

DISCUSSION PAPER SERIES

IZA DP No. 16739

Ethnic Identity and Educational Outcomes

Teresa Randazzo
Matloob Piracha

JANUARY 2024

DISCUSSION PAPER SERIES

IZA DP No. 16739

Ethnic Identity and Educational Outcomes

Teresa Randazzo

University of Naples Parthenope, Fondazione CMCC and RFF-CMCC EIEE

Matloob Piracha

University of Kent and IZA

JANUARY 2024

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Ethnic Identity and Educational Outcomes*

We study the role of immigrant children's ethnic identity in their educational performance and preferences/aspirations in Italy. We find that students with a weak sense of Italian belonging show a low performance in reading and mathematics and higher probability of grade retention. Moreover, children in middle secondary school with a weak sense of Italian identity have a low preference towards academically-oriented high secondary track which normally increases the likelihood of pursuing a university degree. We also find that the intention of immigrant children in high secondary schools to enrol at university decreases if they have a weak Italian identity. We exploit gender heterogeneity finding that females are more adversely affected in their educational aspirations when they have not built a strong sense of Italian identity. Immigrant children will soon form a very important component of the Italian labour force and shedding light on their educational outcomes will help us understand their performance in the Italian labour market better.

JEL Classification: F22, J15, I2, Z13

Keywords: ethnic identity, educational performance, educational preferences

Corresponding author:

Matloob Piracha
School of Economics
University of Kent
CT2 7ND
United Kingdom
E-mail: m.e.piracha@kent.ac.uk

* We would like to thank Raefaele Lagravinese, Vincenzo Lombardo, participants at the GLO Global Conference, Annual Conference of Italian Society of Development Economists (SITES), 34th Annual European Association for Evolutionary Political Economy (EAEPE), Workshop on Trade Unions and Migration (IAAEU), 36th National Conference of Labour Economics (AIEL), 13th Economics of Global Interactions Conference as well as seminar participants at the Universities of Ghent and Lille. The authors are solely responsible for any remaining errors.

1. Introduction

Ethnic identity and its impact on the labour market outcomes of both first- and second- generation migrants has attracted significant attention in economics research in the recent past (see, for instance, Cai and Zimmermann, 2024; Carillo *et al*, 2023; Piracha *et al*, 2023; Casey and Dustmann, 2010; Nekby and Rodin, 2010; Constant *et al*, 2009).¹ A few studies have also analysed the association between ethnic identity of adult immigrants and their education (e.g. Nekby *et al*, 2009; Zimmermann *et al*, 2008) as well as immigrant parents' ethnic identity and their children's education (e.g., Schuller, 2015). We contribute to the literature by analysing the association of ethnic identity of the middle- and high-school pupils on their test scores in reading and mathematics as well as their aspirations for the type of further education, i.e., academic, vocational or technical track after middle school and whether to go to university for those in high secondary school. Since education is an important determinant of economic outcomes, shedding light on immigrants' educational outcomes and choices will help us understand immigrants' investment in education, which has consequences on their labour market performance.

The limited amount of research in economics that is related to children's identity and schooling is primarily on studying the role of racial identity and almost all of that research uses data from the US and focuses on school performance and not on intentions to acquire more education, e.g., going to university.² For instance, Akerlof and Kranton (2002) link identity and schooling by analysing how pupils place themselves in different social categories, e.g., "nerds", "jocks" etc. Race is one such category and it has been observed in the US schooling system that black students tend to form oppositional identities as some of them refuse to "act" white by shunning whites' social norms (including spoken English) while others adopt it (see also Akerlof and Kranton, 2000).³ Similarly, Patacchini and Zenou (2016) analyse the role of networks formed by black and white pupils in a number of school districts in the US. They use friendships at schools to construct a measure of racial identity by observing the number of same-race friends that individuals choose as their best friends. Based on the observation that most prefer to be friends with pupils of their own race, they analyse the impact racial identity has on school performance. They find that having a high percentage of same-

¹Ethnic identity is the feeling of belonging to a specific ethnic or cultural group. At one extreme, immigrants may reject the dominant culture, whereas at the other extreme they may shun their own culture (language and/or religion) in favour of the dominant one (Battu and Zenou, 2010). A particular ethnic group can therefore form 'oppositional identities' where some belong to one group while others belong to the second group.

² There is a vast literature in social psychology that has shown the significance of ethnic identity in educational outcomes (see for instance, Virta and Westin, 1999; Phinney *et al.*, 2001). In the cross-cultural psychology literature, the importance of a strong identification with the background culture for school adaptation as well as individual well-being, self-esteem, and a sense of belonging to the majority (host country) culture are stressed suggesting a positive association between ethnic identity and education.

³ Battu and Zenou (2010) find similar results for the UK.

race friends has a positive impact on white pupils' test score while it has a negative effect on black teenagers' scores, although the negative effect is somewhat mitigated with the black pupil parent's education level.

There are two strands in the scant literature on the association between ethnic identity and education. The first strand looks at whether education acquired in the origin country has an impact on ethnic identity in the destination country. For instance, Constant *et al* (2009) use a two-dimensional identity measure that they call the *ethnosizer*, which assumes four possibilities: assimilation, integration, marginalization and separation⁴ and show that migrants who acquired higher education prior to migrating to Germany tend to integrate rather than assimilate. Using the same German data, Zimmermann (2007) finds that human capital acquired in the country of origin has no effect on migrants' identifying with the majority culture.

The other strand considers the association of ethnic identity with the education attainment of migrants in the host country. There are two related papers to ours in the second strand. The first one, Nekby *et al* (2009), analyses the role of ethnic identity of second generation immigrants in the completion of higher education degree. Using data from Sweden, they construct a two-dimensional measure of ethnic identity *a' la* Berry (1980, 1997) and Constant *et al* (2009). They find that migrant men who identify with both the host and home cultures, i.e., the integrated group, are more likely to attain tertiary education than the assimilated or the marginalised. The second paper, Schuller (2015), explores the association between parents' ethnic identity and their children's education outcomes. Using Germany Socio-economic Panel data she finds that integrated, rather than assimilated or separated family environment is more conducive for second generation pupils' education attainment, which is measured by the type of school chosen – lower, intermediate or upper level. However, unlike the two papers, we study the role of children's ethnic identity in their education outcomes. It has been argued that even after controlling for usual characteristics like parental human capital and other characteristics, neighbourhood, school quality etc, there are still differences in school performance between different groups of pupils, e.g., black and white (see Pattacchini and Zenou, 2016). Literature shows that peer effects and other aspects play a crucial role in academic achievement (Hoxby, 2000; Meschi and Pavese, 2023). It is therefore important to go a step further and study how a pupil's social identity correlates with their educational outcomes. To the best of our knowledge, this is the first

⁴ Assimilation is one extreme where the immigrant fully adopts the native culture, which usually means she prefers the 'white social network', while giving up the culture of her country of origin. Separation, on the other hand, is the opposite of assimilation where the immigrant only identifies with her own ethnic background. The other two possibilities are integration and marginalization, which are defined as follows: in the former case the migrant keeps her own cultural traits (e.g., speaking her native language at home) but adopts significant aspects of the host country culture whereas latter is the weak dedication to both home and host country culture.

paper that explores the association between school going teenage immigrants' ethnic identity and their education performance and aspirations/preferences.

Another marginally related paper to ours is Carlana *et al* (2022), who also use Italian schooling data like us but their focus is different from ours. They show that immigrant pupils are more likely to enrol in vocational or technical education rather than academically-oriented schools compared to the natives of the same ability level. They call this phenomenon "education segregation" (they then analyse the impact of an intervention to help immigrant children achieve better educational outcomes, which is the main focus of their paper).⁵ Since part of the reason for this segregation maybe linked to the network effects and identity, it is important to understand whether, and to what extent, identity plays a part in the educational outcomes of immigrants.

Using survey data from the Istituto Nazionale di Statistica (Istat), we analyse the association between ethnic identity of teenage immigrants on their educational outcomes including plans to pursue further education. To fix ideas, before conducting our main analysis, we first explore the data to show the overall differences in educational performance and aspirations between native born and immigrant pupils. The existence of the ethnic gap in academic achievements is well documented, with existing literature showing that immigrant children tend to be disadvantaged students coming from lower socio-economic backgrounds (Ammermüller, 2005; Frick and Wagner, 2001; Gang and Zimmermann, 2000; Dustmann *et al*, 2012). However, the source of this gap remains poorly understood. Since part of the educational gap between the two groups will be explained by differences in the usual observables, e.g., individual characteristics, family income, parental education etc., we conduct Blinder-Oaxaca decomposition to determine the extent to which the gap is explained by the usual observables. The results show that 50% (or more) of the differences in the math and reading scores for children in middle and high secondary school is unexplained by the observables.⁶ We therefore argue that ethnic identity must be, as it is in the labour market outcomes, one of the key elements in explaining the differences. This is in line with some of the studies cited above.

Next, we conduct our main analysis on the role of ethnic identity in educational outcomes. Since assimilation is across a spectrum of different measures, "it is practically impossible to summarize the overall balance of migrant's commitments by one response" (Constant and Zimmermann, 2008: 428). We therefore use the ethnic identity index along the lines of the one-dimensional 'ethnosizer'

⁵ Their paper is primarily about evaluating the intervention which consisted of (i) psychologically grounded career choice consultancy and (ii) academic tutoring, both provided to immigrant children in various schools. The programme, called "Equality of Opportunity for Immigrant Students" was run in collaboration with the Ministry of Education and three Bank foundations.

⁶ Schindler (2007) also shows that the socio-economic status explains only 50% of the ethnic test score gap between immigrants and natives.

developed by Constant *et al* (2009)⁷, which is an average measure of five indicators: (i) language proficiency, (ii) culture, (iii) ethnic self-identification, (iv) migration history and (v) social network. However, unlike Constant *et al* (2009), we use the Principal Component Analysis (PCA) to construct the index using 11 components we identified in the data.⁸ PCA allows us to address the inherent multicollinearity of the different components of our identity measure. Language, which consists of 4 components, is the average respondent-assigned value for language use. Culture is the average value for visible cultural aspects based on three components. Self-identification is the mean value for ethnic self-identity. Migration history is the average value for migration-related questions, which relates to future plans. Social network consists of two components capturing the friendship group. An individual who scores zero on each of these components is identical to a native. Therefore, the lower the value of the ethnic identity index the higher is the level of assimilation. After controlling for a number of individual and family characteristics, we explore the association of our ethnic identity index to immigrant children's school performance and aspirations for further education. Our results show that having a weak sense of Italian identity is negatively associated with educational outcomes, i.e., weak identity is correlated with lower scores in math and reading, and in terms of aspirations, strong identity is strongly correlated for middle (high) secondary school immigrants opting for academically-oriented (university) track, both of which are likely to be economically more rewarding in later life.

Finally, we explore gender heterogeneity on educational outcomes. Our linear model shows that overall a weak sense of Italian identity has a larger negative correlation for girls' aspiration compared to boys, especially for the choice of going to the university, suggesting that girls are more sensitive to the assimilation process in their potential investments in education.

It is important to point out that none of the results discussed should be considered as causal. This is a common feature in this literature since finding a reasonable instrument in this setting is not possible, especially when the data is limited in terms of information about the immigrants (see, for instance, Piracha *et al*, 2023; Schuller, 2015; Casey and Dustmann, 2010; Pendakur and Pendakur, 2005; Nekby and Rodin, 2010). However, we run a number of tests to check the robustness of our results.

⁷ The data set does not allow us to construct a two-dimensional ethnosizer.

⁸ The 11 components consist of four components linked to language – reading, writing, speaking, understanding; three components of culture – thinking in Italian, speaking Italian at home and watching Italian TV; self-identification, which relates to the question whether the child considers himself or herself as Italian; migration history, which is based on the question whether the child wants to live in Italy in later life; and finally two components linked to social networks – whether they have Italian friends at school and whether they have Italian friends outside of school. More details on each of the categories is provided in Section 4.

Rest of the paper is organised as follows. Section 2 outlines secondary education system in Italy. In Section 3 we use Blinder-Oaxaca decomposition to show how much of the difference in school performance and aspirations for education between natives and immigrants is explained by the observable characteristics. Section 4 explains the data set including the key variables used in the construction of the ethnic identity index. Empirical approach is presented in Section 5 while Section 6 reports the results. Concluding remarks appear in the last section.

2. Education and immigration in Italy

Children in Italy start formal school at the age of 6 years and are required to attend school until the age of 16. Pre-university education comprises five years of primary school (from grade 1 to grade 5), three years of middle school (from grade 6 to grade 8) and five years of high school (from grade 9 to grade 13). The Italian school system is mainly public and it doesn't divide students based on their skills or abilities for subjects like math or science. Instead, all students learn together in the same class for every subject, taught by the same teachers. Classes typically have between 18 and 27 students, and the weekly schedule involves 30 hours of instruction, with reading and math occupying the most instruction hours.⁹ The Ministry of Education defines the academic disciplines, time of instruction, educational programmes and their content ensuring consistency across all Italian schools. However, the assessment in each school is carried out by teachers in the exercise of their professional autonomy, in accordance with the criteria and methods defined by the teaching staff and included in the plan of the educational offering.¹⁰

While primary and middle schools are homogenous in their track, students can choose among three types of high schools: academic-oriented (*liceo*), technical or vocational. The three tracks have the same duration (five years) but they differ in difficulty, prestige and career orientation. Vocational schools provide practical training preparing students for immediate employment in manual low-skilled jobs. Technical and academic schools provide comprehensive knowledge in math, humanities and science. However, academic schools offer a more theoretical background and increase the likelihood to pursue a university degree whereas technical schools complement theory with practical training in non-manual jobs. In general, academic and technical schools offer much better education and employment prospects than vocational schools.

⁹ In the Italian education system, the term "reading" includes focus on developing reading comprehension as well as analytical and interpretative skills. The curriculum typically includes also subjects such as Italian literature, literature from other countries and history. The specific subjects may vary depending on the level of education. Math measures skills in problem-solving, logic and interpretation of quantitative phenomena.

¹⁰ Details could be found at <https://www.gazzettaufficiale.it/eli/id/2017/05/16/17G00070/sg>.

A significant increase in the foreign-born population over the past two decades has had a profound effect on schools and universities in Italy. In terms of migration flows, the foreign born population increased significantly from 1990s to the start of the millennium, going from around 800,000 to over 2 million. According to the 2020 census, that number has more than doubled in the last two decades to over 5 million. If we also take into consideration naturalized citizens (Italians by acquisition) and mixed race individuals, the number goes up by approximately another two million, bringing the foreign population from 10 to over 12 percent of the total population of Italy. The number of migrant children in Italian schools was nearly 900,000 in the academic year 2021/2022.¹¹ In the 2018/2019 school year, students with non-Italian citizenship represented 10% of the overall student population, and this number was even higher for primary and middle schools (11.5%). Compared to 10 years ago, migrant students in secondary schools have increased by 39%, representing 7.4% of students in this category. There are regional disparities in terms of the number of migrant students with a significant majority (65%) being concentrated in the northern regions while 22% are found in the central regions and just over 13% in the South. The largest groups of foreign students are Romanians and Albanians. Among African students, Moroccans are the biggest group. China is the country of origin of over 55,000 students (6.4%), with a growth of almost 80% in the last 10 years (Istat, 2020).

According to the provisions of the Ministry of Education, the number of students with non-Italian citizenship with reduced knowledge of the Italian language should not normally exceed 30% of the students in a class or school. However, there are exceptions. The schools with a number of students with non-Italian citizenship above the 30% threshold are 6.5% and have tripled in the last 10 years.

While all 6-13 year old immigrant children attend school, the rate goes down to 90% for the 14-16 age group and by the age 17 only 2/3rd remain in school (compared to 80.7% of Italians of the same age group). This phenomenon is more worrying for the 17-year-old male immigrant children, whose schooling rate is 59.9%, compared to 76.6% for the females. This is partly influenced by the fact that the failure rate is higher for male immigrant pupils and therefore there is a higher number who repeat a year, which increases the possibility of dropout. Related to this is their academic achievement and the type of education they choose - approximately 44% of the immigrant pupils choose an academic-oriented school, whereas 57% Italians make the same choice (Istat, 2020).

¹¹ “Ministero dell’istruzione e del merito - Ufficio di Statistica”
https://www.miur.gov.it/documents/20182/7715421/NOTIZIARIO_Stranieri_2122.pdf/2593fc66-1397-4133-9471-b76396c2eb97?version=1.1&t=1691593500475

3. The Preamble

To fix ideas, we start by documenting that an ethnic gap exists in educational achievements. We use data from the “integration of second generation survey”¹² conducted in 2015 by “Istituto Nazionale di Statistica” (Istat) and financed by the European Integration Fund.¹³ The survey is based on a sample of state secondary schools with at least 5 migrant students.

The sample consists of 1,419 schools across the country in 821 municipalities. The survey took place during school hours, in classrooms or school laboratories in which students filled out an online questionnaire in the presence of a municipal officer. A total of 68,127 students were interviewed out of which 36,440 were Italians and 31,687 were first or second generation immigrants with a foreign citizenship.¹⁴

Figure 1 shows that native students get a higher score both in reading and math at various points in their educational careers. The gap is larger for students in middle secondary school (from grades 6 to 8) and gets negligible at the end of high secondary school (after grade 10). However, it is not clear that the performance gap between native and immigrants gets smaller only because of the longer time spent by immigrants in the Italian education system. For instance, it could be the case that those with lower educational ability drop out by the time they reach high secondary school. Furthermore, immigrant students have a higher probability of grade retention across the school years.

Figure 2 plots preferences of high secondary school track for students who are enrolled in middle secondary schools. Students make their final track choice at the end of grade 8. Across the different middle secondary school grades, there is a persistence of preference of native students for academically-oriented track and an increasing trend for the preference of a vocational track for immigrant students.¹⁵

Surprisingly, looking at students enrolled in higher secondary school we find that immigrants have a higher interest in pursuing university education compared to natives (Figure 3). This is possibly a result of a selection mechanism. Immigrant students show a higher rate of school dropout and it is

¹² Even though the data set refers to “second generation” it is important to point out that a large number of children were actually not born in Italy and therefore in that respect are not strictly second generation. Approximately 70% of immigrant children were born abroad. However, since a majority of them have attended Italian schools since pre-school or elementary school (82%), they arrived in Italy at a very young age.

¹³ Istat carried out the survey on the integration of immigrants between March and June 2015 in collaboration with the Ministry of Education, University and Research. The survey is available on the Istat website at <http://www.istat.it/it/archive/182866>.

¹⁴ The status of foreign citizenship persists also for children born in Italy if their parents do not hold Italian citizenship.

¹⁵ Enrolment in university is always possible from the technical track and vocational schools. However, very few students who attend vocational education pursue tertiary education (Carlana *et al* 2022).

possible that those who had no intention to pursue further education have already left school. Instead, those who remain in school have a higher motivation for pursuing further education.

One of the main reasons of the gap in educational achievement and aspiration is likely to be related to individual characteristics and family background. Our data shows that, as expected, immigrant students come from families with a lower level of wealth and are more likely to have a parent (mother or father) living outside the household, not working and less educated compared to natives. Immigrant students are also less likely to play a sport and more likely to work after school hours (see Table 1).

However, it is important to explore whether the differences in individual and family background are enough to explain mean educational differences between native and immigrant students. We use the Oaxaca-Blinder decomposition (Blinder 1973; Oaxaca 1973) to show educational outcome differential between native and immigrant students into a part that is “explained” by group differences in observed characteristics and a residual part that cannot be accounted for by such differences in educational determinants. This “unexplained” part is often used as a measure for discrimination and segregation, but it also subsumes the effects of group differences in unobserved predictors (Jann, 2008).

We use the linear model

$$Y = X'\beta + \varepsilon$$

where X is a vector containing the predictors and a constant, β contains the slope parameters and the intercept, and ε is the error. The question now is how much of the mean outcome differences is accounted for by group differences in the predictors.

$$R = E(Y_N) - E(Y_M) = E(X_N)'\beta_N - E(X_M)'\beta_M$$

where $E(Y)$ denotes the expected value of the outcome variable for natives (N) and immigrants (M). The outcome difference can be written as

$$R = \{E(X_N) - E(X_M)\}'\beta^* + \{E(X_N)'\}(\beta_N - \beta^*) + E(X_M)'(\beta^* - \beta_M)$$

We follow a twofold decomposition

$$R = Q + U$$

where the first component $Q = \{E(X_N) - E(X_M)\}'\beta^*$ is the outcome differential that is explained by group differences in the predictors, and the second component $U = \{E(X_N)'\}(\beta_N - \beta^*) + E(X_M)'(\beta^* - \beta_M)$ is the unexplained part which is usually attributed to discrimination or captures potential effects in unobserved variables.

The Oaxaca-Blinder decomposition shows that the gap in educational performance (Table 2) and further educational preferences (Table 3) between natives and immigrants are not fully explained by difference in characteristics and family background. There is a relevant unexplained part which contributes in determining different outcomes between the two groups. For example, in math score a gap of 51% remains unexplained between native and immigrant students in middle secondary school. The math gap is even larger, around 73%, when we focus on native and immigrants students in high secondary school. This suggests that migration status itself plays a role on top of general socioeconomic disadvantage. One of the key elements identified in recent migration literature is that the level of belonging to the host culture may represent a barrier in the labour market (Carillo *et al.*, 2023, Piracha *et al.*, 2023, Constant *et al.*, 2009) and it is very likely that the same barrier leads to the gap in educational achievements and aspirations that results in immigrant students choosing less demanding educational paths (see Virta and Westin, 1999; Phinney *et al.*, 2001). In the rest of the paper we explore how strongly ethnic identity, i.e., the level of belonging to the host culture, is correlated with educational aspirations of the immigrant pupils.

4. Immigrant sample and explanation of variables

In our main analysis we restrict the attention to immigrants in secondary school. Our final sample, after excluding students with missing relevant information, consists of 12,606 students in middle secondary school (6th to 8th grade) and 14,647 students in high secondary school (9th to 13th grade) with no Italian citizenship.¹⁶ Only 19% (41%) of the immigrant sample in high (middle) secondary school were born in Italy.

Table 4 reports the descriptive statistics. In our sample, 47% of students in middle school are female and 53% are male.¹⁷ The percentages are reversed in high secondary school, which suggests that a few male pupils dropout after grade 8. Around 53% of students in middle school play a sport; the number drops to 45% for those in high school. Approximately 30% of middle secondary and 32% of high secondary school pupils work at least a few times a month, though we do not have information on the types of jobs they do. Given the age of children in middle secondary, it is likely that the ‘work’ is to support household activities with simple tasks as child labour in Italy is prohibited under the age of 15. With respect to the location, only 20% and 31% of middle and high secondary school pupils, respectively, live in a big city showing that a large share of immigrant families with children prefers living in small cities or towns. Finally, Veneto, Lombardia Piemonte and Emilia Romagna are the

¹⁶ Some safeguarding measures were adopted to protect the identity of the students, which means that a few key variables are not available in the public dataset, e.g., age of the student. We know students normally start the 6th grade at age 10 or 11 and conclude the 13th grade at age 18 or 19.

¹⁷ Descriptive statistics by gender are presented in Table A2 in the Appendix.

regions with a higher share of immigrant students, which is due to the fact that these regions have the highest number of immigrant population in Italy.

In terms of parental characteristics, immigrant pupils generally come from relatively well educated families. More precisely, the highest level of parental education for around 44% of the high secondary pupils is at the secondary level while 21% of this respondents' group has at least one parent holding a university degree. The number drops for middle secondary pupils to 25% having a secondary education as highest level of parental education and those having a parent holding a degree are just 14%. Mother does not work for around 36% (middle secondary school students) and 34% (high secondary school students) of the respondents and that percentage reduces to 15% (middle secondary school students) and 18% (high secondary school students) when the question refers to non-working fathers.

Looking at the household composition, 20% (25.8%) of our sample in middle (high) secondary school have no sibling. Around 3% (5%) of the pupils in middle (high) secondary school do not live with their mother while the percentage of pupils reporting that their father does not live with them is 20% for middle secondary and 27% for high secondary pupils.

In addition to the individual and household characteristics, Table 4 reports some relevant school information. We know which grade students are attending and whether they trust their teacher.¹⁸ According to 79% (59%) of students in middle (high) secondary school, teachers are a reference point and the students trust them. Finally, on average, immigrant students in middle and high secondary school started Italian education at the elementary school at grade 2 and 4, respectively.

Our outcome variables are related to school performance and aspiration. With respect to school performance, students report the most recent grade in reading and mathematics. Scores go from 1 to 10. In both subjects the average score is around 6, which is just a pass mark. Another indicator of school performance is whether respondent has repeated a grade. Approximately 18% and 30% in the middle and high secondary school respondents, respectively, have repeated a grade. As expected, the incidence of repeating a grade (called "grade retention" in the table) increases with the level of education. In terms of aspirations, the summary statistics show that on average 33% of students in the middle secondary school are planning to go to an academically-oriented school, 24% prefer a technical track and 20% prefer to pursue vocational education. Only about 10% of the cohort is

¹⁸ In education literature, trust has been shown to have a significant effect on student achievement (see, for instance, Adams and Forsyth, 2013; Forsyth *et al*, 2011; Van Maele and Van Houtte, 2011).

intending to go for alternative option (e.g. professional course or working). Finally, on average, almost 45% of migrants in high secondary school have intention to pursue tertiary education.

Table 4 about here

Our variable of interest is an ethnic identity index constructed using 11 variables which fall in the five categories identified by Constant *et al* (2009): (i) language proficiency, (ii) culture, (iii) ethnic self-identification, (iv) migration history and (v) social network. The identity index measures the level of assimilation in the destination country. We construct a set of dummy variables which takes a value of zero when the respondent behaves like a native and 1 when they have not assimilated at all. Since these variables are highly correlated to each other we follow Kalfa and Piracha (2017) and construct the ethnic identity index using the principal component analysis (PCA).

When the index takes the value equal to zero the respondent is fully assimilated in Italy. On the other hand, as the index value moves further from zero, there is a decrease in the level of commitment to Italian culture. More details about how each category is constructed is as follows.

Language: We have information on language proficiency (e.g. how well Italian is spoken, written, understood). The survey indicates for each language skill whether respondents have very good, good, fair, poor or very poor ability. A value of 0 is assigned for those reporting a very good/ good or fair command of Italian. A value of 1 is assigned if ability in Italian is poor or very poor. We consider four abilities in Italian (speak, write, understand and listen) which define the language proficiency.

Culture: It considers the visible Italian cultural elements which we identify in three variables (i) “Do you think in Italian”? (ii) “Do you watch Italian TV”? and (iii) “Do you speak Italian with your parents”? Again, a value of 0 is assigned for each of the three categories and a value of 1 otherwise.

Ethnic self-identification: This category is identified in the variable related to the question on whether the respondent feels Italian. The answer ‘yes’ gets a value of 0 and indicates full assimilation. The answers ‘no’ and ‘do not know’ get a value 1.

Migration history: This category is identified in the variable related to the question “Where do you want to live when you get older?”. The answer ‘want to live in Italy’ takes the value of zero; the alternative ‘want to live where I was born’, ‘want to live where my parents were born’, ‘want to live in another country’ take the value of 1. The willingness to remain in Italy represents the attachment to the country.

Social network: We consider the nationality of friends (i) within school and (ii) outside school. A value of 0 is assigned if friends are exclusively Italian; a value of 1 is given when friends come from mixed origin. The mean assigned values of the two answers referring to friends within and outside school generate the social network category.

Table 5 presents the specific dummy variables related to each category and the related average value. Overall, the mean value of the identity index is around 0.39, showing a good commitment of immigrant children to Italy.¹⁹ However, the level of commitment varies across different measures of identity, proving once again that it is important to capture the multifaceted aspects of assimilation (Constant and Zimmermann, 2008). The average score is close to zero for immigrant children having excellent ability in the Italian language, having Italian friends and being close to Italian cultural elements, showing strong commitment Italian culture. But the question on self-identification and the question asking whether the respondent wishes to live in Italy show that the assimilation process is a bit fragile in these two categories. Around 59% and 65% of immigrants in the middle and high secondary schools, respectively, do not identify themselves as Italian and the numbers for those who do not plan to remain in Italy permanently are 64% and 72% for the two groups of immigrants.

Table 5 around here

5. Empirical strategy

To investigate the relationship between ethnic identity and school outcomes we estimate the following model:

$$y_{ij} = \alpha + \gamma EI_{ij} + \sum_{k=1}^k \beta_k X_{ijk} + \delta_i + \eta_j + \varepsilon_{ij} \quad (1)$$

where y represents the outcome variable for respondent i in school j and region r . In particular, j indicates the secondary school level (middle or high). The outcome variables for school performance are (i) latest score in reading, (ii) latest score in mathematics, and (iii) a binary indicator for grade retention. The outcome variables for aspirations for middle school children are a binary indicator for each alternative type of high school choice: academic-oriented, vocational, technical or other, whereas for high school pupils the aspiration outcome is going to university or not. X_{ik} denotes the k

¹⁹ We present in the Appendix, Table A3, the correlation between the variables used in the PCA, which verifies that the components are sufficiently different from one another and capture various dimensions of ethnic identity. The eigenvalues and the cumulative proportion are shown in Table A4.

exogenous variables, δ and η denote fixed-effect for region and origin country, respectively, and ε is the error term. Our variable of interest is *EI*, the ethnic identity index, and the associated parameter γ captures the effect of ethnic identity on educational outcomes, conditional on individual and socio-economic characteristics. The ethnic identity index is constructed using PCA, a statistical method seeking to reduce multicollinearity by employing an orthogonal transformation. This process converts a group of explanatory variables into principal components that are uncorrelated. This reduction in dimensionality retains maximum variation in the data. Consequently, the first principal component of the chosen variable set captures the larger variation available in the data. For each of the 11 variables used to build the ethnic identity index, Table A5 in the Appendix displays the coefficients of the first principal component.

The first principal component PC_1 used as a proxy for Ethnic identity (*EI*) can be represented as a function of 11 variables and their coefficients:

$$PC_1 = 0.4629 * understand + 0.4806 * speak + 0.4549 * read + 0.4303 * write + 0.1839 * think + 0.1622 * watch\ TV + 0.1369 * speak\ at\ home + 0.1271 * self-identification + 0.0143 * migr.\ history + 0.1789 * peers + 0.1881 * friends$$

The set of characteristics included in X_{ik} are variables related to schooling path (for example, grade in which respondents are enrolled and school information), respondents' characteristics (for example, age and whether born in Italy), household characteristics (for example, parental education). The wealth position of the children's household is captured by a wealth index,²⁰ also constructed using principal component analysis (see Filmer and Pritchett, 2001).

Studying the effect of identity on educational outcomes is empirically challenging. Omitted variables, measurement error in the identity measure, reverse causality are likely to be sources of bias that prevent identification of a causal relationship. Selection of immigrants over individuals' characteristics maybe a further source of bias. In an attempt to address these potential concerns we control for a wide range of individual and family characteristics and we perform a large battery of robustness and sensitivity checks. Despite all the robustness checks we interpret our findings only as correlations without giving them any casual interpretation.²¹

²⁰ The asset ownership indicators used to construct the wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of room

²¹ This is a common feature of most of the research on this topic as finding a suitable instrument is almost impossible.

6. Results

6.1 School performance

We estimate the association between ethnic identity and school performance using three measures: latest score in reading, latest score in mathematics and grade retention. Table 6 reports our baseline estimates. Overall, we find a negative correlation between ethnic identity index and immigrants' school performance. A low sense of Italian belonging is associated with a lower grade both in reading and mathematics and this is true for students attending middle (-26% for reading and -35% for math) and high secondary schools (-7% for reading and 24% for math). The magnitude of the negative impact is slightly larger for students attending middle secondary school. Instead, the probability of having repeated a grade increases for not assimilated immigrants in middle secondary school only (+5%). It is likely that at younger age children are less confident, especially if they are not assimilated, and may have more anxiety as a result, which could have an impact on their school performance.²²

When we move our attention to some school related variables, we find that starting Italian school a year later does not affect student grades in reading. Instead, it decreases math scores by 1.4% in middle secondary school, 4.5% in high secondary school and it increases the probability of grade retention for students in middle (+1.8%) and high (+0.7%) secondary schools. As expected, trusting teachers improves school performance: immigrants get better grades both in reading and math and the probability of grade retention decreases by 4% in middle secondary and almost 6% in high secondary schools. When students trust their teachers, they often feel more comfortable asking questions, seeking help, and engaging in classroom discussions. This fosters an environment where learning is more enjoyable and effective. Although not explicitly centred on the concept of trust between students and teachers, Hanushek and Wößmann (2007) touch upon the importance of teacher quality on educational outcomes, which are closely related to the broader concept of trust in teachers affecting student performance. Furthermore, as mentioned in footnote 18, a number of papers have shown the importance of trust in teachers on students' school performance.

Females in both middle and high secondary schools get higher grades in reading and math and have low probability of grade retention compared to males. This is in line with Carlana *et al* (2022). They find that immigrant boys have lower performance in standardized test scores, lower high school graduation rate and more disruptive behaviour compared to their female counterpart. Born in Italy

²² To dispel any doubts that the effect of our ethnic identity index is driven by the linguistic component only, we ran the model without the language component in the ethnic identity index. The results show that the negative association between low sense of Italian belonging and educational outcomes hold for immigrants in middle secondary school. As expected, the magnitude of the effect gets smaller proving the relevance of language skills. For students in high secondary school the negative correlation between the index and school outcomes hold for math only. Results are available upon request.

decreases the probability of grade retention in both middle and high secondary schools but has a positive effect on reading (+10%) and math (+9%) scores for immigrants in middle secondary schools only. Immigrant students born in Italy might outperform their counterparts born abroad due to potential advantages in exposure to the Italian education system and culture from an early age. They might have a better understanding of the educational norms, teaching styles, and expectations among other factors. Playing a sport, which provides a good opportunity for social interaction, improves school performance and decreases the probability of grade retention. On the contrary, working penalizes students who get lower grades in reading and math and have higher probability of grade retention in both middle and high secondary school.

Student's family background is relevant in explaining school outcomes. Surprisingly, when student's mother does not live in the household, those in high secondary school perform better in reading and math and have a lower probability of grade retention. This is possibly because father puts extra effort in the children's education due to absence of mother, who is generally considered to have stronger influence on children's education when both parents are living together (e.g., Flouri and Buchanan, 2004 found that father's involvement significantly predicted educational attainment for adolescents independently of mother involvement). The impact of a father's absence on academic performance seems to be more noticeable and detrimental for students in middle secondary school, whereas students in high secondary school are less affected. Adolescents in middle secondary school (around ages 11-13) might be more sensitive to disruptions in their family structure. Furthermore, we find a negative impact on the school outcome when father does not work for students in high secondary schools. Usually, a father is the head of the household and the principal contributor to the family income. A family can be financially affected when the father does not live in the household or if he does not work. Low school performance is likely to be a consequence of it. Finally, as expected, there is a positive correlation between parental education and school outcomes. The positive effect is even larger for students in middle secondary school who normally need more support by their parents. For example, when one of the two parents hold a degree it increases the grade in reading by 48%, the grade in math by 42% and it decreases the probability of grade retention by 5.8%.

Tables 7 and 8 show the results by gender. It is clear that ethnic identity does not have significantly differentiated impact across boys and girls in the middle school (Table 7). We find some gender difference in school performance for students in high secondary school. Research in education shows that girls generally perform better than boys in reading (e.g. Hedges and Nowell, 1995; Lynn and Mikk, 2009; Voyer and Voyer, 2014), and it seems that lack of assimilation makes it worse. Indeed, lower assimilation has higher negative impact only on the sample of boys (-11%) in high school reading performance. There is some evidence in the education literature that girls overall perform

worse than boys in science subjects (e.g. Hedges and Nowell, 1995; Voyer and Voyer, 2014), including math, and it seems that the level of assimilation plays an important role. Trust in teacher, playing a sport, born in Italy has similar level of correlation with school performance for both groups, with trust in teacher result more pronounced for boys compared to girls.

6.2 Robustness related to school performance

We perform a set of robustness checks to attenuate the concern related to the identification of the effect of identity on educational outcomes. To exclude the issue of selection of immigrants over individuals' characteristics we restrict the sample to a more homogeneous group of individuals. First, in columns (2)-(4) of Tables 9 and 10 we exclude those who were born in Italy. Next, we try to account for pre-migration economic characteristics. In columns (5)-(7), we exclude immigrant students coming contemporaneously from the wealthier families²³ and at the same time arrived in Italy less than 5 years before the survey. The idea is that wealthier new arrived immigrants may have pre-migration wealth and characteristics that can simultaneously affect their identity formation and their school outcomes. Another robustness check, presented in columns (8)-(10), is to exclude the largest ethnic groups in Italy and see if educational behaviour of minorities is in line with the main results. We exclude Romanians, Albanians, Moroccans and Chinese who together account for up to 40% of the migrant population in Italy.²⁴ Reassuringly, our robustness results are stable in significance and magnitude to our baseline estimates. Finally, columns 11-13 of Tables 9 and 10 show the analysis on the school outcomes for students in their last grade of middle secondary school (grade 8) and high secondary school (grade 13) only. Immigrants in grade 8 have a lower performance in reading (-20%) and mathematics (-29%) as well as a higher probability of grade retention (+6.6%) when they have a low commitment to the Italian culture. Instead, for immigrants who are at the end of their high secondary school (grade 13) the sense of belonging to the Italian culture seems to be negatively associated with the performance in mathematics only (-16.8%).

6.3 Educational aspirations

In this section we present the effect of ethnic identity on two outcome variables. The first is the intentions of middle school pupils about their future plans with the following options: academically-oriented schools, technical schools, vocational education, other (no further education). The second

²³ We use median wealth as threshold.

²⁴ Romanians have historically constituted the largest immigrant group in Italy, followed by Moroccans and Albanians. Chinese immigrants also represent a notable portion of Italy's migrant population, particularly in urban areas where Chinese communities have established themselves in various economic sectors.

outcome variable is the intention of those in high school about whether to pursue tertiary education or not.

Table 11 shows a negative association between ethnic identity and preference for academically-oriented (-8.7%) and technical schools (-5.7%). In other words, those with a weak sense of Italian identity are less likely to aim for high secondary school path that potentially offers better education and employment prospects than vocational schools. On the contrary, a weak sense of Italian identity increases the preference for vocational education. Ethnic identity significantly influences immigrants' educational aspirations, often operating alongside school performance to shape students' attitudes and ambitions, e.g., higher scores in reading and math increase the intention to choose academically-oriented school track and lowers the preference for vocational tracks or the “other” option. Repeating a grade decreases the preference for academically-oriented (-8.5%) and technical tracks (-3.2%) while the intention to enrol to a vocational school increases by 7%. Trust in teacher increases by almost 3% the preference for an academically-oriented track. This is the most challenging track for a student and school performance as well as teacher behaviour and trust influence the attitude toward academic schools.

Females seem more orientated toward academic compared to technical and vocational schools. In fact, for females the preference for academically-oriented school increases by 12.8% while the preference for technical and vocational schools decreases, respectively, by 8% and 2%. Moreover, while playing a sport is positively associated with the preference for academically-oriented and vocational tracks (by 3% and 1.4% respectively), working has the opposite effect (-4.7%). Parental education is also an important determinant of children's educational career. Parents holding high secondary education or a degree increase the children preference for academically-oriented (+6.7% and +10% respectively) and decrease the likelihood to prefer vocational schools. There is also an increasing effect of the level of parental education on preference for a technical track. These figures confirm that parental education has a large influence on how children perceive education (e.g. Solon, 1999; Brown, 2006). Indeed, education is a means for achieving better economic outcomes and parents with a higher level of education are likely to push their children to prefer high secondary school which can facilitate university enrolment. In fact, as expected, we find a negative effect of parental education on “other” preferences (which includes training courses or working). Family wealth can have a profound impact on educational investment. Indeed, wealthier families often have greater financial resources to invest in their children's education. We observe a positive correlation between the wealth index and a preference for academically oriented schools, albeit at a significance level of 10 percent. When considering the variable 'father not working,' we note a decrease of 2.4%

in the intention to pursue academically oriented education. Both these variables represent economic status, indicating its crucial role in shaping children's educational preferences.

The last column of Table 11 presents results for the intention to enrol at the university for students who are attending high secondary schools. At the end of high secondary school, students need to decide whether to pursue further education enrolling to the university or not. The decision of investing in further education rather than entering the labour market immediately is likely to affect economic outcomes. In the case of immigrant students, it is likely that the choice of whether to pursue further education is also determined by the degree of assimilation in the host society which is measured by our ethnic identity index. We find that a weaker sense of Italian identity decreases the preference for enrolling to the university by 10%.

The estimated effects of the other covariates included are in line with the results expected. The intention to enrol at the university increases with the latest grade of high secondary school in reading (+3%) and math (+4.5%). As expected, performing well at school contributes in modelling preferences while grade retention has a negative effect (almost -14%). Trusting school teachers also influence students' choices (+4.5%). Looking at the respondent's characteristics, the preference for university is positively affected by being a female (+19%). Playing a sport allows to meet and interact with peers and this interaction influence and shape the formation of preferences. We find that playing a sport increases preference for university (+3.8%) whereas working, even if not regularly, during school-age decreases the intention to enrol at the university (-6.5%). Living in a family, in which there is no sibling increases the preference for university (+2.7%). Probably because the family financial and time resources are fully orientated for the single child in the household. Finally, as expected, both parental education and family wealth are positively correlated with the intention to pursue further education.

When we run separate estimates for immigrant boys and girls, we find that overall, ethnic identity player a larger role in shaping girls' educational aspirations. The baseline analysis for children in middle secondary shows that a weak sense of Italian identity has a larger negative effect on the likelihood of preferring academically-oriented school for girls (-10%) rather than for boys (-7.4%). Immigrant families often hold specific cultural beliefs and expectations about gender roles and education. If not integrated, these beliefs might be more conservative or traditional, limiting girls' aspirations or opportunities for higher education. Indeed, in some cultures, traditional gender norms may prioritize boys' education over girls'. Lack of integration might mean adherence to these norms, resulting in limited support or encouragement for girls' educational aspirations. Moreover, immigrant

girls facing social isolation due to lack of integration might feel disconnected from mainstream educational opportunities.

In Table 12 we also report the summary of the relevant estimates by gender in middle secondary school. School experience and performance as well as family characteristics seem to have a higher weight for girls. Indeed, the positive effect of reading and math scores on the preference for an academically-oriented track is slightly larger for girls compared to boys as well as the impact of grade retention even if with the reverse sign. Moreover, for example, the dummy variables for parental maximum level of education affects girls more than boys in the aspiration for an academically-oriented school.

Gender differences in the aspiration to pursue further education persist even when we focus on the sample of immigrants attending high secondary schools (Table 13). A weak Italian identity affects girls more than boys. The baseline estimates show that the intention to enrol at university decreases more for girls (-12%) than for boys (-7.7%) when there is a low commitment to the Italian culture. Overall, poor integration might disproportionately affect girls' aspirations for university education due to cultural norms, limited access to supportive networks, differing social pressures, and restricted access to information and resources. For example, girls might face stronger social pressures or expectations regarding familial responsibilities or early marriage when integration is limited. These pressures can discourage girls from aspiring to university education.

The estimated marginal effects of the other covariates included are reported in Table 13. Performance in math has a large effect for girls while the latest grade in reading seems to orientate boys rather than girls in the intention to pursue further education. Everything else equal, born in Italy has a positive effect on the intention to enrol at the university only for boys (+4%) while for girls having born in Italy is negatively correlated with the preference for university. This seems counterintuitive. Maybe it is an indication that for girls cultural norms are more relevant than the place of birth. Instead, playing a sport, which is a way to socialize with native peers outside school, has an effect only for girls (+6%). For boys it might be easier to interact with other natives while for girls, which might have less social opportunity due to some gender inequality within the household, playing a sport might be the only opportunity. Finally, the positive effect of parental education, in particular when parents hold secondary and tertiary education, persist and it is significantly strong for both genders.

6.4 Robustness related to educational aspirations

As for the school performance, we run a number for robustness checks for aspirations as well. Tables 14 and 15 present the results. Again, to exclude the issue of selection of immigrants over individuals'

characteristics we exclude (i) those who were born in Italy; (ii) those coming contemporaneously from the wealthier households and arrived in Italy less than 5 years before the survey; (iii) those who are part the largest ethnic groups in Italy: Romanians, Albanians, Moroccans and Chinese. Our robustness results are stable in significance and magnitude to our baseline estimates. Finally, the last three Columns of Tables 14 and 15 (11)-(13) show the analysis on the school aspiration for students in their last grade of middle secondary school (grade 8) and high secondary school (grade 13) only. In their final year of middle or high secondary school, students are more inclined to translate their aspirations into definitive choices for further education. Indeed, students in their final year of secondary school, when confronted with entrance tests for highly competitive university courses, often devote substantial time to preparing for these exams. Immigrants in grade 8 with a lower Italian identity have a lower intention (-5%) to enrol in an academically-oriented and technical high school tracks (-6%) when they have a low commitment to the Italian culture. Also for immigrants who are at the end of their secondary school (grade 13) the sense of belonging to the Italian culture seems to affect negatively the intention to enrol at the university. In particular, a weak Italian identity decreases the likelihood of pursuing a university degree by almost 13%.

7. Conclusion

This paper analysed the role of ethnic identity on immigrant children's educational outcomes. Our ethnic identity measure was based on the one-dimensional ethnosizer as in Constant *et al* (2009). The five identity measures were language, culture, ethnic self-identification, migration history and social networks. The five measures consist of a total of 11 components. Given that the components are likely to be strongly correlated, we used the Principal Component Analysis to construct an ethnic identity index. The results show a strong negative association between our ethnic identity index and immigrant students' performance in reading and math and their aspirations for further education. More precisely, weak sense of belonging to the Italian culture results in lower grades in reading and math and a higher probability of grade retention, particularly in middle secondary school. While school performance is undoubtedly crucial in shaping educational aspirations, ethnic identity can intersect with and sometimes override academic achievement, influencing how immigrant students perceive their educational trajectories and future possibilities. Poor assimilation into the Italian culture is negatively correlated with aspiration for an academically oriented track and for a technical school for immigrant students in middle secondary school. The result is more pronounced for girls compared to boys. In general, academically-oriented schools and technical education offer better future prospects, skills, and career trajectories compared to vocational schools. We also found that immigrant students with a weaker connection to Italian culture are less inclined to pursue university education. This could

limit their long-term economic opportunities and ability to access higher-paying or skill-intensive occupations, increasing the gap between natives and immigrants.

The large literature on ethnic identity often emphasizes its impact on economic outcomes such as labour market participation, wages, and employment opportunities. However, our paper uniquely shifts this focus to highlight the critical role of ethnic identity in shaping immigrants' educational outcomes. This emphasis is crucial because educational outcomes among immigrant populations serve as a fundamental determinant of future labour force participation and economic contribution. By exploring the link between ethnic identity and educational trajectories, our study sheds light on the potential long-term implications for the future workforce.

References

- Adams C. M. & Forsyth P. B. (2013). Revisiting the collective trust effect in urban elementary schools. *The Elementary School Journal*, 113(4):1-36.
- Akerlof, G. A. & Kranton, R.E. (2002). "Identity and schooling: Some lessons for the economics of education." *Journal of Economic Literature*, 40 (4): 1167-1201.
- Akerlof, G. A. & Kranton, R. E. (2000). Economics and identity. *Quarterly Journal of Economics*, 115(3): 715-753.
- Ammermüller, A. (2005). Poor background or low returns? Why immigrant students in Germany perform so poorly in the programme for international student assessment. *Education Economics*, 15(2): 215-230.
- Battu, H., McDonald, M. & Zenou, Y. (2007). Oppositional identities and the labor market. *Journal of Population Economics*, 20 (3): 643-667.
- Battu, H. & Zenou, Y. (2010). Oppositional identities and employment for ethnic minorities: Evidence from England. *Economic Journal* 120 (542).
- Berry, J. W. (1980). Acculturation as varieties of adaptation. In *Acculturation: Theory, Models and Some New Findings*. Ed. A. M. Padilla. Boulder, CO: Westview Press: 9-25.
- Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology*, 46(1): 5-34.
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8: 436-455.
- Brown, P. H. (2006). Parental education and investment in children's human capital in rural China. *Economic Development and Cultural Change*, 54(4): 759-789.
- Cai, S. & Zimmermann, K. F. (2024). Social assimilation and labor market outcomes of migrants in China. *European Economic Review*. Article in Press. <https://doi.org/10.1016/j.euroecorev.2024.104676>.
- Carlana, M., La Ferrara, E. & Pinotti, P. (2022). Goals and gaps: educational careers of immigrant children. *Econometrica*, 90(1): 1-29.
- Carillo, M.R., Lombardo, V. & Venittelli, T. (2023). Social identity and labour market outcomes of immigrants. *Journal of Population Economics*, 36: 69-113.
- Casey, T. & Dustmann, C. (2010). Immigrants' identity, economic outcomes and the transmission of identity across generations. *Economic Journal*, 120 (542), F31-F51.
- Constant, A. F., & Zimmermann, K. F. (2008). Measuring ethnic identity and its impact on economic behavior. *Journal of the European Economic Association*, 6(2-3), 424-433.
- Constant, A., Gataullina, F. L. & Zimmermann K. F. (2009). Ethnosizing immigrants. *Journal of Economic Behavior & Organization*, 69 (3): 274-287.
- Dustmann, C., Frattini, T., & Lanzara, G. (2012). Educational achievement of second-generation immigrants: An international comparison. *Economic Policy*, 27(69), 143-185.
- Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data or tears: An application to educational enrolments in states of India. *Demography*, 38(1): 115-132.
- Forsyth P. B., Adams C. M., & Hoy W. K. (2011). *Collective trust: Why schools can't improve without it*. Teachers College Press.

- Flouri E., & Buchanan A. (2004). Early father's and mother's involvement and child's later educational outcomes. *British Journal of Educational Psychology*, 74, 141-153.
- Frick, J. R., & Wagner, G. G. (2001). Economic and social perspectives of immigrant children in Germany. E. Currie and T. Wunderlich (eds.), *Deutschland – ein Einwanderungsland? Rückblick, Bilanz und neue Fragen - Festschrift für Friedrich Heckmann*, Stuttgart.
- Gang, I. N., & Zimmermann, K. F. (2000). Is child like parent? Educational attainment and ethnic origin. *Journal of Human Resources*, 550-569.
- Hanushek, E. A., & Wößmann, L. (2007). The role of education quality for economic growth. World Bank Policy Research Working Paper No. 4122.
- Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, 269: 41-45.
- Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. Technical report, National Bureau of Economic Research.
- Islam, A., & Raschky, P. A. (2015). Genetic distance, immigrants' identity, and labor market outcomes. *Journal of Population Economics*, 28:845-868.
- Istat (2020). *Identità e percorsi di integrazione delle seconde generazioni in Italia*. ISBN 978-88-458-2010-6. <https://www.istat.it/it/archivio/240930>.
- Jann, B. (2008). A Stata implementation of the Blinder-Oaxaca decomposition. *Stata Journal*, 8(4): 453-479.
- Kalfa, E., & Piracha, M. (2018). Social networks and the labour market mismatch. *Journal of Population Economics*, 31(3): 877-914.
- Lynn, R., & Mikk, J. (2009). Sex differences in reading achievement. *Frames*, 13: 3-13.
- Meschi, E., & Pavese, C. (2023). Ability composition in the class and the school performance of immigrant students. *Labour Economics*, 85. <https://doi.org/10.1016/j.labeco.2023.102450>
- Nekby, L., & Rodin, M. (2010). Acculturation identity and employment among second and middle generation immigrants. *Journal of Economic Psychology*, 31 (1): 35-50.
- Nekby L, Rödin M., & Özcan G. (2009) Acculturation identity and higher education: is there a trade-off between ethnic identity and education? *International Migration Review*, 43(4):938–973.
- Oaxaca, R. (1973). Male-Female wage differentials in urban labor markets. *International Economic Review*, 14: 693–709.
- Patacchini, E., & Zenou, Y. (2016). Racial identity and education in social networks. *Social Networks*, 44: 85-94.
- Pendakur, K., & Pendakur, R. (2005). Ethnic identity and the labour market. Mimeo, Vancouver Centre of Excellence.
- Phinney, J. S., Horenczyk, G., Liebkind, K., & Vedder, P. (2001). Ethnic identity, immigration, and well-being: An interactional perspective. *Journal of Social Issues*, 57(3): 493-510.
- Piracha, M., Tani, M., Cheng, Z., & Wang, B. Z. (2023). Social assimilation and immigrants' labour market outcomes. *Journal of Population Economics*: 36: 37-67.
- Schuller, S. (2015). Parental ethnic identity and educational attainment of second-generation immigrants. *Journal of Population Economics*, 28: 965-1004.

- Solon, G. (1999). Intergenerational mobility in the labor market. In *Handbook of Labor Economics* (Vol. 3, pp. 1761-1800). Elsevier.
- Van Maele, D., & Van Houtte, M. (2011). The quality of school life: Teacher-student trust relationships and organizational school context. *Social Indicators Research*, 100: 85-100.
- Virta, E., Westin, C. (1999). Psychosocial adjustment of adolescents with immigrant background in Sweden. CEIFO.
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: a meta-analysis. *Psychological Bulletin*, 140(4): 1174.
- Zimmermann, K. F. (2007). Migrant ethnic identity: concept and policy implications. *Ekonomia*, 10(1): 1-17.
- Zimmermann L., Gataullina L., Constant A.F., & Zimmermann K.F. (2008). Human capital and ethnic self-identification of immigrants. *Economics Letters*, 98(3):235–239.

Figure 1 – Educational performance

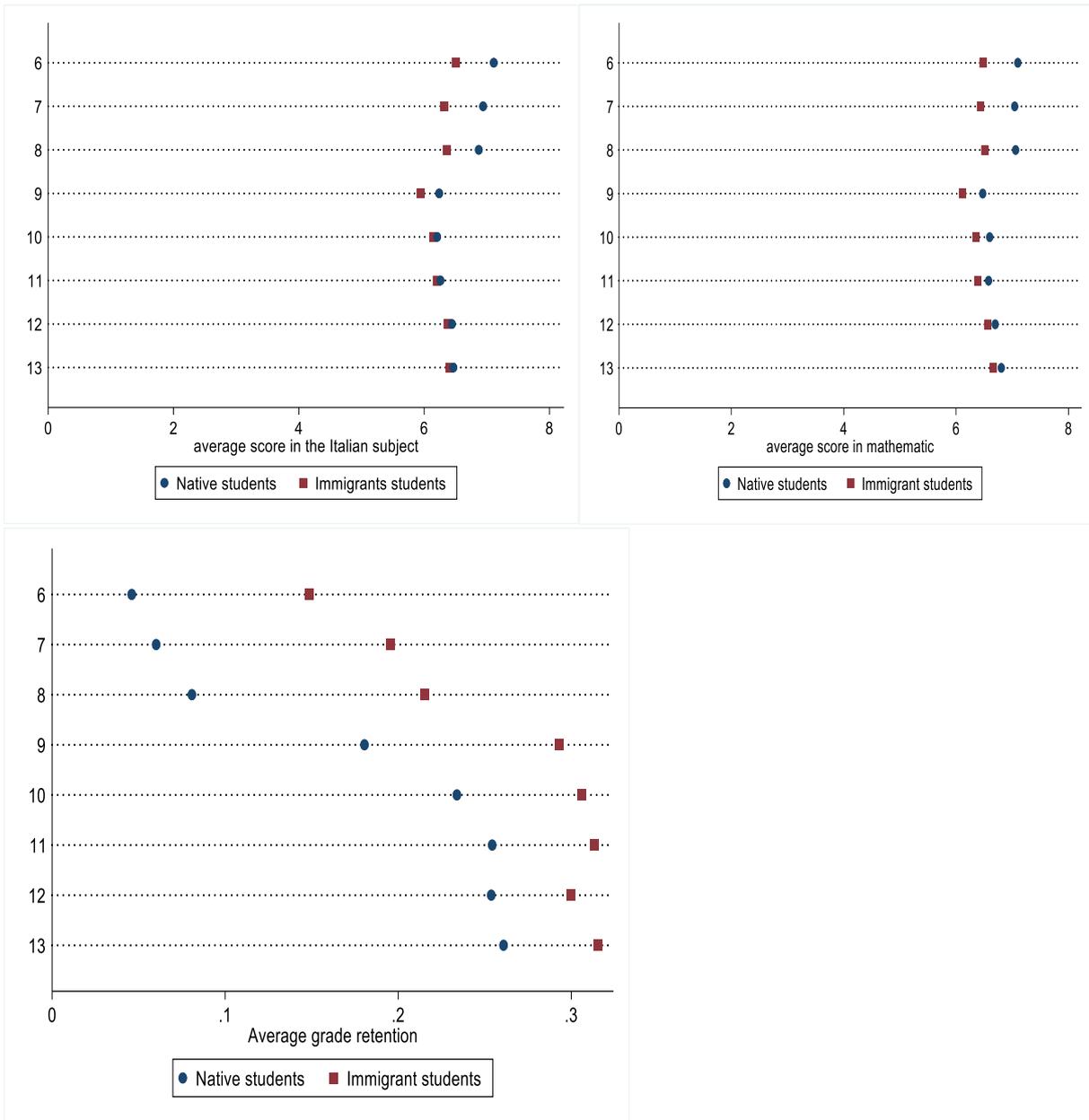


Figure 2 – Educational aspirations of Middle secondary students

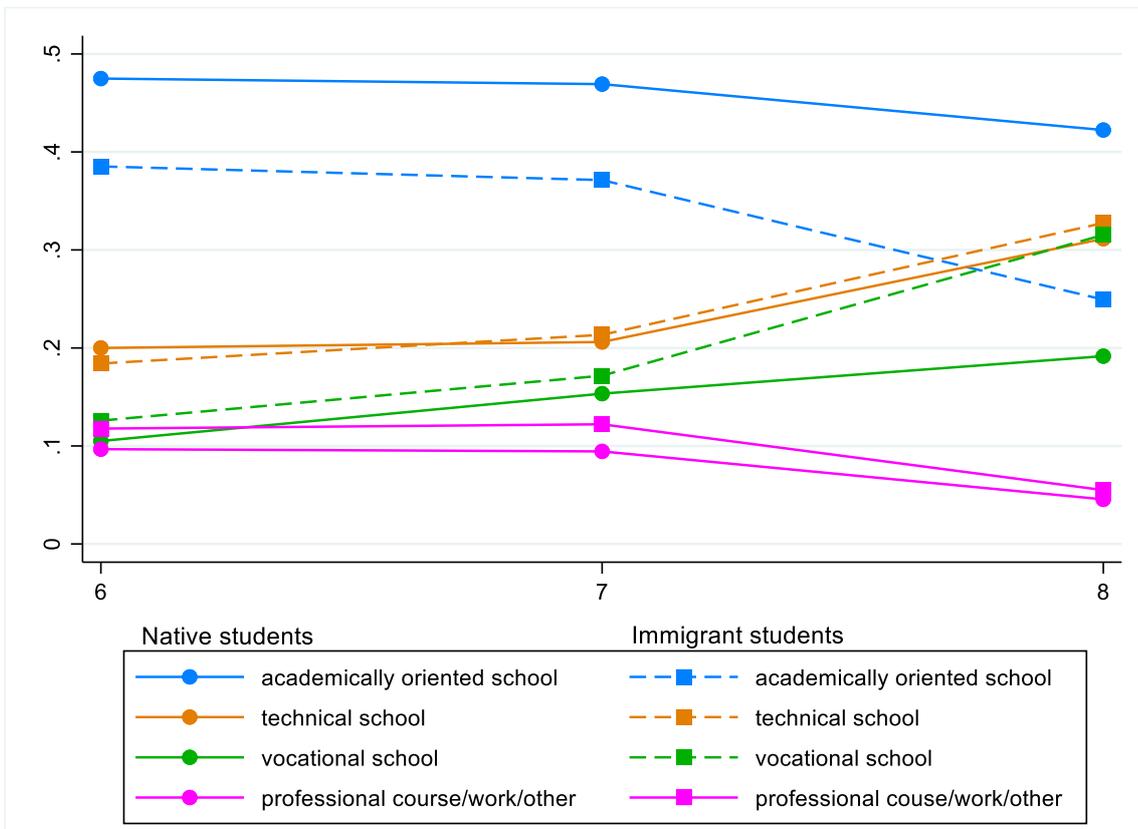


Figure 3 – Aspiration for university for High secondary students

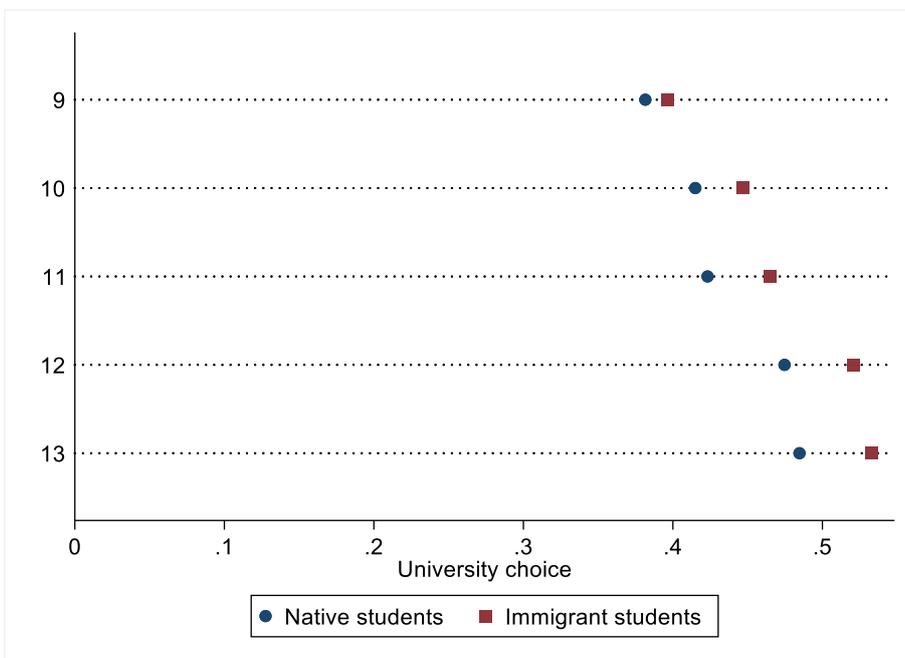


Table 1 - Descriptive differences between native and immigrant students

	Students in middle school			Students in high school		
	Native	Immigrant	diff	Native	Immigrant	diff
Reading score	6.977	6.398	0.579***	6.297	6.159	0.138***
Math score	7.068	6.474	0.594***	6.595	6.357	0.238***
Grade retention	0.062	0.187	-0.125***	0.227	0.304	-0.077***
Academically oriented school	0.457	0.337	0.120***			
Technical school	0.239	0.245	-0.006			
Vocational school	0.151	0.205	-0.055***			
Other	0.079	0.097	-0.018***			
University				0.423	0.456	-0.033***
Grade Italian school started	1.096	2.403	-1.307***	1.144	3.840	-2.696***
Trust teachers	0.785	0.790	-0.004	0.536	0.594	-0.058***
Gender (Female)	0.492	0.468	0.024***	0.485	0.544	-0.060***
Born in Italy	0.938	0.410	0.528***	0.940	0.193	0.747***
Play a sport	0.743	0.538	0.205***	0.641	0.454	0.187***
Work	0.283	0.307	-0.024***	0.300	0.321	-0.021***
Have no sibling	0.207	0.202	0.005	0.249	0.258	-0.009
Mother not in the household	0.018	0.032	-0.013***	0.031	0.051	-0.020***
Father not in the household	0.128	0.200	-0.071***	0.162	0.269	-0.107***
Mother does not work	0.281	0.361	-0.080***	0.298	0.346	-0.048***
Father does not work	0.078	0.152	-0.074***	0.101	0.181	-0.080***
Parental edu (Primary)	0.011	0.028	-0.017***	0.009	0.018	-0.009***
Parental edu (Middle secondary)	0.150	0.156	-0.006	0.204	0.161	0.043***
Parental edu (High secondary)	0.326	0.253	0.073***	0.501	0.446	0.055***
Parental edu (Degree)	0.264	0.145	0.119***	0.222	0.214	0.008
Parental edu (Not reported)	0.211	0.336	-0.125***	0.051	0.116	-0.064***
Family's wealth index	3.371	3.186	0.185***	3.389	3.218	0.171***
Living in a big municipality	0.190	0.204	-0.014**	0.304	0.313	-0.010
N	18254	12781	31035	18923	14757	33680

Table 2 – Oaxaca-Blinder Decomposition: gap in educational outcomes

	Students in middle secondary school			Students in high secondary school		
	Reading	Math	Grade retention	Reading	Math	Grade retention
Natives	6.977*** (698.76)	7.068*** (828.98)	0.0620*** (34.74)	6.297*** (604.88)	6.595*** (848.01)	0.227*** (74.52)
Immigrants	6.398*** (548.38)	6.474*** (662.47)	0.187*** (54.16)	6.159*** (503.46)	6.357*** (704.67)	0.304*** (80.19)
difference	0.579*** (37.74)	0.594*** (45.78)	-0.125*** (-32.12)	0.138*** (8.58)	0.238*** (20.01)	-0.0767*** (-15.78)
explained	0.296*** (21.85)	0.287*** (24.94)	-0.0748*** (-20.63)	0.0585** (2.85)	0.0621*** (4.05)	-0.0801*** (-13.25)
unexplained	0.283*** (14.47)	0.306*** (19.35)	-0.0499*** (-10.02)	0.0793** (3.08)	0.176*** (9.35)	0.00340 (0.45)

Table 3 – Oaxaca-Blinder Decomposition: gap in educational aspirations

	Students in middle secondary school				Students in high secondary school
	Academically oriented	Technical school	Vocational school	Other	University
Natives	0.457*** (123.83)	0.239*** (75.67)	0.151*** (56.91)	0.0788*** (39.51)	0.423*** (117.88)
Immigrants	0.337*** (80.54)	0.245*** (64.34)	0.205*** (57.46)	0.0972*** (37.09)	0.456*** (111.25)
difference	0.120*** (21.51)	-0.00586 (-1.19)	-0.0546*** (-12.28)	-0.0184*** (-5.59)	-0.0327*** (-6.00)
explained	0.155*** (31.74)	-0.00384 (-0.93)	-0.0724*** (-18.79)	-0.0196*** (-6.92)	0.0563*** (8.48)
unexplained	-0.0346*** (-5.07)	-0.00202 (-0.32)	0.0178** (3.17)	0.00124 (0.29)	-0.0890*** (-11.15)

Table 4 – Descriptive statistics for immigrant students only

	Middle secondary school		High secondary school	
	mean	sd	mean	sd
Reading score	6.401	1.318	6.161	1.483
Math score	6.476	1.101	6.357	1.093
Grade retention	0.186	0.389	0.302	0.459
Academically oriented school	0.338	0.473		
Technical school	0.245	0.430		
Vocational school	0.206	0.404		
Other	0.096	0.295		
University			0.457	0.498
Attend grade 6	0.152	0.359		
Attend grade 7	0.331	0.471		
Attend grade 8	0.340	0.474		
Attend grade 9			0.317	0.465
Attend grade 10			0.216	0.411
Attend grade 11			0.191	0.393
Attend grade 12			0.153	0.360
Attend grade 13			0.124	0.329
Grade Italian school started	2.400	2.184	3.838	3.213
Trust teachers	0.791	0.407	0.595	0.491
Gender (Female)	0.469	0.499	0.545	0.498
Born in Italy	0.410	0.492	0.193	0.395
Play a sport	0.537	0.499	0.454	0.498
Work	0.306	0.461	0.320	0.467
Have no sibling	0.201	0.401	0.258	0.438
Mother not in the household	0.031	0.172	0.050	0.219
Father not in the household	0.200	0.400	0.270	0.444
Mother does not work	0.360	0.480	0.345	0.475
Father does not work	0.152	0.359	0.181	0.385
Parental edu (Primary)	0.027	0.163	0.018	0.133
Parental edu (Middle secondary)	0.157	0.364	0.161	0.368
Parental edu (High secondary)	0.253	0.435	0.447	0.497
Parental edu (Degree)	0.146	0.353	0.215	0.411
Parental edu (Not reported)	0.337	0.473	0.115	0.320
Family's wealth index	3.188	0.310	3.219	0.292
Living in a big city	0.204	0.403	0.314	0.464
Region: Piemonte	0.714	0.257	0.078	0.268
Region: Valle D'Aosta	0.014	0.118	0.009	0.092
Region: Lombardia	0.076	0.265	0.091	0.287
Region: Veneto	0.088	0.283	0.102	0.303
Region: Friuli-Venezia Giulia	0.051	0.219	0.055	0.228
Region: Liguria	0.051	0.221	0.068	0.252
Region: Emilia Romagna	0.081	0.272	0.121	0.326
Region: Toscana	0.071	0.258	0.067	0.250
Region: Umbria	0.049	0.215	0.042	0.201
Region: Marche	0.047	0.212	0.038	0.192
Region: Lazio	0.055	0.229	0.061	0.240
Region: Abruzzo	0.048	0.213	0.037	0.189
Region: Molise	0.009	0.093	0.008	0.090
Region: Campania	0.064	0.246	0.038	0.190
Region: Puglia	0.034	0.182	0.038	0.192
Region: Basilicata	0.013	0.111	0.008	0.090
Region: Calabria	0.036	0.187	0.036	0.187
Region: Sicilia	0.051	0.220	0.041	0.199

Region: Sardegna	0.015	0.121	0.014	0.117
Region: Autonomous Province of Bolzano	0.033	0.179	0.017	0.128
Region: Autonomous Province of Trento	0.042	0.201	0.031	0.172
Mother's origin country: Albania	0.161	0.367	0.163	0.369
Mother's origin country: Bangladesh	0.013	0.115	0.013	0.114
Mother's origin country: China	0.067	0.250	0.044	0.205
Mother's origin country: Ecuador	0.023	0.151	0.034	0.180
Mother's origin country: Egypt	0.009	0.092	0.014	0.119
Mother's origin country: Philippine	0.037	0.188	0.035	0.185
Mother's origin country: India	0.021	0.145	0.016	0.127
Mother's origin country: Kosovo	0.018	0.132	0.014	0.117
Mother's origin country: Macedonia	0.032	0.177	0.019	0.137
Mother's origin country: Marocco	0.094	0.291	0.081	0.273
Mother's origin country: Pakistan	0.019	0.136	0.016	0.126
Mother's origin country: Peru	0.021	0.144	0.029	0.167
Mother's origin country: Polonia	0.016	0.124	0.021	0.145
Mother's origin country: Romania	0.220	0.414	0.218	0.413
Mother's origin country: Sri Lanka	0.013	0.114	0.013	0.113
Mother's origin country: Tunisia	0.020	0.140	0.019	0.135
Mother's origin country: Ukraine	0.023	0.149	0.039	0.194
N	12606		14647	

Notes: We present here only the main countries of origin but we control for all countries of origin reported in the data. Full list of reported mother's origin countries is presented in the Appendix, Table A1. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms.

Table 5 – Ethnic identity index components

	Middle secondary		High secondary	
	mean	sd	mean	sd
Language				
Understand: Italian understanding (yes=0)	0.020	0.139	0.011	0.105
Speak: Italian speaking (yes=0)	0.026	0.158	0.014	0.119
Read: Italian reading (yes=0)	0.032	0.177	0.016	0.126
Write: Italian writing (yes=0)	0.038	0.191	0.024	0.153
Culture				
Think: Thinking in Italian (yes=0)	0.370	0.483	0.353	0.478
Speak at home: Speaking Italian at home (yes=0)	0.502	0.500	0.512	0.500
Watch TV: watching Italian TV (yes=0)	0.116	0.320	0.130	0.336
Ethnic self-identification				
Identifying as Italian (yes=0)	0.591	0.492	0.658	0.474
Migration history				
Wishing to live in Italy (yes=0)	0.649	0.477	0.722	0.448
Social Network				
Peers: Meeting Italian schoolmates outside school (yes=0)	0.311	0.463	0.325	0.469
Friends: Having Italian friends (yes=0)	0.245	0.430	0.220	0.414
Ethnic identity index (PCA)	0.394	0.388	0.379	0.339
N	12606		14647	
<i>Notes: (1) Language: for each language component, a value of 0 is assigned for those reporting a very good/ good or fair command of Italian. A value of 1 is assigned if ability in Italian is poor or very poor. (2) Culture: for the three visible cultural aspects a value of 0 is assigned if respondent are close to Italian culture and 1 otherwise. (3) Ethnic self-identification: feeling Italian gets a value of 0 (it indicates full assimilation) and 0 otherwise. (4) Migration history: it refers to the question “Where do you want to live when you get older?”. The answer ‘want to live in Italy’ takes the value of 0; any possible alternative takes the value of 1. (5) Social network: We consider the nationality of friends (i) within school and (ii) outside school. A value of 0 is assigned if friends are exclusively Italian; a value of 1 is given when friends come from mixed origin. The ethnic identity index is constructed using the Principal Component Analysis.</i>				

Table 6 - Baseline Results: School performance

	Middle secondary school			High secondary school		
	Reading	Math	Grade retention	Reading	Math	Grade retention
Ethnic identity index	-0.265*** (0.034)	-0.354*** (0.029)	0.054*** (0.011)	-0.070* (0.043)	-0.247*** (0.032)	0.017 (0.013)
Attend grade 7	-0.200*** (0.028)	-0.049** (0.023)	0.034*** (0.008)			
Attend grade 8	-0.188*** (0.029)	-0.001 (0.024)	0.040*** (0.009)			
Attend grade 10				0.210*** (0.034)	0.201*** (0.024)	0.013 (0.010)
Attend grade 11				0.244*** (0.036)	0.236*** (0.025)	0.018 (0.011)
Attend grade 12				0.444*** (0.037)	0.414*** (0.027)	-0.006 (0.012)
Attend grade 13				0.463*** (0.041)	0.536*** (0.029)	-0.002 (0.013)
Grade Italian school started	-0.008 (0.007)	-0.014*** (0.005)	0.018*** (0.002)	-0.000 (0.005)	-0.045*** (0.003)	0.007*** (0.002)
Trust teachers	0.183*** (0.028)	0.201*** (0.023)	-0.042*** (0.009)	0.297*** (0.025)	0.157*** (0.018)	-0.059*** (0.008)
Gender (Female)	0.213*** (0.023)	0.414*** (0.019)	-0.104*** (0.007)	0.217*** (0.026)	0.302*** (0.018)	-0.095*** (0.008)
Born in Italy	0.103*** (0.030)	0.093*** (0.024)	-0.058*** (0.009)	-0.014 (0.038)	-0.067** (0.027)	-0.091*** (0.011)
Play a sport	0.109*** (0.024)	0.085*** (0.020)	-0.003 (0.007)	0.109*** (0.026)	0.068*** (0.018)	-0.019** (0.008)
Work	-0.202*** (0.024)	-0.128*** (0.020)	0.037*** (0.008)	-0.061** (0.026)	-0.043** (0.019)	0.098*** (0.008)
Have no sibling	0.098*** (0.030)	0.053** (0.025)	-0.021** (0.009)	-0.011 (0.030)	0.044** (0.021)	0.003 (0.009)
Mother not in the household	-0.059 (0.069)	-0.117** (0.055)	0.022 (0.024)	0.164*** (0.064)	0.113** (0.045)	-0.035* (0.020)
Father not in the household	-0.088*** (0.033)	-0.077*** (0.027)	0.035*** (0.010)	-0.079** (0.033)	0.036 (0.023)	0.006 (0.010)
Mother does not work	0.035 (0.027)	0.039* (0.022)	0.008 (0.008)	0.058** (0.029)	0.022 (0.021)	-0.006 (0.009)
Father does not work	-0.075** (0.033)	-0.031 (0.027)	-0.001 (0.010)	-0.094*** (0.033)	-0.070*** (0.024)	0.034*** (0.011)
Parental edu (Primary)	0.191** (0.078)	0.129** (0.061)	0.023 (0.027)	0.112 (0.108)	0.048 (0.079)	0.024 (0.035)
Parental edu (Mid. Secondary)	0.162*** (0.047)	0.129*** (0.038)	0.024 (0.016)	0.048 (0.068)	0.153*** (0.050)	-0.009 (0.022)
Parental edu (High secondary)	0.411*** (0.046)	0.372*** (0.037)	-0.035** (0.015)	0.136** (0.065)	0.265*** (0.048)	-0.039* (0.021)
Parental edu (Degree)	0.481*** (0.050)	0.423*** (0.041)	-0.058*** (0.015)	0.233*** (0.068)	0.347*** (0.050)	-0.040* (0.022)
Parental edu (Not reported)	0.106** (0.043)	0.089** (0.035)	-0.028* (0.014)	-0.065 (0.071)	0.053 (0.051)	-0.011 (0.023)
Family's wealth index	0.118*** (0.040)	0.114*** (0.034)	-0.003 (0.012)	-0.023 (0.047)	0.089** (0.035)	0.005 (0.014)
Living in a big city	0.136*** (0.033)	0.066** (0.028)	-0.052*** (0.010)	-0.085*** (0.032)	-0.083*** (0.023)	-0.013 (0.010)
Italian region	yes	yes	yes	yes	yes	yes
Mother's origin country	yes	yes	yes	yes	yes	yes
Constant	5.814*** (0.149)	5.745*** (0.126)	0.224*** (0.046)	5.752*** (0.178)	5.743*** (0.130)	0.364*** (0.054)
N	12606	12606	12606	14647	14647	14647
R-squared	0.118	0.159	0.111	0.088	0.144	0.083

Notes: Standard errors in parentheses. * p<0.10 ** p<0.05 ***p<0.01. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the index: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms. The Italian regions are: Piemonte, Valle D'Aosta, Lombardia, Veneto, Friuli-Venezia Giulia, Liguria, Emilia Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna, Autonomous Province of Bolzano, Autonomous Province of Trento. The list of mother's origin country is presented in the Appendix, Table A1.

Table 7 – School performance in middle secondary school by gender

	Males			Females		
	Reading	Math score	Grade retention	Reading	Math score	Grade retention
Ethnic identity index	-0.265*** (0.046)	-0.340*** (0.040)	0.056*** (0.016)	-0.256*** (0.053)	-0.357*** (0.044)	0.055*** (0.015)
Attend grade 7	-0.228*** (0.038)	-0.064** (0.032)	0.043*** (0.012)	-0.165*** (0.040)	-0.026 (0.033)	0.026** (0.010)
Attend grade 8	-0.242*** (0.040)	-0.050 (0.032)	0.060*** (0.013)	-0.132*** (0.043)	0.063* (0.035)	0.019* (0.011)
Grade Italian school started	-0.007 (0.009)	-0.003 (0.008)	0.021*** (0.003)	-0.009 (0.010)	-0.028*** (0.008)	0.016*** (0.003)
Trust teachers	0.165*** (0.039)	0.195*** (0.031)	-0.051*** (0.013)	0.209*** (0.042)	0.213*** (0.034)	-0.027** (0.011)
Born in Italy	0.091** (0.041)	0.105*** (0.033)	-0.075*** (0.013)	0.104** (0.044)	0.070** (0.034)	-0.035*** (0.011)
Play a sport	0.141*** (0.034)	0.081*** (0.028)	-0.006 (0.011)	0.085** (0.035)	0.092*** (0.029)	-0.005 (0.009)
Work	-0.165*** (0.033)	-0.098*** (0.027)	0.040*** (0.011)	-0.243*** (0.037)	-0.153*** (0.030)	0.034*** (0.010)
Have no sibling	0.058 (0.041)	0.060* (0.034)	-0.043*** (0.013)	0.159*** (0.045)	0.060* (0.036)	0.000 (0.011)
Mother not in the household	-0.114 (0.091)	-0.135* (0.075)	-0.015 (0.031)	0.010 (0.106)	-0.111 (0.079)	0.077** (0.036)
Father not in the household	-0.068 (0.047)	-0.079** (0.039)	0.043*** (0.016)	-0.099** (0.046)	-0.073** (0.037)	0.030** (0.013)
Mother does not work	0.023 (0.039)	-0.006 (0.032)	0.004 (0.012)	0.048 (0.039)	0.093*** (0.031)	0.011 (0.010)
Father does not work	-0.068 (0.048)	-0.038 (0.039)	0.002 (0.016)	-0.077 (0.047)	-0.015 (0.038)	-0.006 (0.013)
Parental edu (Primary)	0.144 (0.100)	-0.001 (0.078)	0.007 (0.038)	0.256** (0.123)	0.290*** (0.097)	0.045 (0.037)
Parental edu (Mid. secondary)	0.169*** (0.064)	0.110** (0.051)	0.031 (0.023)	0.155** (0.071)	0.160*** (0.057)	0.021 (0.022)
Parental edu (High secondary)	0.426*** (0.062)	0.344*** (0.049)	-0.048** (0.022)	0.393*** (0.069)	0.410*** (0.056)	-0.019 (0.020)
Parental edu (Degree)	0.475*** (0.069)	0.369*** (0.057)	-0.072*** (0.022)	0.499*** (0.076)	0.495*** (0.062)	-0.037* (0.021)
Parental edu (Not reported)	0.115** (0.058)	0.090* (0.047)	-0.036* (0.020)	0.101 (0.065)	0.097* (0.053)	-0.019 (0.019)
Family's wealth index	0.064 (0.053)	0.026 (0.048)	0.032* (0.018)	0.178*** (0.060)	0.221*** (0.049)	-0.044*** (0.017)
Living in a big city	0.131*** (0.047)	0.093** (0.039)	-0.074*** (0.015)	0.140*** (0.049)	0.029 (0.040)	-0.029** (0.012)
Italian region	yes	yes	yes	yes	yes	yes
Mother's origin	yes	yes	yes	yes	yes	yes
Constant	6.006*** (0.198)	6.066*** (0.175)	0.134** (0.066)	5.802*** (0.225)	5.757*** (0.182)	0.220*** (0.063)
N	6689	6689	6689	5917	5917	5917
R-squared	0.121	0.122	0.104	0.125	0.164	0.107

Notes: Standard errors in parentheses. * p<0.10 ** p<0.05 ***p<0.01. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the index: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms. The Italian regions are: Piemonte, Valle D'Aosta, Lombardia, Veneto, Friuli-Venezia Giulia, Liguria, Emilia Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna, Autonomous Province of Bolzano, Autonomous Province of Trento. The list of mother's is presented in the Appendix, Table A1.

Table 8 - School performance in High secondary school by gender

	Males			Females		
	Reading	Math	Grade retention	Reading	Math	Grade retention
Ethnic identity index	-0.116* (0.065)	-0.211*** (0.049)	0.008 (0.019)	-0.026 (0.057)	-0.259*** (0.042)	0.021 (0.017)
Attend grade 10	0.223*** (0.051)	0.200*** (0.035)	0.039** (0.016)	0.196*** (0.046)	0.206*** (0.033)	-0.011 (0.013)
Attend grade 11	0.190*** (0.054)	0.212*** (0.038)	0.032* (0.017)	0.292*** (0.049)	0.268*** (0.034)	0.006 (0.014)
Attend grade 12	0.447*** (0.056)	0.333*** (0.041)	0.018 (0.019)	0.444*** (0.050)	0.477*** (0.037)	-0.024 (0.015)
Attend grade 13	0.442*** (0.063)	0.521*** (0.043)	0.000 (0.020)	0.480*** (0.054)	0.547*** (0.039)	-0.006 (0.017)
Grade Italian school started	0.000 (0.007)	-0.038*** (0.005)	0.006** (0.002)	-0.002 (0.006)	-0.053*** (0.005)	0.008*** (0.002)
Trust teachers	0.316*** (0.039)	0.141** (0.027)	-0.078*** (0.012)	0.285*** (0.033)	0.168*** (0.024)	-0.043*** (0.010)
Born in Italy	0.008 (0.058)	-0.089** (0.040)	-0.102*** (0.017)	-0.025 (0.050)	-0.039 (0.036)	-0.086*** (0.014)
Play a sport	0.110*** (0.038)	0.056** (0.027)	-0.008 (0.012)	0.087** (0.036)	0.061** (0.026)	-0.029*** (0.011)
Work	-0.016 (0.038)	-0.018 (0.027)	0.097*** (0.012)	-0.112*** (0.037)	-0.052** (0.026)	0.098*** (0.012)
Have no sibling	0.049 (0.045)	0.094*** (0.032)	0.006 (0.014)	-0.059 (0.040)	-0.002 (0.029)	0.000 (0.012)
Mother not in the household	0.153 (0.094)	0.120* (0.062)	-0.061** (0.028)	0.151* (0.089)	0.040 (0.068)	0.017 (0.028)
Father not in the household	-0.099* (0.053)	0.054 (0.037)	-0.008 (0.016)	-0.062 (0.044)	0.025 (0.031)	0.021 (0.013)
Mother does not work	0.085** (0.043)	0.002 (0.031)	-0.023 (0.014)	0.035 (0.039)	0.040 (0.028)	0.008 (0.012)
Father does not work	-0.048 (0.052)	-0.085** (0.038)	0.039** (0.016)	-0.138*** (0.044)	-0.061* (0.031)	0.031** (0.014)
Parental edu (Primary)	0.003 (0.160)	0.069 (0.099)	0.038 (0.055)	0.247* (0.148)	0.051 (0.115)	0.019 (0.046)
Parental edu (Mid. secondary)	-0.049 (0.096)	0.177*** (0.069)	-0.011 (0.030)	0.163 (0.099)	0.151** (0.073)	-0.000 (0.032)
Parental edu (High secondary)	0.044 (0.091)	0.259*** (0.065)	-0.060** (0.029)	0.250*** (0.097)	0.293*** (0.070)	-0.015 (0.031)
Parental edu (Degree)	0.155 (0.095)	0.328*** (0.068)	-0.039 (0.030)	0.341*** (0.101)	0.384*** (0.073)	-0.036 (0.032)
Parental edu (Not reported)	-0.097 (0.097)	0.070 (0.068)	-0.018 (0.031)	-0.006 (0.104)	0.053 (0.077)	0.001 (0.034)
Family's wealth index	-0.089 (0.063)	0.038 (0.049)	0.019 (0.019)	0.063 (0.069)	0.173*** (0.049)	-0.015 (0.021)
Living in a big city	-0.112** (0.050)	-0.061* (0.035)	-0.017 (0.016)	-0.063 (0.042)	-0.095*** (0.030)	-0.012 (0.013)
Italian region	yes	yes	yes	yes	yes	yes
Mother's origin	yes	yes	yes	yes	yes	yes
Constant	5.956*** (0.245)	5.813*** (0.183)	0.374*** (0.075)	5.642*** (0.263)	5.822*** (0.185)	0.281*** (0.077)
N	6664	6664	6664	7983	7983	7983
R-squared	0.092	0.119	0.089	0.091	0.157	0.081

Notes: Standard errors in parentheses. * p<0.10 ** p<0.05 ***p<0.01. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the index: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms. The Italian regions are: Piemonte, Valle D'Aosta, Lombardia, Veneto, Friuli-Venezia Giulia, Liguria, Emilia Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna, Autonomous Province of Bolzano, Autonomous Province of Trento. The list of mother's is presented in the Appendix, Table A1.

Table 9 - Robustness: school performance in middle secondary school

	Robustness 1			Robustness 2			Robustness 3			Robustness 4		
	Reading	Math	Grade retention	Reading	Math	Grade retention	Reading	Math	Grade retention	Reading	Math	Grade retention
Ethnic identity index	-0.280*** (0.043)	-0.325*** (0.038)	0.051*** (0.015)	-0.253*** (0.037)	-0.361*** (0.031)	0.055*** (0.012)	-0.315*** (0.052)	-0.325*** (0.044)	0.042** (0.017)	-0.202*** (0.062)	-0.296*** (0.050)	0.066*** (0.020)
N	7434	7434	7434	11605	11605	11605	5788	5788	5788	4289	4289	4289
R-squared	0.121	0.174	0.124	0.117	0.160	0.108	0.127	0.175	0.115	0.136	0.192	0.132

Notes: Standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01; The table presents the main estimation results. We control for grade immigrant is attending, grade Italian school has started, trust teacher, born in Italy, play a sport, work, have no sibling, mother/father not in the household, parental education, living in a big city, Italian region, mother's origin country. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the ethnic identity index using the PCA analysis: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms.

Robustness analysis 1 excludes those immigrant students who where born in Italy

Robustness analysis 2 excludes those immigrant students who come from wealthier households and arrived in Italy in the last 5 year

Robustness analysis 3 excludes the largest ethnic group in Italy: Romanian, Albanian, Marocco, China

Robustness analysis 4 perform the analysis for immigrant students in their last middle secondary year (grade 8)

Table 10 - Robustness: school performance in high secondary school

	Robustness 1			Robustness 2			Robustness 3			Robustness 4		
	Reading	Math	Grade retention	Reading	Math	Grade retention	Reading	Math	Grade retention	Reading	Math	Grade retention
Ethnic identity index	-0.078* (0.047)	-0.232*** (0.034)	0.003 (0.014)	-0.077* (0.045)	-0.264*** (0.034)	0.015 (0.013)	-0.162*** (0.059)	-0.207*** (0.043)	0.029 (0.018)	0.015 (0.114)	-0.168** (0.084)	-0.020 (0.034)
N	11817	11817	11817	13618	13618	13618	7234	7234	7234	1812	1812	1812
R-squared	0.089	0.152	0.088	0.087	0.144	0.086	0.104	0.157	0.088	0.092	0.162	0.117

Notes: Standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01; The table presents the main estimation results. We control for grade immigrant is attending, grade Italian school has started, trust teacher, born in Italy, play a sport, work, have no sibling, mother/father not in the household, parental education, living in a big city, Italian region, mother's origin country. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the ethnic identity index using the PCA analysis: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms.

Robustness analysis 1 excludes those immigrant students who where born in Italy

Robustness analysis 2 excludes those immigrant students who come from wealthier households and arrived in Italy in the last 5 year

Robustness analysis 3 excludes the largest ethnic group in Italy: Romanian, Albanian, Marocco, China

Robustness analysis 4 perform the analysis for immigrant students in their last middle secondary year (grade 8)

Table 11 - Baseline results: Educational aspiration

	Middle secondary school				High secondary
	Academic.	Technical	Vocational	Other	University
Ethnic identity index	-0.087*** (0.011)	-0.057*** (0.011)	0.018* (0.011)	-0.011 (0.007)	-0.104*** (0.013)
Attend grade 7	0.000 (0.010)	0.027*** (0.009)	0.032*** (0.008)	0.000 (0.007)	
Attend grade 8	-0.133*** (0.010)	0.139*** (0.010)	0.175*** (0.009)	-0.062*** (0.007)	
Attend grade 10					0.019* (0.010)
Attend grade 11					0.040*** (0.011)
Attend grade 12					0.070*** (0.012)
Attend grade 13					0.088*** (0.013)
Reading score	0.040*** (0.004)	0.010*** (0.004)	-0.031*** (0.003)	-0.008*** (0.003)	0.031*** (0.003)
Math score	0.057*** (0.005)	-0.015*** (0.005)	-0.016*** (0.004)	-0.008*** (0.003)	0.045*** (0.004)
Grade retention	-0.085*** (0.010)	-0.032*** (0.010)	0.072*** (0.011)	0.009 (0.008)	-0.139*** (0.008)
Grade Italian school started	0.002 (0.002)	-0.003 (0.002)	0.004* (0.002)	-0.000 (0.002)	-0.004*** (0.002)
Trust teachers	0.029*** (0.010)	-0.001 (0.010)	0.015* (0.009)	-0.006 (0.007)	0.045*** (0.008)
Gender (Female)	0.128*** (0.009)	-0.088*** (0.008)	-0.022*** (0.007)	0.008 (0.006)	0.194*** (0.008)
Born in Italy	0.046*** (0.010)	-0.024** (0.010)	-0.019** (0.009)	-0.008 (0.007)	0.006 (0.012)
Play a sport	0.030*** (0.008)	0.007 (0.008)	-0.009 (0.007)	-0.003 (0.006)	0.038*** (0.008)
Work	-0.047*** (0.008)	0.003 (0.008)	0.014* (0.008)	0.002 (0.006)	-0.065*** (0.008)
Have no sibling	0.003 (0.010)	-0.001 (0.010)	-0.015* (0.009)	0.013* (0.007)	0.027*** (0.010)
Mother not in the household	-0.045** (0.021)	-0.019 (0.022)	0.055** (0.024)	-0.009 (0.015)	-0.018 (0.019)
Father not in the household	0.021* (0.012)	-0.009 (0.011)	0.002 (0.010)	0.002 (0.008)	0.010 (0.011)
Mother does not work	-0.012 (0.009)	0.017* (0.009)	-0.013 (0.008)	-0.001 (0.006)	-0.001 (0.009)
Father does not work	-0.024** (0.011)	0.017 (0.011)	-0.006 (0.011)	0.003 (0.008)	-0.033*** (0.011)
Parental edu (Primary)	-0.016 (0.025)	0.064** (0.026)	-0.001 (0.026)	-0.046*** (0.018)	0.049 (0.032)
Parental edu (Mid. secondary)	0.026 (0.016)	0.037** (0.016)	0.019 (0.016)	-0.029** (0.013)	0.049** (0.019)
Parental edu (High secondary)	0.067*** (0.016)	0.055*** (0.015)	-0.023 (0.015)	-0.037*** (0.012)	0.135*** (0.018)
Parental edu (Degree)	0.105*** (0.018)	0.053*** (0.017)	-0.057*** (0.016)	-0.032** (0.013)	0.224*** (0.019)
Parental edu (Not reported)	0.015 (0.015)	0.040*** (0.014)	-0.013 (0.014)	-0.015 (0.012)	0.019 (0.020)
Family's wealth index	0.026* (0.013)	0.043*** (0.012)	0.016 (0.011)	-0.018* (0.009)	0.046*** (0.013)
Living in a big city	0.042*** (0.012)	0.003 (0.012)	-0.038*** (0.010)	-0.011 (0.008)	0.006 (0.010)
Italian region	yes	yes	yes	yes	yes
Mother's origin	yes	yes	yes	yes	yes
Constant	-0.412*** (0.056)	0.078 (0.052)	0.415*** (0.048)	0.317*** (0.038)	-0.375*** (0.054)
N	12606	12606	12606	12606	14647
R-squared	0.169	0.055	0.103	0.032	0.197

Notes: Standard errors in parentheses. * p<0.10 ** p<0.05 ***p<0.01. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the index: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms. The Italian regions are presented in Table 4. The list of mother's origin country is presented in Table A1.

Table 12 – Educational aspiration in middle secondary school by gender

	Males				Females			
	Academic.	Technical	Vocational	Other	Academic.	Technical	Vocational	Other
Ethnic identity index	-0.074*** (0.014)	-0.063*** (0.015)	0.015 (0.015)	-0.009 (0.010)	-0.109*** (0.018)	-0.044*** (0.016)	0.022 (0.017)	-0.014 (0.011)
Attend grade 7	-0.008 (0.014)	0.041*** (0.013)	0.033*** (0.011)	-0.001 (0.010)	0.011 (0.016)	0.011 (0.012)	0.027** (0.011)	0.002 (0.011)
Attend grade 8	-0.165*** (0.013)	0.164*** (0.014)	0.196*** (0.013)	-0.061*** (0.009)	-0.101*** (0.016)	0.113*** (0.014)	0.153*** (0.013)	-0.062*** (0.010)
Reading score	0.038*** (0.005)	0.017*** (0.005)	-0.036*** (0.005)	-0.005 (0.003)	0.042*** (0.006)	0.001 (0.005)	-0.023*** (0.005)	-0.010*** (0.004)
Math score	0.048*** (0.006)	-0.017*** (0.006)	-0.004 (0.006)	-0.008* (0.004)	0.064*** (0.007)	-0.011* (0.006)	-0.030*** (0.006)	-0.009** (0.005)
Grade retention	-0.067*** (0.011)	-0.039*** (0.014)	0.064*** (0.013)	0.013 (0.009)	-0.117*** (0.018)	-0.023 (0.016)	0.086*** (0.018)	0.004 (0.012)
Grade Italian school started	0.003 (0.003)	-0.008** (0.003)	0.003 (0.003)	0.002 (0.002)	0.001 (0.004)	0.002 (0.003)	0.004 (0.003)	-0.002 (0.002)
Trust teachers	0.033*** (0.012)	-0.003 (0.014)	0.020* (0.012)	-0.004 (0.009)	0.025 (0.015)	0.001 (0.014)	0.008 (0.013)	-0.009 (0.010)
Born in Italy	0.049*** (0.013)	-0.053*** (0.015)	-0.008 (0.013)	0.004 (0.009)	0.043*** (0.016)	0.007 (0.014)	-0.031** (0.013)	-0.017* (0.010)
Play a sport	0.041*** (0.011)	0.009 (0.012)	-0.006 (0.011)	-0.009 (0.008)	0.020 (0.013)	0.001 (0.011)	-0.012 (0.010)	0.003 (0.008)
Work (yes=1)	-0.044*** (0.011)	-0.000 (0.012)	0.012 (0.011)	0.005 (0.008)	-0.049*** (0.014)	0.004 (0.012)	0.017 (0.012)	-0.004 (0.009)
Have no sibling	-0.003 (0.014)	-0.006 (0.015)	-0.007 (0.013)	0.020** (0.010)	0.016 (0.016)	0.004 (0.014)	-0.029** (0.012)	0.005 (0.010)
Mother not in the hh	-0.027 (0.026)	-0.022 (0.030)	0.055* (0.030)	-0.018 (0.019)	-0.074** (0.036)	-0.015 (0.033)	0.056 (0.038)	0.008 (0.024)
Father not in the hh	0.002 (0.015)	-0.005 (0.017)	0.001 (0.015)	0.018 (0.011)	0.036** (0.017)	-0.008 (0.015)	0.003 (0.014)	-0.015 (0.011)
Mother does not work	-0.015 (0.012)	0.005 (0.013)	0.001 (0.012)	0.000 (0.009)	-0.007 (0.014)	0.028** (0.012)	-0.026** (0.012)	-0.002 (0.009)
Father does not work	0.004 (0.015)	0.019 (0.017)	-0.010 (0.015)	-0.015 (0.011)	-0.048*** (0.018)	0.012 (0.016)	-0.005 (0.015)	0.021* (0.011)
Parental edu (Primary)	-0.016 (0.031)	0.042 (0.036)	-0.003 (0.034)	-0.033 (0.022)	-0.010 (0.044)	0.088** (0.039)	0.003 (0.039)	-0.071** (0.029)
Parental edu (Mid. secondary)	0.004 (0.021)	0.037 (0.023)	0.021 (0.022)	-0.004 (0.016)	0.059** (0.026)	0.037* (0.022)	0.012 (0.024)	-0.061*** (0.020)
Parental edu (High secondary)	0.053** (0.021)	0.051** (0.022)	-0.005 (0.020)	-0.020 (0.015)	0.092*** (0.026)	0.054** (0.021)	-0.042* (0.023)	-0.060*** (0.020)
Parental edu (Degree)	0.093*** (0.023)	0.054** (0.024)	-0.053** (0.022)	-0.017 (0.016)	0.129*** (0.028)	0.051** (0.023)	-0.067*** (0.024)	-0.054** (0.021)
Parental edu (Not reported)	0.006 (0.019)	0.021 (0.020)	-0.014 (0.019)	0.015 (0.015)	0.037 (0.025)	0.056*** (0.020)	-0.015 (0.022)	-0.055*** (0.019)
Family's wealth index	-0.003 (0.017)	0.083*** (0.017)	0.012 (0.015)	-0.032** (0.014)	0.061*** (0.022)	-0.002 (0.018)	0.017 (0.017)	-0.003 (0.013)
Living in a big city	0.025 (0.016)	0.014 (0.017)	-0.034** (0.015)	-0.008 (0.011)	0.065*** (0.018)	-0.007 (0.016)	-0.049*** (0.014)	-0.015 (0.012)
Italian region	yes							
Mother's origin	yes							
Constant	-0.269*** (0.073)	-0.060 (0.073)	0.375*** (0.066)	0.317*** (0.054)	-0.460*** (0.087)	0.149** (0.076)	0.444*** (0.070)	0.337*** (0.055)
N	6689	6689	6689	6689	5917	5917	5917	5917
R-squared	0.151	0.065	0.111	0.045	0.162	0.047	0.112	0.048

Notes: Standard errors in parentheses. * p<0.10 ** p<0.05 ***p<0.01. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the index: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms. The Italian regions are: Piemonte, Valle D'Aosta, Lombardia, Veneto, Friuli-Venezia Giulia, Liguria, Emilia Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna, Autonomous Province of Bolzano, Autonomous Province of Trento. The list of mother's origin country is presented in the Appendix, Table A1.

Table 13 - Educational aspiration in high secondary school by gender

	University	
	Males	Females
Ethnic identity index	-0.077*** (0.017)	-0.127*** (0.018)
Attend grade 10	0.020 (0.015)	0.014 (0.015)
Attend grade 11	0.082*** (0.016)	0.007 (0.016)
Attend grade 12	0.112*** (0.018)	0.035** (0.017)
Attend grade 13	0.117*** (0.019)	0.064*** (0.018)
Reading score	0.036*** (0.004)	0.026*** (0.004)
Math score	0.033*** (0.006)	0.053*** (0.006)
Grade retention	-0.128*** (0.011)	-0.148*** (0.013)
Grade Italian school started	-0.004* (0.002)	-0.003 (0.002)
Trust teachers	0.036*** (0.011)	0.053*** (0.011)
Born in Italy	0.051*** (0.017)	-0.034** (0.016)
Play a sport	0.010 (0.011)	0.066*** (0.012)
Work	-0.053*** (0.011)	-0.071*** (0.012)
Have no sibling	0.022 (0.014)	0.029** (0.013)
Mother not in the household	0.019 (0.026)	-0.046 (0.029)
Father not in the household	0.005 (0.016)	0.010 (0.014)
Mother does not work	0.005 (0.013)	-0.006 (0.013)
Father does not work	-0.022 (0.015)	-0.041*** (0.015)
Parental edu (Primary)	0.067 (0.045)	0.064 (0.046)
Parental edu (Middle secondary)	0.049** (0.023)	0.065** (0.031)
Parental edu (High secondary)	0.128*** (0.022)	0.154*** (0.030)
Parental edu (Degree)	0.225*** (0.025)	0.232*** (0.032)
Parental edu (Not reported)	0.020 (0.024)	0.027 (0.033)
Family's wealth index	0.029* (0.016)	0.069*** (0.020)
Living in a big city	0.033** (0.015)	-0.011 (0.014)
Italian region	yes	yes
Mother's origin	yes	yes
Constant	-0.308*** (0.072)	-0.267*** (0.082)
N	6664	7983
R-squared	0.173	0.162

Notes: Standard errors in parentheses. * p<0.10 ** p<0.05 ***p<0.01. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the index: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms. The Italian regions are: Piemonte, Valle D'Aosta, Lombardia, Veneto, Friuli-Venezia Giulia, Liguria, Emilia Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna, Autonomous Province of Bolzano, Autonomous Province of Trento. The list of mother's origin country is presented in the Appendix, Table A1.

Table 14 – Robustness: school aspiration in middle secondary school

	Robustness 1				Robustness 2				Robustness 3				Robustness 4			
	Academic.	Technical	Voc.	Other	Academic.	Technical	Voc.	Other	Academic.	Technical	Voc.	Other	Academic.	Technical	Voc.	Other
Ethnic identity index	-0.064*** (0.013)	-0.064*** (0.014)	0.016 (0.014)	-0.021** (0.009)	-0.093*** (0.012)	-0.061*** (0.011)	0.023** (0.011)	-0.011 (0.008)	-0.095*** (0.018)	-0.037** (0.017)	0.033* (0.018)	-0.016 (0.011)	-0.050*** (0.016)	-0.067*** (0.021)	0.024 (0.022)	0.000 (0.011)
N	7434	7434	7434	7434	11605	11605	11605	11605	5788	5788	5788	5788	4289	4289	4289	4289

Notes: Standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01. The table presents the main estimation results. We control for grade immigrant is attending, grade Italian school has started, trust teacher, born in Italy, play a sport, work, have no sibling, mother/father not in the household, parental education, living in a big city, Italian region, mother's origin country. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the ethnic identity index using the PCA analysis: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms.

Robustness analysis 1 excludes those immigrant students who where born in Italy

Robustness analysis 2 excludes those immigrant students who come from wealthier households and arrived in Italy in the last 5 year

Robustness analysis 3 excludes the largest ethnic group in Italy: Romanian, Albanian, Marocco, China

Robustness analysis 4 perform the analysis for immigrant students in their last middle secondary year (grade 13)

Table 16 – Robustness: Educational aspiration in high secondary school

	Robustness 1	Robustness 2	Robustness 3	Robustness 4
	University	University	University	University
Ethnic identity index	-0.101*** (0.014)	-0.105*** (0.013)	-0.075*** (0.018)	-0.129*** (0.043)
N	11817	13618	7234	1812
R-squared	0.205	0.197	0.197	0.246

Notes: Standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01. The table presents the main estimation results. We control for grade immigrant is attending, grade Italian school has started, trust teacher, born in Italy, play a sport, work, have no sibling, mother/father not in the household, parental education, living in a big city, Italian region, mother's origin country. The ethnic identity index is constructed using the PCA analysis. 11 indicators are used to construct the ethnic identity index using the PCA analysis: understand; speak; read; write; think in Italian; speak Italian at home; watch Italian TV; identify as Italian; wish to live in Italy; Italian peers; Italian friends. The asset ownership indicators used to construct the family's wealth index are: separate kitchen; toilet; shower; hot water, drinkable water; heating system; refrigerator; washing machine; dishwasher; television; computer; motorcycle; car; own private room; number of rooms.

Robustness analysis 1 excludes those immigrant students who where born in Italy

Robustness analysis 2 excludes those immigrant students who come from wealthier households and arrived in Italy in the last 5 year

Robustness analysis 3 excludes the largest ethnic group in Italy: Romanian, Albanian, Marocco, China

Robustness analysis 4 perform the analysis for immigrant students in their last middle secondary year (grade 13)

APPENDIX

Table A1 – Full list of reported origin countries

Albania	Czech Republic	Ivory Coast	Niger	Stateless
Afghanistan	Czechoslovakia	Jamaica	Nigeria	Sudan
Algeria	Democratic Rep.	Japan	North Korea	Suriname
Andorra	Denmark	Jersey Island	North Macedonia	Sweden
Angola	Djibouti	Jordan	Norway	Switzerland
Antigua & Barbuda	Dominica	Kazakhstan	Nuova Guinea	Syria
Argentina	Dominican Republic	Kenya	Oman	Taiwan
Armenia	Dutch Antilles	Kiribati	Pakistan	Tajikistan
Australia	East Timor	Kuwait	Palau	Tanzania
Austria	Ecuador	Kyrgyzstan	Palestine	Thailand
Azerbaijan	Egypt	Laos	Panama	Togo
Bahamas	El Salvador	Latvia	Papua New	Tonga
Bahrain	Equatorial Guinea	Latvians non-citizen	Paraguay	Trinidad & Tobago
Bangladesh	Eritrea	Lebanon	Peru	Tunisia
Barbados	Estonia	Lesotho	Philippines	Turkey
Belarus	Eswatini	Liberia	Poland	Turkmenistan
Belgium	Ethiopia	Libya	Portugal	Tuvalu
Belize	Micronesia	Liechtenstein	Qatar	Uganda
Benin	Fiji	Lithuania	Republic of	Ukraine
Benin	Finland	Luxembourg	Republic of	United Arab
Bhutan	Former Yugoslavia	Madagascar	Romania	United Kingdom
Bolivia	France	Malawi	Russian	United States of
Bosnia-	Gabon	Malaysia	Rwanda	Uruguay
Botswana	Gambia	Maldives	Saint Kitts and	Uzbekistan
Brazil	Georgia	Mali	Saint Lucia	Vanuatu
Brunei	Germany	Malta	Saint Vincent &	Vatican City
Bulgaria	Ghana	Marshall Islands	Salomone	Venezuela
Burkina Faso	Greece	Mauritania	Samoa Islands	Vietnam
Burundi	Grenada	Mauritius	San Marino	Yemen
Cambodia	Guatemala	Mexico	São Tomé &	Zambia
Cameroon	Guinea	Moldova	Saudi Arabia	Zimbabwe
Canada	Guinea Bissau	Monaco	Senegal	
Cape Verde	Guyana	Mongolia	Serbia	
Central African	Honduras	Montenegro	Seychelles	
Chad	Hungary	Morocco	Sierra Leone	
Chile	Iceland	Mozambique	Singapore	
China	India	Myanmar	Slovakia	
Colombia	Indonesia	Namibia	Slovenia	
Comoros	Iran	Nauru	Somalia	
Costa Rica	Iraq	Nepal	South Korea	
Croatia	Ireland	Netherlands	South Sudan	
Cuba	Isle of Man	New Zealand	Spain	
Cyprus	Israel	Nicaragua	Sri Lanka	

Table A2 – Descriptive statistics for immigrants students by gender

	Middle secondary school				High secondary school			
	Males		Females		Males		Females	
	mean	sd	mean	sd	mean	sd	mean	sd
Reading score	6.303	1.324	6.513	1.303	6.042	1.514	6.260	1.449
Math score	6.281	1.077	6.697	1.086	6.183	1.077	6.503	1.086
Grade retention	0.237	0.425	0.128	0.334	0.357	0.479	0.256	0.437
Academically oriented school	0.258	0.438	0.428	0.495				
Technical school	0.287	0.452	0.199	0.399				
Vocational school	0.225	0.418	0.183	0.387				
Professional course/work/other	0.096	0.295	0.096	0.295				
Aspiration for university					0.330	0.470	0.564	0.496
Attend grade 6	0.333	0.471	0.323	0.470				
Attend grade 7	0.333	0.471	0.329	0.470	0.000	0.000	0.000	0.000
Attend grade 8	0.334	0.472	0.347	0.476	0.000	0.000	0.000	0.000
Attend grade 9	0.000	0.000	0.000	0.000	0.348	0.476	0.291	0.454
Attend grade 10	0.000	0.000	0.000	0.000	0.211	0.408	0.220	0.414
Attend grade 11	0.000	0.000	0.000	0.000	0.185	0.388	0.196	0.397
Attend grade 12	0.000	0.000	0.000	0.000	0.139	0.346	0.164	0.370
Attend grade 13	0.000	0.000	0.000	0.000	0.117	0.322	0.129	0.335
Grade Italian school started	2.409	2.168	2.391	2.201	3.867	3.227	3.814	3.201
Trust teachers	0.783	0.412	0.800	0.400	0.608	0.488	0.584	0.493
Born in Italy	0.402	0.490	0.419	0.494	0.195	0.396	0.192	0.394
Play a sport	0.656	0.475	0.402	0.490	0.624	0.484	0.311	0.463
Work	0.352	0.478	0.253	0.435	0.384	0.486	0.267	0.443
Have no sibling	0.203	0.402	0.200	0.400	0.274	0.446	0.245	0.430
Mother not in the household	0.035	0.183	0.026	0.159	0.066	0.248	0.038	0.190
Father not in the household	0.177	0.381	0.226	0.418	0.250	0.433	0.287	0.452
Mother does not work	0.357	0.479	0.364	0.481	0.362	0.481	0.331	0.471
Father does not work	0.145	0.352	0.160	0.366	0.176	0.381	0.184	0.388
Parental educ: primary	0.029	0.168	0.025	0.157	0.015	0.123	0.020	0.141
Parental educ: low secondary	0.152	0.359	0.163	0.369	0.154	0.361	0.167	0.373
Parental educ: high secondary	0.240	0.427	0.268	0.443	0.413	0.492	0.475	0.499
Parental educ: degree	0.145	0.352	0.146	0.353	0.226	0.418	0.205	0.404
parent missing	0.344	0.475	0.328	0.470	0.137	0.343	0.098	0.297
Family's wealth index	3.197	0.316	3.177	0.304	3.226	0.319	3.214	0.268
Living in a big city	0.203	0.402	0.205	0.404	0.299	0.458	0.327	0.469
Region: Piemonte	0.073	0.260	0.070	0.255	0.071	0.257	0.084	0.277
Region: Valle D'Aosta	0.015	0.123	0.013	0.113	0.009	0.093	0.009	0.092
Region: Lombardia	0.078	0.268	0.074	0.262	0.091	0.288	0.090	0.286
Region: Veneto	0.087	0.282	0.090	0.286	0.098	0.297	0.106	0.308
Region: Friuli-Venezia Giulia	0.052	0.222	0.049	0.216	0.062	0.241	0.050	0.217
Region: Liguria	0.052	0.222	0.051	0.219	0.056	0.230	0.079	0.269
Region: Emilia Romagna	0.077	0.266	0.085	0.279	0.117	0.321	0.124	0.329
Region: Toscana	0.074	0.262	0.068	0.252	0.071	0.257	0.064	0.244
Region: Umbria	0.048	0.213	0.050	0.218	0.041	0.199	0.042	0.202
Region: Marche	0.046	0.211	0.048	0.214	0.048	0.213	0.030	0.171
Region: Lazio	0.055	0.227	0.056	0.231	0.050	0.218	0.071	0.257
Region: Abruzzo	0.047	0.212	0.048	0.213	0.040	0.196	0.035	0.183
Region: Molise	0.009	0.096	0.008	0.089	0.007	0.082	0.009	0.096
Region: Campania	0.062	0.241	0.067	0.251	0.038	0.191	0.037	0.190
Region: Puglia	0.032	0.177	0.036	0.187	0.040	0.196	0.037	0.189
Region: Basilicata	0.014	0.118	0.011	0.103	0.010	0.101	0.006	0.080
Region: Calabria	0.037	0.190	0.035	0.184	0.042	0.201	0.031	0.174
Region: Sicilia	0.051	0.219	0.051	0.221	0.047	0.213	0.036	0.186
Region: Sardegna	0.015	0.121	0.015	0.120	0.014	0.117	0.014	0.117
Region: Provincia autonoma di Bolzano	0.033	0.180	0.033	0.179	0.015	0.120	0.018	0.134
Region: Provincia autonoma di Trento	0.042	0.201	0.042	0.201	0.033	0.180	0.028	0.166
	6689		5917		6664		7983	

Table A3 – Pearson correlation

	Underst.	Speak	Read	Write	Think	Speak home	Watch TV	Identif.	Migr. history	Peers	Friends
Understand	1.000										
Speaking	0.7279	1.000									
Read	0.5538	0.5756	1.000								
Write	0.4803	0.5096	0.5715	1.000							
Think	0.0629	0.0963	0.0885	0.0953	1.000						
Speak home	0.0932	0.1205	0.0924	0.0933	0.2151	1.000					
Watch TV	0.0400	0.0533	0.0599	0.0621	0.2632	0.1529	1.000				
Identification	0.0180	0.0357	0.0387	0.0448	0.3114	0.1292	0.2172	1.000			
Migr. history	-0.0199	-0.0274	-0.0150	-0.0178	0.1022	0.0475	0.0407	0.1456	1.000		
Peers	0.0665	0.0880	0.0828	0.0899	0.1997	0.1145	0.1508	0.1696	0.0309	1.000	
Friends	0.0747	0.1030	0.0837	0.0935	0.2163	0.1230	0.1618	0.1896	0.0401	0.4700	1.000

Table A4 - Eigenvalues and cumulative proportion

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	2.87652	0.947389	0.2615	0.2615
Comp2	1.92913	0.815845	0.1754	0.4369
Comp3	1.11328	0.14494	0.1012	0.5381
Comp4	0.968343	0.0991741	0.0880	0.6261
Comp5	0.869169	0.104962	0.0790	0.7051
Comp6	0.764207	0.0972599	0.0695	0.7746
Comp7	0.666947	0.0672641	0.0606	0.8352
Comp8	0.599683	0.0706452	0.0545	0.8898
Comp9	0.529038	0.11502	0.0481	0.9378
Comp10	0.414018	0.144353	0.0376	0.9755
Comp11	0.269666		0.0245	1.0000

Table A5 – Principal components (eigenvectors)

Variables	Comp 1
Understand	0.4629
Speak	0.4806
Read	0.4549
Write	0.4303
Think in Italian	0.1839
Watch Italian TV	0.1622
Speak Italian at home	0.1369
Self-identification	0.1271
Migr. history	0.0143
Peers	0.1789
Friends	0.1881