas that receive, as expected, the vast majority of migration flows.[16] In fact, around 60% of all immigrant students (both first and second generation) are located in this area of the country, while in the South these percentages range from only 9% (second generation immigrants among 2nd graders) to a maximum of 21% for 10th graders.

In Table 3 we identify the distribution of first generation immigrant students considering, for each grade, their place of birth. We also include information on second generation that, needless to say, are born in Italy. Unfortunately, for the former INVALSI has collected information only for few large geographical macro-areas and we are only able to identify immigrant students who are born 1) in a EU27 country, 2) in a European country outside EU27, or 3) outside Europe. Thus, despite the vast literature that stresses that differences in educational attainments vary significantly across ethnic communities, data disaggregated by country of birth are unfortunately not available. However, from recent data on Italian migration flows by country of origin it is possible to infer that the first group, EU27, mainly consist of children born in one the New EU member states. In fact, the number of EU27 citizens migrating to a Member State other than their own has significantly increased during the last years and peaked in 2007, but the largest group is formed by the newer EU countries: Romanians, followed by Poles and Bulgarians.[17] At the same time from migration flows data we may infer that the second group includes first generation students born in one of the East-European countries that are involved into the European Neighbourhood Policy (ENP). The ENP is a specific policy that supports political and economic reforms in sixteen of Europe's neighbouring countries from the East and Southern borders of EU. In particular, this policy was conceived after the 2004 enlargement of the EU with 10 new member countries, in order to "avoid creating new borders in Europe" and promote good governance and social development in Europe's neighbour. Therefore, it should also work as an effective integration tool for the management of migration flows for the European Union since EU countries are the main destinations of migrants from the sixteen ENP countries. Finally, the third category, students from outside Europe, is the most heterogeneous and does not even enable us to distinguish immigrant students by continent of birth. Considering first generation students sample, we observe that, with

[16] This is well documented also at the country level: immigrants sort across countries and the more developed countries usually host a higher share. See Brunello and Rocco (2012). Note also that the geographic distribution of the total number of Italian students (both natives and immigrants) is similar, around 40% between North and South, with less than 20% in the Centre.

[17] The EU-27 Member States received 384 000 Romanian citizens, 266 000 Polish citizens and 91 000 Bulgarian citizens. See European Union (2011).

the exception of second graders, the largest group of immigrant students (more than 40%) comes from non-European countries. Among younger children (2nd grade) EU27 immigrants are more numerous and, due to the enlargement of the European Union, they are expected to further increase in the near future.

Table 4 provides statistical evidence of the differences in the Language test score results between native and immigrant students. First, as observed in most countries, we find that native students obtain on average a significantly higher score than migrants in all years, and that 2nd generation immigrant students perform better than first generation. Second, we identify the place of birth of first generation students in order to investigate if even at this macro-area level, it affects students' test results. For students on all grades, data suggest that the test performance of European, both EU27 and other European countries students, is consistently better than that achieved by students from other countries. In particular, for primary school students (both 2nd and 5th graders), the average test scores of EU27 first generation students is even higher than that achieved on average by the second generation students. In general, these data suggest that, as also found in other countries analysis, in Italy differences in educational attainments vary significantly across the immigrant students' place of origin. This suggests that the quality of the country of origin schooling system and, more broadly, institutional and cultural factors play a role.

In Table 5 we classify first generation immigrant students based on how long they have been living in Italy before starting school. INVALSI data provide information on specific age ranges and we are able to distinguish between pupils who have lived in Italy before starting school for i) only one year, ii) between 2-4 years, iii) between 5 and 7 years, and iv) for more than 7 years. Table 5 includes their language test results correspondingly. [18] Moreover, in order to ease the analysis, we show the test score results for natives and second generation once more. As expected, we find that the length of stay of first generation immigrants explain a significant part of their observed achievement gap with the second generation group. In general, for all schooling grades we observe a similar pattern: as expected, the longer their stay in Italy, the lower the achievement gap.

Moreover, interesting results arise when we compare the data on the specific levels of schooling. First, for primary school children (2nd and 5th graders) we find that after having spent 5 or more years in Italy, the percentage of correct answers of a first generation student is almost identical to

[18] More precisely, the Questionnaire asks these students how old they were when they arrived in the country of assessment but classifies the data differently depending on the students' grades. For example, while for 2nd grade pupils INVALSI we know the exact age at which first generation immigrants arrived in Italy, for upper secondary (10th grade) students we only have information on specific age ranges: up to 3 years, 4 to 6 years, 7 to 9 years, 10 to 12 years, 13 to 15 years and, finally, 16 years or older.

that obtained by second generation. Note that these are almost certainly students who have never attended any other school system but the Italian one. Conversely, for children attending secondary schools (sixth and tenth graders) differences between first and second generation persist even after more than 5 years spent in the host country. Note that these students are likely to have previously attended a different schooling system before the Italian one. In particular, for sixth graders the achievement gap between first and second generation immigrants is 6% (54.7 vs 51.1), while for tenth graders it increases to 9%. This evidence therefore confirms what has been found in other studies. First, we find that in terms of educational achievements, first-generation immigrant students who arrived in Italy at a young age are almost indistinguishable from second generation immigrants (see Van Ours and Veenman, 2006 among others). In particular, if immigrant students have only attended the Italian school system, they are very likely to get the same test results irrespective to their first or second-generation status.

Further, Table 5 also shows that the sub-group of New EU member states first generation immigrant students, or EU27, performs better than the two remaining groups: the former group average test score results are almost identical to the second generation results. Conversely, the largest gap is observed for non-European immigrants. Overall, these differences across area of origin indicate that other factors, such as language, institutional structures and cultural features of the country of origin, are likely to play a role in the observed educational disadvantage.

In sum, comparing the test results across different grades we firstly find that the gap between natives immigrant students is significant and persistent in all years of the Italian school system. This is not surprising since the gap in the educational attainment of Italians versus immigrants has been already found as one of the highest across OECD countries in other studies. As shown by Dustmann et al. (2011), this evidence cannot be simply explained by the fact that Italy tends to attract immigrants with low qualifications. In this case, the gap would be mainly due to the socio-economic background of immigrant families. The formal skills gap across the two groups of natives and immigrants is in fact similar: that is, the observed skills of the Italian-native labour force are also low compared to other OECD countries. Moreover, these descriptives also indicate that first generation tend to catch up with second generation immigrant students. It seems to be a matter of time, but only if immigrant students arrived in Italy when very young and they have almost exclusively attended the Italian school system. Thus, being born abroad does not seem to cause a permanent disadvantage for first generation immigrant children with respect to second generation students. We will further investigate these issues in the following section.

4. MAIN RESULTS

We estimate a standard education production function where student test performance in language is modelled as a function of the native vs. immigrant first and second generation status, and a set of additional variables that control for student characteristics (gender, socio-economic background, native/I or II generation immigrants and area of origin), school characteristics (size, school type if in upper secondary school) and area characteristics (macro-area dummies). Table 6a sums up for each level of schooling the major characteristics (mean and standard deviation) of these additional variables for our overall sample. We also replicate the Table for the specific sample of immigrant students (Table 6b).

More precisely, we examine the relationship between the immigrant status and students' outcomes using two alternative regression settings, one of which takes into account the length of stay in the Italian school system. In details:

$$Y_{ij} = \alpha + \beta first_{ij} + \gamma second_{ij} + X_j'\delta + Z_j{}'\theta + \varepsilon_{ij} \qquad (1)$$

$$Y_{ij} = \alpha + years_{ij}{}'\beta + \gamma second_{ij} + X_j'\delta + Z_j{}'\theta + \varepsilon_{ij} \qquad (2)$$

In both specifications, Y_{ij} is the result obtained at the language national standardized test of student i attending school j; *first* and *second* are two dummy variables indicating, respectively, first and second generation immigrants; *X* is a set of individual and family additional characteristics and *Z* are school and area controls. Unlike eq. 1, in model 2 the simple dummy *first* is substituted by $years_{ij}$, a set of dummy variables indicating the length of stay in Italy of first generation immigrants. These dummies separately identify if these students have spent a) one year, b) between 2 to 4 years, c) more than five years in the host country. An exception is found for 10th year students, for whom we have identified four rather than three dummies/categories for *years*: in this case, we distinguish the last category between c) 5 to 7 years in Italy and d) over 7 years.[19]

This analysis is performed using student data for three different stages of schooling: the fifth year (last year of primary school, ISCED 1), the sixth grade in lower secondary (ISCED 2) and the tenth grade of upper secondary school (ISCED 3). Comparing the results of several stages of schooling enables us, even if imperfectly, to disentangle the effect on language performance of students' age at arrival from the effect of how long immigrant children have been in Italy. In particular, we try to identify if pupils are particularly at risk of suffering a long-lasting gap if they joined the Italian school system at different ages. Note that our year's dummies capture both a)

[19] Immigrant students who reside in Italy for less than one year do not have to attend the test and are therefore excluded from the analysis.

years in Italy before performing the test and b) age at immigration. In fact, these are two sides of the same coin: for example, a first generation immigrant student who has spent one year in Italy and is attending the fifth year in primary schooling, has arrived in Italy around the age of nine.[20]

4.1 FIRST AND SECOND GENERATION IMMIGRANT STUDENTS

We start the analysis in Table 7 where we include the OLS results when estimating equation 1. In our results we always report in parenthesis robust standard errors, clustered at school level. Models from 1 to 4 show evidence for fifth year students and use a sample including approximately four hundred thousand students attending the 2010-11 school year.

Model 1 introduces the results for our most parsimonious specification. Together with immigrants dummies, it includes a series of additional controls for the students main characteristics, that is, gender, a measure of her/his socio-economic background, if she/he speaks a foreign language at home or an Italian dialect, and the number of students per class. First of all, estimates confirm that in Italy children of immigrants face important gaps in language test results, with first-generation immigrants being the most disadvantaged group. We find that the coefficient on first generation immigrant is -5.2, while for second generation is -3.6. An F-test on the equality of these two coefficients largely rejects the null. Note that our dependent variable, the test scores results for language, is expressed as percentage of right answers. Given that the test results are between 0 and 100 the estimated coefficients can be interpreted in terms of decreased test score results: that is, *ceteris paribus*, the percentage of correct answers for first and second generation immigrant student is on average 5.2% and 3.6% below the natives'.

Second, the index of socioeconomic background, ESCS, is positive and strongly associated with student achievement. This variable, analogous to the same one computed by OECD for the PISA project, is created by INVALSI on the basis of the occupational and educational level of the student's parents and her/his home educational and cultural resources. The individual scores of this index are obtained by a principal component analysis, with normalized zero mean and unit standard deviation.[21] The indicator of socioeconomic background is one of the most important controls in this analysis since it takes into account the influence of family on achievement and choices: in addition to their cognitive abilities, immigrant (and natives) students sorting is certainly significantly driven by the family background (Brunello et al., 2007). Further, we also

[20] This is not true for students that are repeating a year or for students that attend a year lower or higher than the one that correspond to her/his age.

[21] See also INVALSI (2011) and Campodifiori et al. (2012) for more details on the ESCS index.

control for gender. In general, cross country analysis shows that language gender gaps often results in favour of girls and this is also confirmed by our analysis.[22]

Third, since our focus is on language skills, we also include a dummy, foreign language, identifying if the surveyed student speaks a different language from Italian at home. Using this dummy we try to disentangle the specific role of the family language and, possibly, of cultural attachment to one's nation of origin, from other immigrant students characteristics. In general, it has been shown that if a foreign language is spoken at home, second-generation immigrant children are more likely to find difficulties during their school career and to drop out early. Moreover, even if many immigrant students speaks their language of origin at home, our data stresses that a significant proportion report that the home language is the same as their destination country. We also observe a small percentage of Italian students, 1,7%, speaking a foreign language. In particular, the share of foreign language speakers at home is largest among 5th graders while decreases for older students. For the former, the percentage of first and second generation students who declare to speak Italian at home is, respectively, 30% and 50%. The coefficients on this dummy are always negative and significant and their values confirm that foreign students from diverse language background could encounter difficulties in the host country schools.

Model 1 also includes a dummy called dialect, equal to one if Italian students speak a dialect at home. Unlike other countries, Italy shows a significant percentage of dialect speakers, between 15 to 16 % across all school grades. Few studies focus on native students speaking a dialect at home and results on this dummy are therefore not obvious. In fact, differently from immigrants, dialect speakers have attended exclusively the Italian school system and are unlikely to meet significant social integration costs. Accordingly, we may expect that dialect speaker students are able to overcome possible difficulties. However, our results show that this is not the case. Controlling for different family's socio-economic background variables and additional student and school characteristics we find that speaking a dialect at home is significantly and negatively related to students' standardized test results.

Finally, note that in all specifications, including model 1, we also introduce macro-area dummies.[23] In fact, previous studies show that geographical location is an important determinant of Italian students test scores, with students in the North-East area usually outperforming the

[22] For example, PISA 2009 results report higher mean reading performance for girls in most countries.
[23] We identify the following dummy variables: North-East, North-West, Centre, South, South-Islands.

others, and those from the South being substantially behind. Our results confirm the expected territorial patterns.[24] Overall, model 1 results are largely consistent with the literature. One possible exception is the number of students per class: its coefficient is never significant for younger children.

In model 2 we further focus on two specific features of the students' socioeconomic background. In particular, we separately identify the role of specific family cultural upbringing, possibly not fully captured by the ESCS variable. The latter has been created using the (first) principal component analysis, a useful approach for creating a new variable that are linear combinations of a set of highly correlated original variables but that also has several shortcomings with high dimensional data or large numbers of data points as in our case. In particular, reducing a large number of variables into a smaller set of linear combinations (components), as done for the ESCS index, may hide role of specific important components. The INVALSI dataset enables us to separately identify two important elements of the immigrant students' socioeconomic status. The first is the number of siblings (no. siblings). This indicator is important since statistics show that immigrant women's' fertility rates are significantly higher than native women: respectively, 2.23 for immigrants and 1.31 for natives.[25] Thus, immigrant families' size is usually larger than that of natives and in terms of educational achievements, larger families imply worse studying conditions at home, such as the absence of a quiet place for studying or less time dedicated from parents to each children. Moreover, international studies also show that, unlike natives, the presence of dependent children among migrants significantly increases the risk of poverty and also of being in a less favourable situation with regard to housing conditions: the proportions of migrants living in overcrowded dwellings is high in almost all EU countries.[26] Thus, our dummy on immigrant student may also capture the effect of children living in overcrowded dwellings.

For the second, we follow Hanushek and Woessmann (2011) who argue that the number of books at home is the best single predictor of students' performance and include a dummy equal to 1 for students with more than 100 books at home. The variable manybooks is included among the determinants of the socioeconomic index but its correlation with ESCS is not high and

[24] On this see Di Liberto et al. (2013), Cipollone et al. (2010) and Bratti et al. (2007). Results are available upon request.

[25] Considering the mothers' citizenship, in 2009 second generation children born in Italy were mainly from Romanian (16,727), followed by Moroccan (14,370), Albanian (9,937), and Chinese mothers (just over 5,000 births). See ISTAT (2011).

[26] The overcrowding indicator relates the number of rooms in the house to the number of people. See Eurostat (2011a).

multicollinearity does not represent an issue here.[27] Due to missing values on the new variables in this case we are able to estimate the models on fewer observations with respect to the previous parsimonious specification. Including no. siblings and *manybooks* in the model does not significantly change the results on the ESCS coefficient, while these additional variables are both significant and with the expected sign, negative for siblings, positive for books, and cause the coefficients of both first and second generation students to decrease. Overall, this evidence casts some doubts on the use of ESCS as a synthetic index of the individual socio-economic background. It also suggests that specific policies directed towards improving studying conditions in schools, such as providing adequate study spaces and efficient libraries, are likely to have large returns for the more disadvantaged students.

Finally, in Model 3 we include two important school characteristics, the school size and the average index of socioeconomic background at school level, or *escs_school*, while in model 4 we further introduce two dummies that identify the countries (area) of origin of first generation students. Model 3 additional variables are never significant and do not change the results seen in the previous models. Conversely, the two included area of origin dummies are negative and significant. Our reference category is represented by students born in one of the EU27 countries, while European non-EU27 and non Europeans are the included dummies. As said in section three, we may roughly identify the first group, EU27, as mainly consisting of children born in one the New EU member states, while the second group should mainly include children born in one of the East-European countries currently part of the European Neighbourhood Policy (ENP). We find that, with respect to the new EU27 member states, the schooling performance of both ENP-East European Countries and the remaining group of non European countries is always worst. In particular, the latter group shows the largest gap in language test outcomes.

These models are replicated for the sample of 6th year students in models from 5 to 8. In terms of students' characteristics, our analysis confirms for the most part the results on primary school children, but these results also show a widening achievement gap between first and second generation immigrant: in model 5, the parsimonious specification, the percentage of correct answers for first and second generation immigrant student, is on average 8.3% and 3.7% below the natives'. Moreover, another important exception is observed for our three school characteristics indicators, class size, school size and *escs_school*. The latter indicator, *escs_school*, should take into account the socio-economic composition of the school and peer effects. Table 7 therefore

[27] Correlation coefficients are 0.41 for 5th graders, 0.48 for 6th graders and 0.52 for upper secondary students. We've also checked the correlation between ESCS and no. Siblings and, for all levels of schooling, is always negative and very low (around -0.10).

shows that, unlike primary school years, for lower secondary students these coefficients are always positive and significant. In particular, this result could be driven by the presence of a sorting process, possibly related to the family background, of best students into best schools, a process absent or confined at primary school level. This evidence is consistent with other findings from recent studies that stress how in Italy the inequality of opportunities for students, while almost absent at primary school level, would arise in the early years of secondary school (years 6-8). [28] In other words, our results confirm other recent evidence suggesting that, starting from the lower secondary school level, the family background in Italy is likely to be significantly related with students' educational achievements. That is, unlike other industrialized countries, Italian schools seem unable to stop possible adverse mechanisms of intergenerational inheritance, with high intergenerational educational persistence that ultimately translates into very low intergenerational income mobility, a sign that the system suffers a significant problem of equity.

Furthermore, our results on tenth grade students (Models from 9 to 12) indicate that this initial sorting would also translate into a social tracking along the upper secondary's tracks: the positive and significant coefficient on *escs_school* is also found for upper secondary students. Models from 9 to 12 include also two additional dummies, Lyceum and Vocational. In fact, unlike their younger peers, Italian students face, at the start of upper secondary school in year 9, the choice between different possible curricula and we therefore need to include these additional variables that identify the school type. Indeed, as said above, previous studies on the Italian case reveal that at this level of schooling the educational track plays a significant role for educational outcomes. Italian students choose schools that specialize in each of these three main curricula: Lyceum, Technical and Vocational. That is, unlike other countries, Italian upper secondary school tracking is not determined by a formal assignment process to academic or vocational courses depending on students' past performance or by any alternative selection processes. The vocational/academic intensity is at its lowest/highest level in the Lyceum (with almost no vocational component) and at its highest/lowest level in Vocational schools. In between these two curricula there is the curriculum offered by Technical schools. Moreover, only Vocational schools can last for 3 rather than 5 years, even if graduates from all three school types, after five years, may continue to tertiary education. In general, empirical studies show that students in general/academic track in most cases attain higher achievement than those in vocational tracks and this is true also for the Italian case (Cipollone et al., 2010, Di Liberto et al., 2013).

[28] See Triventi et al. (2009), Mocetti (2012), Di Paolo (2012), and De Simone (2013).

In general, our results on 10th graders are very similar to those previously discussed for 6th grade students. With respect to school type, in our regression analysis we use the Technical school dummy as reference and find, consistently with other studies, a positive and significant coefficient for Lyceum and a negative one for Vocational schools. Again, this does imply that the choice of the type of school and curriculum is not random, while it is related to family background with immigrant students more likely to end up in lower-performing Vocational schools, because they often originate from lower social strata. However, results for 10th grade students also show that both 1) the gap between immigrant and natives and 2) that between first and second generation immigrant students are lower than that observed for 5th and 6th grades students. Reasons are twofold. First, as already said in Section 3, the language test for tenth graders was designed by INVALSI to be easier than normal and this may affect the natives-immigrants performance gap. Second, this result may be also explained by the presence of higher drop-out rates in upper secondary school than lower levels of schooling. In other words, it is possible that the selection mechanism already described in section 3, is at work with the most disadvantaged (mainly immigrant) students leaving the Italian school system at the end of the eighth grade level.

4.2 THE LENGTH OF STAY OF FIRST GENERATION IMMIGRANT STUDENTS

We now turn our analysis to the model specification of equation 2. Results are shown in Table 8 that replicates the previous analysis substituting the single dummy for first generation students with separate dummies that also take into account their length of stay in Italy. That is, three dummies separately identify for how long first generation students have been living in the host country. In details, we classify as *late arrival* children those who have spent one year in Italy, long staying or early arrivals as those who have spent more than five years, while an intermediate category is represented by first generation children who stayed in Italy between 2 to 4 years. Indeed, previous studies show that age at immigration significantly affect the educational attainments of first-generation immigrants, with *immigrants arriving in their teen ages reaching lower educational attainment levels than those who arrived earlier.* Note that first generation results are also likely to affect the educational attainments of (future) second generation immigrants through intergeneration transmission mechanisms.

As before, models from 1 to 4 display primary school children results. This evidence suggests that the late-arrival penalty for first generation immigrant students is significant but that, after a relatively short period in the Italian school system, this gap decreases rapidly: the estimated coefficient in model 1 drops from -11.3 for late arrival children to -3.8 and -3.4 for, respectively, immigrant children who are enrolled in the Italian school system between two to four years before the test and those that enrolled since the start of primary school (five or more years). In particular, we find that after four years in the Italian school system the attainment gap of first generation immigrant students is almost identical to that observed for second generation students (-3.41). Further, considering that the observed mean value of the language test of all five graders is 73.9, our results imply a 15% gap for a newcomer in the Italian school system that decreases to less than 5% for early arrivals children. Very similar results are found in models 2 and 3 specifications.

In addition, model 4 results confirm once more that the score disadvantage of first generation immigrant children significantly depend on their country-of-origin. Including the two area of origin dummies in our regression model we find that the value of the coefficients on the length of stay dummies decreases substantially: coefficients in model 4 implies a 8% gap for a newcomer from a EU27 country that decreases to less than 1% for earlier (five or more years) arrivals children. That is, the early arrivals point estimate of -0.13 implies that with respect to the average student, the difference of this group of immigrant children in test results is only -0.2%. In sum, considering the country of origin the coefficient of the first generation is even lower than that observed for second generation children for whom we cannot control the area of origin.

Models from 5 to 8 show the 6th graders sample results. The decreasing pattern of the length of stay dummies is similar to that observed for primary school children. However, for sixth graders the estimated gap of late-arrivals is significantly larger and the pace at which first generation immigrants close the gap during their stay in Italy is slower: with an observed mean value of the language test of 60.9, these coefficients imply a 25% gap for a newcomer, 10% for those enrolled between two to four years before the test and 9% for the early arrivals. Moreover, even after controlling for many important demographic and school characteristics (models 6 and 7) the achievement gap of long staying first generation is larger than second generation students. However, as before, when we include the area of origin dummies (model 8) we observe a significant drop in the length of stay coefficients.

For tenth grade students (models from 9 to 12) we identify four rather than three categories for the length of stay and split the early arrival children category between c) 5 to 7 years in Italy and d) over 7 years. Unlike the youngest cohort, the tenth graders sample delivers a longer age-at-immigration performance profile. Although in an indirect and imperfect way, for this reason the tenth grade students cohort offers the most robust sample to test for the presence of a critical age above which first generation immigrant students arriving in the destination country face strong negative impact on their school performance. Given that the mean value of the tenth graders language test is 67.5, model 9 coefficient values imply a 20% gap for a newcomer (15-16 years old students), 10% for those enrolled between two to four years before the test (about 14-12 years old), 4% for those enrolled between five to seven years (11-9 years old), and only 2% the early arrivals (8 years or younger). Second, once we control for the area of origin (model 12) we find that, after having spent more than four years in Italy, the first generation achievement gap disappears, while the seven_more dummy coefficient even shows a positive sign. These results would imply a turning point around the age ten: that is, children arriving in Italy up to about the end of the primary school are able to catch up in terms of language skills. Thus, our findings seem to corroborate those of previous studies that estimate a critical age at immigration of about 9.

Finally, comparing the findings across the three level of schooling we observe that the late-arrival penalty is different for the three levels of schooling analysed: it is at its lowest for primary school children, it peaks for sixth graders, and then decreases again. That is, the estimated school attainment gap decreases when we move from lower to upper secondary school results, but it is possible to explain this puzzling result with the specificities of upper secondary school (easier test and drop outs) described above. Second, while the achievement gap always decreases for all cohorts with time, the speed at which decreases depends also on the children age at arrival. In fact, comparing the coefficients of immigrant students who arrived in Italy one year and those who arrived 2 to 4 years before the test we observe a 66% decrease for fifth graders, 56% for sixth graders while the gap for tenth graders is only reduced by 48%. That is, the estimated pace at which the gap closes is slower the later the children arrive. Once more, this evidence corroborates the critical period hypothesis (Blakeley and Chin, 2004 and 2010) that assumes that children are able to learn new languages in an easy way when they are younger.

5. ROBUSTNESS CHECKS

In this section we perform a set of robustness checks of the findings discussed above. First concerns are selection issues. Even if our set of covariates allows us to control for many likely sources of endogeneity, when the focus is on educational outcomes of immigrant students, selection issues are likely to play a role in our previous OLS analysis and the interpretation of our results in causal terms always should be taken with caution.

In order to take into account possible unobserved factors that might affect the native-immigrants test score gap and are common within each school or class, we performed the same analysis introducing both school and classroom fixed effect. To save on space, in Table 9 and 10 we only report the results we obtain when we replicate the previous analysis with class fixed effects. This specification has the advantage that both observed and unobserved class (and school) variables are removed, overcoming many issues of self-selection. In particular, school and class fixed effects enable to control for problems related to the non-random allocation of immigrant students across schools and classes and omitted variable problems.[29] Interestingly, inspection of Table 9 reveals very small differences for our full set of students' controls with respect to results reported in Table 7 and the same results arise when we compare the point estimates obtained with class fixed effects in Table 10 with that of Table 8.[30] Even our dialect dummy remains negative and significant and this result was not obvious. In fact, it is likely that most dialect speaker families also live in rural areas where they can take advantage of worse educational infrastructures than those living in urban areas. In this case, our OLS results would also reflect the effect of living in rural locations while FE estimates confirm that this was not the case. Overall, we take these results as fully corroborating the previous ones, even if the same fixed effects estimation is not free of other sources of selection bias.

As a second robustness check we substitute our measure of students' socioeconomic status, ESCS, with the family educational attainment level and check if there are substantial changes in our estimates of the immigrant-native differentials.[31] Parental education is calculated as dummy variables, reflecting 4 different levels of education: tertiary, post-compulsory secondary, compulsory and less than compulsory, using the father or mother maximum educational attainment level, whichever higher. Indeed, numerous studies consider parental education as the most important predictor of school performance and educational attainments. Moreover, parental education is found to be the main mechanism through which parents may achieve a better socioeconomic status.[32] Tables 9 and 10 shows the results when using these variables instead of ESCS in our less parsimonious specification (see models 4, 8 and 12 for fifth, sixth and tenth graders respectively). The less than compulsory schooling attainment levels is the reference category and we find, as expected, that all coefficients are positive and significant but one single exception of a non significant coefficient for the degree level for tenth graders. In general, results suggest that substituting ESCS with parental education does not significantly affect our main results.

In the following, we discuss our additional robustness evidence without including the corresponding Tables in the Appendix (available upon request). We have also replicated our

[29] A recent paper for the Italian case, Ballatore et al. (2013), assumes that immigrant students are allocated more in disadvantaged schools and classes in which natives have a less favourable socio-economic background.

[30] When comparing OLS and FE results, note that, with respect to Tables 7 and 8, the exclusion of all class and school variables in Tables 9 and 10 cause the exclusion of one estimated model for each grade.

[31] Note that ESCS is a composite index created also with the educational level of the parents. Unlike *manybooks*, the two variables, ESCS and parental education, are highly positively correlated (0.81 for fifth graders, very similar for sixth and tenth graders).

[32] See Di Paolo (2012) among the many others.

analysis selecting a specific sub-sample of schools. Indeed, in the above analysis we have used the entire student population data but, in addition to this, INVALSI also conduct a specific nationally-representative survey, where the same tests are administered under the supervision of observers in each class of the sample. This survey is conducted in order to prevent and control for cheating, mainly observed in the southern areas of the country, and facilitate the procedures of data collection available on students' achievements. Despite the possible advantages - first of all, better quality data - the reduction of the sample size is significant: for example, for year 5 students, the sample reduces from almost four hundred thousand to only thirty thousand and correspondingly, our immigrant students' observations drops from more than 45 thousand to 2856. Nevertheless, the use of this high quality data sub-sample does not change our analysis.

As a final robustness check we have run the same regression analysis using an alternative performance variable, the test score results in mathematics. Previous studies on the impacts on multiple subjects usually find differential impacts of immigrant students across them, with typically worst performance in language rather than math or science test.[33] We confirm previous studies findings: all the results concerning mathematics test scores are qualitatively similar to that found for the Language test but smaller in size. We also confirm that the pace at which the gap closes is slower the later the children arrive: when we compare the coefficients of immigrant students who arrived in Italy one year before the test and those who arrived 2 to 4 years before the math test we observe a very similar and large decrease for both fifth and sixth graders (58% and 57% respectively), while the gap for tenth graders is only reduced by 43%.

5. CONCLUSIONS

Using a standard education production function setting, this paper investigates whether the length of stay plays a role in the host country language skills acquisition of immigrant students in Italy. We use the students' outcomes in a language standardized test for different cohorts of Italian students collected during the school year 2010-11. In particular, our regression analysis focuses on three different stages of students' school life: end of primary school (fifth grade), first year of lower secondary (sixth grade) and second year of upper secondary (tenth grade). Our results are very much consistent with the literature and are also robust to the inclusion of fixed effect at school and classroom level, in order to control for likely endogeneity problems, the use of a specific subsample of students that enables to control for problems arising from cheating, the use of a different control set and the use of math test scores as dependent variable.

[33] Recent references are Ohinata et al. (2012) Ballatore et al. (2013). This result is plausibly explained by the assumption that immigrant learning difficulties are more sizeable during Italian language lectures than during mathematics lectures.

First of all, we find a significant gap between native and immigrant students in school outcomes for all grades. In particular, the second generation achievement gap is large and corroborates the lack of adequate integration policies in Italy already stressed in other studies.[34] However, we also find that the acquisition of language skills represents a problem also for native students speaking a dialect at home and that, starting from the lower secondary school level, the socio-economic composition of the school and peer effects represent important determinants of the students' performance. In other words, the Italian school system seems unable to integrate disadvantaged students: due to their high social integration and language acquisition costs, immigrant students are those more at risk of poor results and social exclusion.

Second, for all levels of schooling and specifications we observe the same expected pattern: newly arrived immigrant children show the poorest performance in terms of test score outcomes, a result that can be is easily explained by the lack of familiarity with the new language and more precarious living conditions with respect to early-arrivals. Furthermore, the late-arrival penalty is different for the three levels of schooling analysed: it is at its lowest for primary school children, while it peaks for sixth graders. Also, this gap decreases with time and it depends on how long they have been in Italy and results suggest that the pace at which the gap closes is slower the later the children arrive. In other words, comparing the findings across the different school grades, we see that interventions at younger ages are likely to be more effective. In particular, our results corroborate those of other recent studies that estimate a critical age at immigration of about 9.

Third, this empirical analysis indicate that the area of origin and, thus, institutional and cultural factors play a role on immigrant students' outcomes and integration. In particular, we are able to identify, even if very imperfectly, and compare the schooling performance of two different groups of children born in relatively close geographical areas, the groups of New EU member states and European country non EU27. The database enables us to further identify a third highly heterogeneous group that includes all the rest of first generation migrants born outside Europe. This latter group always shows the largest gap in language test outcomes. In addition, when we compare the two European groups, we find that the students' performance of those born in the most geographically remote group, the non-EU27 group, is below that observed for EU27 children. Thus, it seems that the Italian school system has been able to integrate more easily the increased intra-European migration due to the EU enlargement into Central and Eastern Europe in 2004 rather than that the flow of students from other non-EU27 countries. There are many possible explanations for the different outcomes observed in these two culturally apparently

[34] Contini (2013).

similar groups. First, non-EU27 institutions are very unlike the EU27 ones: many non-EU27 countries are non-democratic while the process of democratization in New EU27 member countries has started long ago. However, these different outcomes may also be the result of a different migrant selection process due to the different immigration policies and rules between the two groups. However, our dataset and, in particular, its very broad level of aggregation, does not allow us to further investigate these important issues here. Indeed, it is fair to say that these results are more suggestive rather than conclusive.

In sum, since the acquisition of the native language by immigrants is of utmost importance for their integration, this analysis implies that new and effective integration policies need to be urgently implemented in Italian schools. It also suggests that, if the foreign children's late arrival is the result of national migration policies on family reunification, the possible benefit of delaying immigrant family reunification could be offset by the possibly large costs of students' lower school performance. Finally, our evidence seems to indicate that possible future implementation of new policies directed to integrate foreign students into the Italian schooling system should take into account cultural differences of its immigrant students and possibly avoid "one size fits all" approaches. Indeed, the success or failure of immigrants and their children to integrate in the destination country can potentially intensify conflicts within societies and, through this, affect economic growth.

In general, more should be done in order to assess which are the specific channels through which this feature works. Our evidence offers some clue on what can be done to foster immigrant students' school outcomes, suggesting that improving studying conditions in schools may be an effective policy for improving the acquisition of their language skills. However, much more research should focus in the future on what kind of specific interventions works.

Acknowledgments (if applicable):
The research leading to these results has received funding from the European Union's Seventh Framework Programme FP7-SSH-2010-2.2-1 (2011-2014), under grant agreement n° 266834 SEARCH project. I wish to thank INVALSI and, in particular, Patrizia Falzetti for providing the data on student outcomes. I also thank Benedetta De Magistris for her excellent work as research assistant.

REFERENCES

Alesina, A., La Ferrara, E (2002), Participation in Heterogeneous Communities, *The Quarterly Journal of Economics*, 115: 847-904.

Ballatore, R., Fort, M. & Ichino, A. (2013). "The Tower of Babel in the classroom: immigrants and natives in Italian schools", unpublished working paper (available at http://www.andreaichino.it/research_progress/ichino_ballatore_fort_babel.pdf).

Bleakley, H. & Chin, A. (2004). "Language Skills and Earnings: Evidence from Childhood Immigrants," *The Review of Economics and Statistics*, vol. 86(2): 481-496.

Bleakley, H. & Chin, A. (2010). "Age at Arrival, English Proficiency, and Social Assimilation among US Immigrants," *American Economic Journal: Applied Economics,* vol. 2(1): 165-92.

Böhlmark, A. (2008), Age at immigration and school performance: A siblings analysis using Swedish register data, *Labour Economics*, 15 (6): 1366-1387.

Böhlmark, A. (2009) Integration of childhood immigrants in the short and long run – Swedish evidence, *International Migration Review*, 43(2): 387-409.

Borjas, G. J. (1995). Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s? *Journal of Labor Economics*, 13, 201–245.

Bratsberg, B., Raaum, O., Roed, K. (2011). Educating children of immigrants: closing the gap in Norwegian schools. IZA Discussion Paper, 6138, Bonn.

Bratti, M., Checchi, D., Filippin, A. (2007) Territorial Differences in Italian Students' Mathematical Competencies: Evidence from PISA 2003, IZA Discussion Papers 2603, Institute for the Study of Labour (IZA).

Brunello, G., Checchi, D. (2007). "Does school tracking affect equality of opportunity? New international evidence." *Economic Policy*, 22:781–861.

Brunello, G., Rocco L. (2008), Educational Standards in Private and Public Schools. *The Economic Journal*, 118 (533):1866–1887.

Brunello, G., Rocco L. (2012), The effect of immigration on the school performance of natives: Cross country evidence using PISA test scores. *Economics of education Review*, 32: 234-242.

Cahan, D., Davis, D., Staub, R. (2001) Age at immigration and scholastic achievement in school-age children: Is there a vulnerable age? *International Migration Review*, 35 (2): 587-595.

Campodifiori E., Figura E., Papini M., Ricci R. (2010), Un indicatore di status socio-economicoculturale degli allievi della quinta primaria in Italia, INVALSI Working Paper n. 02/2010.

Cipollone, P., Montanaro, P., Sestito P. (2010), Value-Added Measures in Italian High Schools: Problems and Findings. *Giornale degli Economisti*,69 (2):81–114.

Chiswick, B. R., and DebBurman, N. (2004). Educational Attainment: Analysis by Immigrant Generation. *Economics of Education Review*, 23(4), 361–379.

Contini, C. (2013), Immigrant background peer effects in Italian schools, *Social Science Research* vol. 42 issue 4 July, 2013. p. 1122-42.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, F., Mood, A.M., Weinfeld, G.D., & York. R.L. (1966). *Equality of Educational Opportunity*. Washington, DC: U.S. Government Printing Office.

Corak, M., (2012) Age at immigration and the educational outcomes of children. In A. Masten, K. Liebkind, and D. J. Hernandez (Eds.), *Realizing the Potential of Immigrant Youth* (pp. 90–116). Cambridge: Cambridge University Press.

Cortes K. (2006), The effects of age at arrival and enclave schools on the academic performance of immigrant children, *Economics of Education Review*, 25: 121–132.

De Simone, G. (2013), Render unto Primary the Things which are Primary's. Inherited and Fresh Learning Divides in Italian Lower Secondary Education, *Economics of Education Review*, 23: 12-23.

Di Liberto A., Schivardi F., Sulis G., (2013), Managerial Practices and Students' Performance, *mimeo*.

Di Paolo, A. (2012). "Parental Background and Educational Opportunities: Evidences across Cohorts for Italy and Spain", *Revista de Economía Aplicada*, vol. 58-(XX-2012): 119-146.

Dustmann, C., Frattini T., Lanzara G. (2011) Educational Achievement of Second Generation Immigrants: An International Comparison. *Economic Policy*, vol. 27(69), 143-185.

European Union (2011), Migrants in Europe: A statistical portrait of the first and second generation, Luxembourg: Publications Office of the European Union.

Eurostat (2011a), *Migrants in Europe: A statistical portrait of the first and second generation*, Luxembourg: Publications Office of the European Union.

Eurostat (2011b), *Indicators of Immigrant Integration*, Luxembourg: Publications Office of the European Union.

Glaeser, E. (1999), The Formation of Social Capital, *mimeo*.

Gonzalez, A. (2003). The Education and Wages of Immigrant Children: The Impact of Age at Arrival. *Economics of Education Review*, 22, 203–212.

Hanushek, E.A., Woessmann L. (2011), The Economics of International Differences in Educational Achievement, in: Hanushek E.A., Machin S., Woessmann L. (eds.) *Handbook of the Economics of Education, Vol. 3*, Amsterdam: North Holland.

Heath, A. and E. Kilpi-Jakonen (2012), Immigrant Children's Age at Arrival and Assessment Results", OECD Education Working Papers, No. 75, OECD Publishing.

Helliwell J. and Putnam R., (1999), Education and Social Capital. NBER WP 7121. Cambridge: National Bureau of Economic Research.

INVALSI (2011), *Rapporto tecnico sulle caratteristiche delle prove INVALSI 2011*. Technical Report, INVALSI.

ISTAT, (2011), Natalità e fecondità della popolazione residente, *Statistiche Report*.

ISTAT, (2012), Il censimento della popolazione straniera, *15° Censimento della popolazione*.

Johnson, J. S., and E. L. Newport. (1989) Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language. *Cognitive Psychology*, 21 (1): 60-99.

Mocetti, S. (2012), Educational choices and the selection process: before and after compulsory schooling, *Education Economics*, vol. 20(2): 189-209.

Myers, D., Gao, X., Emeka, A. (2009) The gradient of immigrant age-at-arrival effects on socioeconomic outcomes in the U.S., *International Migration Review*,43 (1): 205-229.

OECD (2010a), *Closing the Gap for Immigrant Students: Policies, Practice and Performance, OECD Reviews of Migrant Education*, Paris: OECD Publishing.

OECD (2010b), *PISA 2009 Results: Overcoming Social Background – Equity in Learning Opportunities and Outcomes (Vol II),* Paris: OECD Publishing.

OECD (2012) Untapped Skills. Realizing the potential of immigrant students, Paris: OECD Publishing.

Ohinata, and Van Ours, J. C. (2012), Young immigrant children and their educational attainment. *Economics Letters*, 3, 288-290.

Patkowski, M. S. (1990), Age and Accent in a Second Language: A Reply to James Emil Flege. *Applied Linguistics*, 11 (1): 73–90.

Schaafsma, J., and Sweetman, A. (2001). Immigrant Earnings: Age at Immigration Matters. The *Canadian Journal of Economics*, 34(4), 1066–1099.

Schumann, J. H. (1976), Social distance as a factor in second language acquisition, *Language Learning*, vol. 26, 1, 135-143.

Sweetman, A., (2010) Immigrant children in elementary school: an international perspective, in: McDonaldd, T., Ruddick, E., Sweetman, A., Worswick, C. (eds.), *Canadian Immigration: Economic Evidence for a Dynamic Policy Environment*. Montral and Kingston: Queen's Policy Studies Series, McGill-Queen's University Press.

Sweetman, A., Van Ours J. C. (2014) Immigration: What About The Children And Grandchildren?, Tilburgh University Discussion Papers, No. 2014-009.

Tonello, M. (2011), Immigrant background peer effects in Italian schools, WP n. 14/2011, Università di Torino.

Triventi, M., Trivellato, P. (2009) Participation, performance and inequality in Italian higher education in the 20th century, *Higher Education*, vol. 58, pp. 285-306.

Van Ours, J. C., Veenman, J. (2006) Age at immigration and educational attainment of young immigrants. *Economics Letters*, 90 (3): 310–316

APPENDIX

Description of Variables

Source: Invalsi data for school year 2010-11.

Dependent Variables:

- **Language_test**: INVALSI Italian test, normalized scores.
- **Math_test**: INVALSI Mathematics test, normalized scores.

Test scores results are measured as the fraction of correct answers.

Additional controls:

- **gender**: Dummy=1 if female.
- **ESCS**: it follows the lines of PISA's homologous index. It is based on parental education, occupational status and a number of home possessions. The individual values are obtained by a principal component analysis. By construction, the average of the ESCS index is equal to 0 and its standard deviation is 1.
- ***_parents**: father or mother max educational attainment level (whichever higher) dummies reflecting 4 different levels of education: tertiary, post-compulsory secondary, compulsory and less than compulsory.
- **dialect**: Dummy=1 if language spoken at home is dialect.
- **foreign language**: Dummy=1 if language spoken at home is not Italian.
- **no. stud_class**: Number of students per class.
- **non europe**: Dummy=1 if born outside Europe.
- **other european**: Dummy=1 if born in Europe but no EU27.
- **no. siblings**: Number of siblings, 4 indicates 4 or more.
- **manybooks**: Dummy=1 if more than 100 books at home.
- **school_size**: Number of students per school.
- **escs_school**: Average School Level ESCS Index.
- **foreign 1st generation**: Dummy=1 if students born abroad of foreign-born parents.
- **foreign 2nd generation**: Dummy=1 if native-born children of foreign-born parents.
- **Lyceum**: Dummy=1 if the upper secondary school type is "Licei".
- **Technical**: Dummy=1 if the upper secondary school type is "Tecnici".
- **Vocational**: Dummy=1 if the upper secondary school type is "Professionali".
- **Campione**: Dummy=1 if school selected for monitoring by Invalsi.

The length of stay variables are constrained by the data on upper secondary school first generation immigrant students produced by Invalsi. This dataset enable us to identify first generation immigrant students by their age of arrivals in Italy according to the following age intervals: 16 years old (or older), between 13 and 15 years old, between 10 to 12 years old, between 7 to 9 years old, between 4 to 6 years old, 3 years old or younger.

Table 1. Distribution of natives and immigrant students by grade attended

	Native students	%	Total Immigrants	%	Total number
2nd grade (primary)	453591	91.3	43325	8.7	496916
5th grade (primary)	462483	91.1	45090	8.9	507573
6th grade (lower sec.)	467687	90.3	50038	9.7	517725
10th grade (upper sec.)	398421	92.7	31339	7.3	429760
	Immigrants 1st generation	%	Immigrants 2nd generation	%	
2nd grade (primary)	14168	2.9	29157	5.9	
5th grade (primary)	23895	4.7	21195	4.2	
6th grade (lower sec.)	30935	6.0	19103	3.7	
10th grade (upper sec.)	23017	5.4	8322	1.9	

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data.

Table 2. Distribution of immigrant students by macro-areas (percentage values)

Macroregions	Immigrants 1st generation	Immigrants 2nd generation
2nd grade primary school		
North	59.8	68.1
Centre	22.3	22.9
South	17.9	9.0
Italy	100.0	100.0
5th grade primary school		
North	63.0	66.0
Centre	23.0	22.9
South	14.0	11.1
Italy	100.0	100.0
6th grade lower secondary school		
North	64.4	66.4
Centre	22.8	21.6
South	12.7	12.0
Italy	100.0	100.0
10th grade upper secondary school		
North	64.6	58.8
Centre	22.8	20.3
South	12.6	20.9
Italy	100.0	100.0

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data.

Table 3. Distribution of immigrant students by place of birth.

				2nd generation		1st generation						
	Total No.	%		Italy	%	EU Countries (EU27)	%	Other European (non EU)	%	Non Europe	%	Tot. %
2nd grade primary school												
1st generation	14168	2.8				5728	40.3	3257	22.9	5183	36.5	100.0
2nd generation	29157	5.7		29157	100.0							100.0
Total immigrants	43325	8.4		29157	67.3	5728	13.2	3257	7.5	5183	12.0	100.0
5th grade primary school												
1st generation	23895	4.9				8319	34.4	5777	23.8	9799	41.7	100.0
2nd generation	21195	4.1		21195	100.0							100.0
Total immigrants	45090	8.9		21195	45.4	8777	18.8	6067	13.0	10640	22.8	100.0
6th grade lower secondary school												
1st generation	30935	6.4				9317	30.0	7911	25.2	13707	44.7	100.0
2nd generation	19103	3.7		19103	100.0							100.0
Total immigrants	50038	9.6		19103	36.8	9868	19.0	8277	15.9	14670	28.2	100.0
10th grade upper secondary school												
1st generation	23017	5.2				6033	26.2	7375	31.6	9609	42.2	100.0
2nd generation	8322	1.6		8322	100.0							100.0
Total immigrants	31339	8.0		8322	23.8	6962	19.9	8390	24.0	11211	32.1	100.0

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data.
.

Table 4. Average test scores: Language test results (by immigrant status and place of birth)

	Place of birth			
	Italy	EU Countries (EU27)	Other European (non EU)	Non Europe
2nd grade primary school				
Natives	73.2			
Immigrants (1st generation)		63.7	59.7	55.7
Immigrants (2nd generation)	61.6			
5th grade primary school				
Natives	74.7			
Immigrants (1st generation)		68.7	65.2	61.3
Immigrants (2nd generation)	67.8			
6th grade lower secondary school				
Natives	62.2			
Immigrants (1st generation)		52.0	49.6	44.1
Immigrants (2nd generation)	54.7			
10th grade upper secondary school				
Natives	68.3			
Immigrants (1st generation)		60.7	60.1	53.4
Immigrants (2nd generation)	63.8			

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data.

Table 5. Language average test scores: first- generation immigrants by place of birth and time spent in Italy before the test

		Time spent in Italy before the test			
		1 year	2/4 years	>5 years*	>7 years
2nd grade primary school					
EU Countries (EU27)		61.3	63.3	65.1	
Other European (non EU)		55.4	59.9	60.9	
Non Europe		51.2	55.2	58.0	
Total 1st generation		55.8	59.7	61.5	
2nd generation	*61.6*				
Natives	*73.2*				
5th grade primary school					
EU Countries (EU27)		63.4	70.0	69.9	
Other European (non EU)		58.8	65.6	67.2	
Non Europe		52.8	62.2	64.3	
Total 1st generation		57.6	66.0	66.9	
2nd generation	*67.8*				
Natives	*74.7*				
6th grade lower secondary school					
EU Countries (EU27)		43.5	53.9	54.6	
Other European (non EU)		42.4	50.6	52.0	
Non Europe		35.4	45.2	47.9	
Total 1st generation		39.0	49.4	51.1	
2nd generation	*54.7*				
Natives	*62.2*				
10th grade upper secondary school					
EU Countries (EU27)		47.3	57.2	62.0	62.9
Other European (non EU)		49.3	55.6	60.1	62.7
Non Europe		41.1	48.3	53.3	57.0
Total 1st generation		44.3	52.9	58.0	60.4
2nd generation	*63.8*				
Natives	*68.3*				

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data. *For 10th grade students the column on more than 5 years correspond to 5 to 7 years.

Table 6a. Descriptive statistics: whole sample

Variables	5th grade		6th grade		10th grade	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
gender	0.49	0.50	0.48	0.50	0.49	0.50
ESCS	0.00	1.04	-0.02	1.05	0.02	1.00
dialect	0.15	0.35	0.16	0.37	0.15	0.35
foreign language	0.07	0.26	0.07	0.26	0.05	0.21
no. stud_class	19.3	4.37	21.7	3.90	21.4	4.67
no. Siblings	1.24	0.88	1.25	0.92	1.24	0.87
manybooks	0.25	0.43	0.30	0.46	0.31	0.46
campione	0.06	0.24	0.07	0.26	0.09	0.29
foreign 2nd generation	0.04	0.19	0.03	0.18	0.02	0.13
foreign 1st generation	0.05	0.21	0.06	0.23	0.06	0.23
one_year	0.01	0.10	0.01	0.11	0.00	0.05
two_4years	0.01	0.12	0.01	0.11	0.01	0.08
five_more/five_7years	0.02	0.14	0.03	0.17	0.01	0.09
seven_more					0.02	0.14
other European	0.01	0.11	0.02	0.12	0.02	0.13
non Europe	0.02	0.14	0.03	0.16	0.02	0.15
EU27	0.02	0.13	0.02	0.13	0.01	0.12
school_size	102.6	45.9	147.0	77.7	179.0	77.5
escs_school	-0.01	0.47	-0.03	0.48	0.00	0.47
Lyceum					0.47	0.50
Technical					0.33	0.47
Vocational					0.20	0.40
North_East	0.18	0.38	0.18	0.38	0.18	0.38
North_West	0.25	0.43	0.25	0.43	0.24	0.43
Centre_North	0.18	0.39	0.18	0.38	0.18	0.38
Centre_South	0.23	0.42	0.23	0.42	0.24	0.43
Islands_South	0.16	0.37	0.17	0.37	0.17	0.37

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data.

Table 6b. Descriptive statistics: immigrant students

Variables	5th grade		6th grade		10th grade	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
gender	0.49	0.50	0.47	0.50	0.51	0.50
ESCS	-0.52	0.93	-0.54	0.93	-0.41	0.95
foreign language	0.63	0.48	0.62	0.49	0.59	0.49
no. stud_class	19.02	4.09	21.23	3.50	20.99	4.43
no. Siblings	1.53	1.08	1.55	1.11	1.54	1.12
manybooks	0.11	0.32	0.13	0.33	0.13	0.34
school_size	110	44	147	76	183	72
escs_school	-0.01	0.39	-0.02	0.40	-0.13	0.40
Lyceum					0.24	0.43
Technical					0.38	0.48
Vocational					0.38	0.49
foreign 1st generation						
one_year	0.10	0.30	0.12	0.33	0.03	0.18
two_4years	0.15	0.36	0.14	0.34	0.10	0.29
five_more/five_7years	0.23	0.42	0.29	0.46	0.12	0.32
seven_more					0.29	0.45
other European	0.13	0.34	0.16	0.37	0.24	0.43
non Europe	0.23	0.42	0.28	0.45	0.32	0.47
EU27	0.19	0.39	0.19	0.39	0.20	0.40
North_East	0.28	0.45	0.28	0.45	0.28	0.45
North_West	0.36	0.48	0.38	0.48	0.35	0.48
Centre_North	0.23	0.42	0.22	0.42	0.22	0.42
Centre_South	0.07	0.26	0.07	0.25	0.09	0.28
Islands_South	0.06	0.23	0.06	0.23	0.06	0.24

Source: Invalsi data for academic year 2010-11. See the Appendix for additional details regarding variables.

Table 7. OLS estimates: main results

Dependent variable: standardized language test results	5th grade - primary school				6th grade - lower secondary school				10th grade - upper secondary school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
gender	0.57***	0.55***	0.55***	0.57***	2.94***	2.91***	2.93***	2.92***	2.48***	2.43***	2.61***	2.61***
	(0.04)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.11)	(0.11)	(0.11)	(0.11)
escs	2.90***	2.45***	2.45***	2.46***	4.88***	4.07***	3.80***	3.81***	1.35***	0.75***	0.40***	0.38***
	(0.04)	(0.04)	(0.03)	(0.03)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
dialect	-1.55***	-1.40***	-1.39***	-1.36***	-3.44***	-3.16***	-3.06***	-3.04***	-1.12***	-1.12***	-0.87***	-0.88***
	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.12)	(0.12)	(0.12)	(0.12)
foreign language	-3.24***	-3.06***	-3.06***	-2.86***	-5.13***	-4.77***	-4.79***	-4.36***	-3.07***	-2.91***	-2.85***	-2.61***
	(0.14)	(0.14)	(0.14)	(0.15)	(0.14)	(0.14)	(0.14)	(0.15)	(0.18)	(0.18)	(0.18)	(0.18)
no. stud_class	0.01	0.00	-0.00	-0.01	0.17***	0.16***	0.13***	0.13***	0.16***	0.16***	0.09***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
other european				-2.91***				-1.87***				-0.47
				(0.32)				(0.32)				(0.29)
non europe				-5.23***				-4.82***				-4.55***
				(0.30)				(0.29)				(0.30)
n. siblings		-1.02***	-1.02***	-0.97***		-1.58***	-1.58***	-1.53***		-0.32***	-0.29***	-0.24***
		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)
manybooks		2.14***	2.14***	2.11***		3.06***	3.10***	3.00***		2.24***	2.08***	2.08***
		(0.06)	(0.06)	(0.06)		(0.08)	(0.08)	(0.08)		(0.07)	(0.07)	(0.07)
school_size			0.00	0.00			0.00**	0.00**			0.01***	0.01***
			(0.00)	(0.00)			(0.00)	(0.00)			(0.00)	(0.00)
escs_school			0.05	0.10			1.40***	1.50***			4.94***	4.95***
			(0.17)	(0.17)			(0.18)	(0.19)			(0.31)	(0.31)
foreign1st generation	-5.20***	-4.76***	-4.76***	-2.10***	-8.30***	-7.64***	-7.66***	-5.46***	-4.50***	-4.22***	-4.22***	-2.39***
	(0.18)	(0.18)	(0.18)	(0.24)	(0.18)	(0.18)	(0.18)	(0.24)	(0.17)	(0.17)	(0.17)	(0.24)
foreign 2nd generation	-3.57***	-3.11***	-3.12***	-3.24***	-3.74***	-3.02***	-3.05***	-3.33***	-2.22***	-1.96***	-2.08***	-2.19***
	(0.16)	(0.16)	(0.16)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.19)	(0.20)	(0.20)	(0.20)
Vocational									-9.77***	-9.80***	-8.77***	-8.73***
									(0.28)	(0.28)	(0.29)	(0.29)
Lyceum									9.15***	8.90***	6.50***	6.51***
									(0.24)	(0.24)	(0.29)	(0.29)
Constant	74.33***	74.92***	74.78***	74.78***	59.91***	60.80***	61.02***	61.02***	65.37***	65.14***	64.78***	64.70***
	(0.26)	(0.26)	(0.30)	(0.31)	(0.36)	(0.36)	(0.36)	(0.37)	(0.45)	(0.44)	(0.45)	(0.45)
Macro area dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	410800	379405	379405	368234	436670	412474	412474	399229	388451	371575	371575	371250
R-squared	0.078	0.086	0.086	0.089	0.167	0.178	0.179	0.182	0.303	0.308	0.316	0.317

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data. Standard errors are clustered at school level; *** p<0.01, ** p<0.05, * p<0.1.

Table 8. OLS estimates: Length of stay of first generation immigrants

Dependent variable: standardized language test results	5th grade - primary school				6th grade - lower secondary school				10th grade - upper secondary school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
gender	0.59***	0.58***	0.58***	0.57***	2.95***	2.93***	2.94***	2.93***	2.48***	2.43***	2.61***	2.61***
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.11)	(0.11)	(0.11)	(0.11)
escs	2.93***	2.48***	2.47***	2.45***	4.91***	4.11***	3.83***	3.81***	1.37***	0.76***	0.40***	0.38***
	(0.04)	(0.04)	(0.03)	(0.03)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
dialect	-1.50***	-1.36***	-1.34***	-1.35***	-3.37***	-3.10***	-2.99***	-3.01***	-1.09***	-1.09***	-0.83***	-0.87***
	(0.09)	(0.09)	(0.09)	(0.09)	(0.11)	(0.10)	(0.10)	(0.10)	(0.12)	(0.12)	(0.12)	(0.12)
foreign language	-3.43***	-3.22***	-3.22***	-2.95***	-5.66***	-5.26***	-5.29***	-4.86***	-4.41***	-4.19***	-4.14***	-2.93***
	(0.14)	(0.14)	(0.14)	(0.15)	(0.15)	(0.15)	(0.15)	(0.14)	(0.20)	(0.20)	(0.20)	(0.18)
no. stud_class	0.00	0.00	-0.00	-0.00	0.17***	0.16***	0.13***	0.13***	0.16***	0.16***	0.09***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
other european				-3.39***				-3.78***				-2.25***
				(0.31)				(0.31)				(0.27)
non europe				-5.55***				-6.28***				-5.91***
				(0.28)				(0.27)				(0.28)
n. siblings		-1.01***	-1.01***	-0.97***		-1.59***	-1.59***	-1.52***		-0.34***	-0.31***	-0.23***
		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)
manybooks		2.13***	2.13***	2.11***		3.00***	3.03***	3.01***		2.28***	2.11***	2.10***
		(0.06)	(0.06)	(0.06)		(0.08)	(0.08)	(0.08)		(0.07)	(0.07)	(0.07)
school_size			0.00	0.00			0.00**	0.00**			0.01***	0.01***
			(0.00)	(0.00)			(0.00)	(0.00)			(0.00)	(0.00)
escs_school			0.09	0.11			1.49***	1.51***			4.95***	4.95***
			(0.17)	(0.17)			(0.19)	(0.19)			(0.31)	(0.31)
foreign 2nd generation	-3.41***	-2.98***	-2.99***	-3.19***	-3.42***	-2.73***	-2.75***	-3.04***	-1.71***	-1.47***	-1.59***	-2.07***
	(0.16)	(0.16)	(0.16)	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.19)	(0.20)	(0.19)	(0.19)
one_year	-11.32***	-10.88***	-10.89***	-7.95***	-15.23***	-14.57***	-14.61***	-11.01***	-13.27***	-12.98***	-12.96***	-10.12***
	(0.39)	(0.39)	(0.39)	(0.41)	(0.31)	(0.31)	(0.31)	(0.34)	(0.64)	(0.65)	(0.65)	(0.67)
two_4years	-3.80***	-3.43***	-3.44***	-0.83***	-6.67***	-6.05***	-6.09***	-2.98***	-6.86***	-6.69***	-6.68***	-4.55***
	(0.26)	(0.26)	(0.26)	(0.28)	(0.29)	(0.29)	(0.29)	(0.31)	(0.37)	(0.37)	(0.36)	(0.38)
five_7years	-3.44***	-3.04***	-3.05***	-0.13	-5.56***	-4.96***	-4.98***	-1.66***	-2.64***	-2.40***	-2.38***	-0.43
	(0.21)	(0.22)	(0.22)	(0.26)	(0.22)	(0.21)	(0.21)	(0.26)	(0.32)	(0.32)	(0.32)	(0.35)
seven_more									-1.47***	-1.14***	-1.12***	1.25***
									(0.25)	(0.25)	(0.25)	(0.31)
Vocational									-9.79***	-9.81***	-8.78***	-8.71***
									(0.28)	(0.28)	(0.29)	(0.29)
Lyceum									9.16***	8.91***	6.51***	6.51***
									(0.24)	(0.24)	(0.29)	(0.29)
Constant	74.33***	74.93***	74.78***	74.77***	59.86***	60.77***	61.01***	60.98***	65.32***	65.10***	64.74***	64.67***
	(0.26)	(0.26)	(0.31)	(0.31)	(0.37)	(0.36)	(0.37)	(0.37)	(0.45)	(0.44)	(0.45)	(0.45)
Macro area dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	399343	368826	368826	368234	423262	399856	399856	399229	388451	371575	371575	371250
R-squared	0.081	0.089	0.089	0.090	0.170	0.180	0.182	0.183	0.303	0.308	0.317	0.318

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data. Standard errors are clustered at school level; *** p<0.01, ** p<0.05, * p<0.1.

Table 9. Fixed effect estimates: main results

Dependent variable: standardized language test results	5th grade - primary school				6th grade - lower secondary school				10th grade - upper secondary school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
gender	0.52***	0.50***	0.51***	0.52***	2.89***	2.86***	2.86***	2.88***	2.17***	2.14***	2.14***	2.11***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
escs	2.84***	2.43***	2.42***		4.50***	3.80***	3.78***		0.35***	0.01	-0.00	
	(0.02)	(0.02)	(0.02)		(0.03)	(0.03)	(0.03)		(0.02)	(0.02)	(0.02)	
degree_parents				6.88***				10.21***				-0.03
				(0.07)				(0.08)				(0.06)
high school_parents				4.61***				7.30***				0.83***
				(0.06)				(0.07)				(0.05)
compulsory_parents				1.93***				3.22***				0.31***
				(0.08)				(0.09)				(0.07)
dialect	-1.45***	-1.37***	-1.36***	-1.24***	-2.89***	-2.73***	-2.71***	-2.55***	-0.74***	-0.73***	-0.73***	-0.70***
	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)	(0.06)	(0.06)
foreign language	-2.95***	-2.81***	-2.62***	-2.80***	-4.97***	-4.67***	-4.32***	-4.59***	-2.65***	-2.59***	-2.41***	-2.28***
	(0.11)	(0.11)	(0.12)	(0.13)	(0.13)	(0.14)	(0.14)	(0.15)	(0.14)	(0.14)	(0.14)	(0.14)
foreign1st generation	-5.06***	-4.63***	-2.11***	-2.67***	-7.92***	-7.35***	-4.83***	-5.98***	-3.59***	-3.44***	-1.74***	-1.78***
	(0.14)	(0.15)	(0.20)	(0.21)	(0.15)	(0.16)	(0.23)	(0.24)	(0.13)	(0.13)	(0.19)	(0.19)
foreign 2nd generation	-3.44***	-2.98***	-3.15***	-3.55***	-3.44***	-2.85***	-3.11***	-4.03***	-1.55***	-1.44***	-1.54***	-1.46***
	(0.12)	(0.13)	(0.13)	(0.14)	(0.15)	(0.16)	(0.16)	(0.17)	(0.14)	(0.14)	(0.14)	(0.14)
other european			-2.88***	-2.68***			-2.28***	-1.64***			-0.83***	-0.66***
			(0.29)	(0.31)			(0.30)	(0.33)			(0.22)	(0.22)
non europe			-4.98***	-4.60***			-5.21***	-4.60***			-3.89***	-3.70***
			(0.26)	(0.29)			(0.27)	(0.30)			(0.23)	(0.23)
n. siblings		-0.97***	-0.92***	-0.91***		-1.39***	-1.34***	-1.35***		-0.16***	-0.11***	-0.09***
		(0.02)	(0.02)	(0.02)		(0.03)	(0.03)	(0.03)		(0.02)	(0.02)	(0.02)
manybooks		2.01***	1.99***	2.56***		2.61***	2.53***	3.76***		1.46***	1.46***	1.50***
		(0.05)	(0.05)	(0.05)		(0.06)	(0.06)	(0.06)		(0.04)	(0.04)	(0.04)
Constant	74.97***	75.66***	75.59***	72.25***	62.13***	63.05***	63.04***	57.96***	67.94***	67.80***	67.75***	67.58***
	(0.02)	(0.04)	(0.04)	(0.06)	(0.03)	(0.05)	(0.05)	(0.07)	(0.03)	(0.04)	(0.04)	(0.05)
Classroom fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	410800	379405	368234	348474	436670	412474	399229	372005	388451	371575	371250	363381
R-squared	0.081	0.090	0.093	0.094	0.129	0.138	0.141	0.145	0.023	0.026	0.027	0.028

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data. Classroom-level fixed effect estimates. *** p<0.01, ** p<0.05, * p<0.1.

Table 10. Fixed effect: length of stay of first generation immigrants

Dependent variable: standardized language test results	5th grade - primary school				6th grade - lower secondary school				10th grade - upper secondary school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
gender	0.53***	0.51***	0.50***	0.52***	2.90***	2.87***	2.86***	2.88***	2.17***	2.14***	2.14***	2.11***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
escs	2.84***	2.44***	2.42***		4.51***	3.81***	3.78***		0.36***	0.01	-0.00	
	(0.02)	(0.02)	(0.02)		(0.03)	(0.03)	(0.03)		(0.02)	(0.02)	(0.02)	
degree_parents				6.87***				10.18***				-0.03
				(0.07)				(0.08)				(0.06)
high school_parents				4.60***				7.28***				0.82***
				(0.06)				(0.07)				(0.05)
lower secondary_parents				1.93***				3.20***				0.31***
				(0.08)				(0.09)				(0.07)
dialect	-1.43***	-1.35***	-1.35***	-1.24***	-2.83***	-2.67***	-2.68***	-2.54***	-0.72***	-0.71***	-0.73***	-0.69***
	(0.06)	(0.06)	(0.06)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)	(0.06)	(0.06)
foreign language	-3.13***	-2.97***	-2.72***	-2.81***	-5.49***	-5.14***	-4.73***	-4.70***	-3.39***	-3.30***	-2.43***	-2.32***
	(0.11)	(0.11)	(0.11)	(0.12)	(0.13)	(0.14)	(0.13)	(0.15)	(0.13)	(0.13)	(0.13)	(0.13)
foreign 2nd generation	-3.31***	-2.87***	-3.09***	-3.54***	-3.11***	-2.55***	-2.88***	-3.97***	-1.25***	-1.15***	-1.53***	-1.43***
	(0.12)	(0.12)	(0.13)	(0.14)	(0.15)	(0.16)	(0.16)	(0.17)	(0.14)	(0.14)	(0.14)	(0.14)
other european			-3.39***	-2.89***			-3.97***	-2.41***			-1.83***	-1.67***
			(0.28)	(0.30)			(0.28)	(0.32)			(0.19)	(0.20)
non europe			-5.34***	-4.68***			-6.47***	-4.97***			-4.57***	-4.39***
			(0.24)	(0.28)			(0.25)	(0.29)			(0.20)	(0.20)
n. siblings		-0.97***	-0.92***	-0.91***		-1.40***	-1.34***	-1.35***		-0.17***	-0.11***	-0.09***
		(0.02)	(0.02)	(0.02)		(0.03)	(0.03)	(0.03)		(0.02)	(0.02)	(0.02)
manybooks		2.01***	1.99***	2.56***		2.56***	2.54***	3.76***		1.48***	1.47***	1.50***
		(0.05)	(0.05)	(0.05)		(0.06)	(0.06)	(0.06)		(0.04)	(0.04)	(0.04)
one_year	-10.96***	-10.53***	-7.68***	-8.32***	-14.64***	-14.12***	-10.41***	-12.40***	-12.64***	-12.44***	-10.20***	-9.99***
	(0.33)	(0.34)	(0.35)	(0.38)	(0.29)	(0.29)	(0.32)	(0.36)	(0.56)	(0.57)	(0.57)	(0.60)
two_4years	-3.74***	-3.36***	-0.82***	-1.80***	-6.44***	-5.92***	-2.67***	-5.14***	-6.73***	-6.59***	-4.88***	-4.95***
	(0.22)	(0.22)	(0.24)	(0.26)	(0.26)	(0.26)	(0.29)	(0.31)	(0.28)	(0.28)	(0.29)	(0.30)
five_7years	-3.44***	-3.01***	-0.17	-1.20***	-5.25***	-4.76***	-1.31***	-3.57***	-2.65***	-2.49***	-0.90***	-0.88***
	(0.18)	(0.18)	(0.22)	(0.24)	(0.19)	(0.19)	(0.24)	(0.26)	(0.23)	(0.24)	(0.25)	(0.25)
seven_more									-1.37***	-1.20***	0.73***	0.60***
									(0.16)	(0.16)	(0.19)	(0.20)
Constant	74.92***	75.60***	75.58***	72.25***	62.09***	63.03***	63.00***	57.97***	67.90***	67.77***	67.74***	67.56***
	(0.02)	(0.04)	(0.04)	(0.06)	(0.03)	(0.05)	(0.05)	(0.07)	(0.03)	(0.04)	(0.04)	(0.05)
Classroom fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	399343	368826	368234	348474	423262	399856	399229	372005	388451	371575	371250	363381
R-squared	0.084	0.093	0.095	0.096	0.131	0.141	0.143	0.147	0.025	0.028	0.030	0.030

Notes: Source: Invalsi data for academic year 2010-11. See the List of Variables in the Appendix for additional details regarding data. Classroom-level fixed effect estimates. *** p<0.01, ** p<0.05, * p<0.1.