

Ethnic identity and educational aspirations

Very preliminary not complete draft – do not cite without permission

Teresa Randazzo & Matloob Piracha

Abstract

We study the role of ethnic identity on educational preferences of children of immigrants in Italy. We find that a weak Italian identity negatively affects immigrant educational careers. In particular, students in middle secondary school with a weak sense of Italian belonging show a low preference towards academically-oriented high secondary tracks which normally increase the likelihood to pursue a university degree. Moreover, the intention of immigrant children in high secondary schools to enrol at university decreases if they have a weak Italian identity. The effects are much larger for girls who show to be more sensitive to the assimilation process. Immigrant children will soon form a very important component of the adult population of Italian society and shedding light on their educational careers will help understand their performance in the Italian labour market.

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, Carillo et al, 2021; Piracha et al, 2021; Cai and Zimmermann, 2020, Casey and Dustmann, 2010; Nekby and Rodin 2010; Constant et al, 2009). However, to the best of our knowledge there is no study that considers the role ethnic identity plays in the performance and aspirations of school going children of immigrants. There are only two related papers, Nekby et al (2009) and Schuller (2015). The first one analyses the correlation between acculturation and education attainment, but their focus is on those who manage to complete university education, while the second explores the correlation between parents' ethnic identity and their children's education outcomes. Unlike both papers, we study the impact of ethnic identity of the middle- and high-school pupils on their intentions to pursue further education.

The limited amount of research in economics that is related to 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.¹ 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).² Similarly, Patacchini and Zenou (2015) 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-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.

Association between ethnic identity and education has been indirectly explored in the economics literature. Following Berry et al (1989), Constant and Zimmermann (2009) use a two dimensional identity measure that they call the *ethnosizer*, which assumes four possibilities: assimilation, integration, marginalization and separation. 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. Using data from Germany they find that migrants who acquired higher education prior to migration tend to integrate rather than assimilate. Using the same German data, Zimmermann et al (2006) find that human capital acquired in the country of origin has a detrimental effect on migrants’ identifying with the majority culture.

¹ 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.

² Battu and Zenou (2010)and ... look at similar aspects using UK data.

Nekby and Rodin (2010) depart from the previous literature by analysing the role of ethnic identity of second generation immigrants in the completion of higher education degrees. Using data from Sweden they construct a two-dimensional measure of ethnic identity *a' la* Berry (1998) and Constant et al (2008).

Using Italian schooling data, Carlana *et al* (2021) show that immigrant pupils are more likely to enrol in vocational or technical education rather than academic-oriented one compared to the natives of the same ability level. They call this phenomenon “education segregation”. 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 outcome of immigrants. We, therefore, analyse the impact of ethnic identity of teenage immigrants on their plans to pursue further education.

To our knowledge this is the first paper that explores the role of school going teenage immigrants’ ethnic identity on their education intentions. Our results show that having a weak sense of Italian identity negatively affects the preference of middle high school immigrants for academically-oriented high schools which offer a more theoretical background and increase the likelihood to pursue a university degree. When we move our attention to high secondary school immigrants, we find that a low commitment to the Italian culture has also a negative effect on the intention to enrol at the university. Education is an important determinant of economic outcomes and shedding light on immigrants’ educational careers will help us to understand how immigrants perform into the Italian labour market.

2. Theory and measurement of ethnic identity

The underlying notion of how ethnic identity manifests itself in economic outcomes through its importance in the formation of social networks. Battu *et al* (2007) argue that since whites don’t suffer from discrimination, the non-whites (immigrants) might prefer interacting with whites as that social network is likely to be more rewarding in the labour market. However, because of peer pressure and social preferences within their own “community/group”, each non-white person’s interaction with the white person or group decreases the non-white group’s overall utility. The individuals for whom the non-white group utility is more important than their own labour market performance might form *oppositional identities*, i.e., stay closer to their own ethnic group by not mixing with whites.

We follow Constant et al. (2009) and build an index along the lines of the one-dimensional *ethnosizer* in their paper.³ We use five measures: (i) language proficiency, (ii) culture, (iii) ethnic self-identification, (iv) migration history and (v) social network. Below we explain the details of each category and the construction of the ethnosizer. The lower the value of the ethnosizer the higher is the level of assimilation.

3. Secondary education in Italy

The average age of the foreign population residing in Italy is about 35 years old and the foreign presence is more concentrated in the regions of the Centre-North.⁴

In Italy, children 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, three years of middle school and five years of high school. 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. In particular, while academic schools offer a more theoretical background and increase the likelihood to pursue a university degree, 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.

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 the children of mixed couples, we should raise the figure by approximately another two million, bringing the foreign population from 10 to over 12 percent of the total population of Italy. Since migrants have a higher fertility rate⁵, the number of migrant children in Italian schools and universities were approximately one million at the beginning of 2020. 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

³ ISTAT does not have the kind of data that is needed to build a *two-dimensional ethnosizer*.

⁴ Lombardy is the region with the highest migratory presence followed by Lazio, Emilia-Romagna and Veneto.

⁵ According to the latest data from Istat, the fertility rate of migrants is 1.98 children against an Italian fertility rate of 1.27.

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, the most represented in Italy are Moroccans. China is the country of origin of over 55,000 students (6.4%), with a growth of almost 80% in the last 10 years.

According to the provisions of the Ministry of Education, the number of students with non-Italian citizenship with reduced knowledge of the Italian language must not normally exceed 30% of the students in a class or school. 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 years 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. The dropout phenomenon is probably partly influenced by another negative trend: school delay which is much more pronounced for foreigners. Related to this is their academic achievement and type of education they choose: academic-oriented, technical or vocation. Approximately 40% of the immigrant pupils choose an academic-oriented school, whereas ??% Italians make the same choice. Consistent with the drop out rate, more female immigrants choose academic-oriented schools compared to their male counterparts.

Immigrant children will soon form a very important component of the adult population of Italian society and shedding light on their choice of education will help understand their occupational outcomes.

4. Data and set variable description

The empirical analysis uses data from the “integration of second generation survey” conducted in 2015 by “Istituto Nazionale di Statistica” (Istat) and financed by the European Integration Fund.⁶ The survey is based on a sample of state secondary schools with at least 5 migrant students. In Italy, secondary school is articulated in low or middle secondary school, from grade 6 to grade 8, and high secondary school, from grade 9 to grade 13.

⁶ Istat carried out the survey on the integration of second generations between March and June 2015 generations in collaboration with the Ministry of Education, University and Research. The survey is available in the Istat website at <http://www.istat.it/it/archive/182866>

The sample considers 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 are immigrants of first or second generation with a foreign citizenship.⁷ In our study we restrict the attention to immigrants in secondary school. Our final sample, after excluding students with missing relevant information, considers 12,945 students between the grade 6th and 8th [middle secondary school] and 14790 students between the 9th and 13th grades [high secondary school]⁸ with no Italian citizenship. Only 19% of this latter group of students were born in Italy. The percentage rises to 39% when we consider immigrants in middle secondary school.⁹

Table 1 reports the descriptive statistics for the sample of interest. In our sample, 46% of students in middle school are female and 54% are male. The percentages are reversed in high secondary school which implies that a few male pupils dropout after grade 8. Around 55% of students in middle school play a sport; the percentage reduces to 46% in case of students in high school. Students working at least a few times a month are respectively 31% for middle secondary and 32% for high secondary. We do not have information on students' types of jobs. 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 30% for middle and high secondary schools, respectively, live in a big city showing that a large share of immigrant families with children prefers living in small cities or towns. Finally, as expected, Veneto, Lombardia and Piemonte are the regions with a higher share of immigrant students.

Families contribute to understand children outcomes and in the case of immigrants they play a decisive role in the integration processes of their young members. Looking at the household characteristics, 42% have one sibling. This percentage reduces respectively to 36%, for middle school students, and 31%, for high secondary students, when we consider having at least two siblings showing that also immigrants' families are not very large. Mother does not work for around 35% of the respondents and that percentage reduces to 13-15% when the question refers to non-working fathers.

⁷ The status of foreign citizenship persists also for children born in Italy if their parents do not hold Italian citizenship.

⁸ Some protective measures were adopted to protect the identity of the young interviewers, which means that some key variables are not available in the public dataset. For instance, the exact age is not available in the public data. We know students normally start the 6th grade at age 10 or 11 and conclude the 13th grade at age 18 or 19.

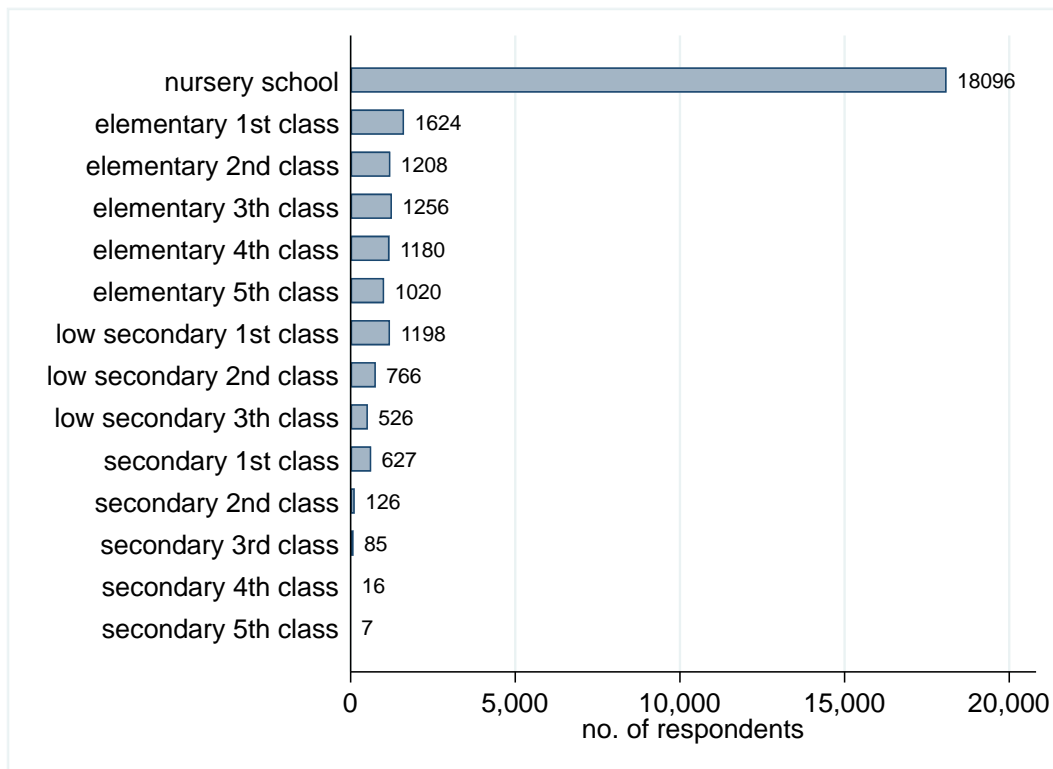
⁹ The information on the type of high secondary education in which students are enrolled is not provided.

Immigrant students come from well-educated families. The highest level of parental education, for around 44% of the interviewed students in high secondary, is at the secondary level while 22% of this respondents' group have at least a parent holding a university degree. Among students in middle secondary, only 25% of them have secondary education as highest level of parental education and those having a parent holding a degree are just 16%. The differences in parental education between the two groups of students may be due to the fact that students in middle secondary might not remember the level of parental education.¹⁰ Together with individual and household characteristics, Table 1 reports some school information. The school is, in fact, an important place for socialization outside the family control. Often for immigrant pupils the school is the social space where they experience the cultural environment of the host country. As shown in Figure 1, around 18,096 immigrant students started Italian education at the nursery school (what we call the zero grade) and on average the survey respondents have been in the Italian educational system since the 2nd or 3rd grade of the elementary school.

Table 1 also shows that respectively 18% and 30% of the middle secondary and high secondary respondents have repeated a grade. As expected, grade retention increases with the level of education. With respect to school performance, students report the latter grade in the subjects Italian and math. Scores go from 1 to 10. In both subjects the average score is around 6 which is the minimum score for not failing the subject. Finally, according to 18% of students in middle secondary, teachers adopt a discriminatory behaviour in class. Those who report a teachers' discriminatory behaviour in class are 30% in the case of high secondary students. This is a quite large percentage that could affect educational outcomes and aspirations.

¹⁰ In fact, this information is not reported for 32% of them

Figure 1: Grade at which Italian school has started



Our outcome variables are (i) for immigrants in middle secondary: preference for high secondary track, and (ii) for students in high secondary: intention of enrol to University, which represent an investment in future education. The summary statistics show that on average 34% of students in middle secondary is planning to go to an academically-oriented school, 24% prefer a technical track and 20% prefer to pursue vocational education. Only 10% is thinking for alternative options (e.g. working). Finally, on average almost 50% of migrants in high secondary have intention to pursue tertiary education. The willingness to invest in further education could represent a desire for social redemption which is certainly stronger when teenagers are comfortable in the country where they live.

Table 1 about here

Our variable of interest is an ethnic identity index constructed along the lines of the one-dimensional ethnosizer described in Constant et al. (2009). Our index, therefore, comprises five categories which together define the individual ethnic identity: (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. When the affinity with the destination country is at its maximum the index takes the value equal to 0 and the respondent is fully assimilated in Italy. On the

other hand, when there is least commitment to the Italian culture, and therefore the original ethnic identity has not been altered at all, the identity index takes value equal to 1. More details about how each category is constructed is as follows.

Language: We have information on how well Italian is spoken, written, understood and listened. 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 if the ability is very good; a value of 0.25 is given for those reporting good command of Italian; those who have fair knowledge of Italian get a value of 0.5; a value of 0.75 is assigned if ability in Italian is poor; and, finally, a value of 1 is assigned if the ability is very poor. The overall average value of our four abilities in Italian (speak, write, understand and listen) define the language proficiency.

Culture: It considers the visible Italian cultural elements which we identify in two variables (i) thinking in Italian and (ii) watching Italian TV. Again, a value of 0 is assigned for each of the two categories and a value of 1 otherwise. Culture is the mean assigned value of the two questions related to cultural elements.

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 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 0.5 is given when friends are a combination of Italian and foreigners; finally, if friends are exclusively foreigners, with the same or different origin, we attach a value of 1. The mean assigned values of the two answers referring to friends within and outside school generate the social network category.

Table 2 presents the specific variables related to each category and the related average value. As discussed above, each variable will take a value between 0 (maximum commitment and assimilation to the Italian culture) and 1 (no commitment to the Italian culture) and the ethnic identity index is simply the mean assigned value of answers to the questions of the five categories. Overall, its mean value is around 0.30 showing a good commitment of immigrant children to Italy. However,

differences exist between the five categories. On the one hand, immigrants have excellent ability in the Italian language, they show a good Italian social network and they are close to Italian cultural elements. For these three categories the average score is close to zero. In particular, migrant children have a good social network. Only 26% of them do not meet Italian friends (schoolmates or friends outside school).

On the other hand, the question on self-identification and the question asking whether the respondent wish to live in Italy show that the assimilation process is a bit fragile in these two categories. Around 56% and 63% immigrants in 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 2 around here

5. Empirical strategy

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

$$y_{ijr} = \alpha + \gamma EI_{ijr} + \sum_{k=1}^k \beta_k X_{ijkkr} + \delta_r + \varepsilon_{ijr} \quad (1)$$

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 two outcome variables are a binary indicator for each alternative high school choice in which a student in middle secondary plans to enrol and a binary indicator for whether the respondent in high secondary intends to pursue further education after finishing school. X_{ikr} denotes the k exogenous variables, δ is a region fixed-effect and ε is the error component. 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 socio-economic characteristics. The set of characteristics included in X_{ikr} are variables related to the schooling path (for example, grade in which respondents are enrolled and school performance), respondents' characteristics (for example, age and born in Italy), household characteristics (for example, parental education). The wealth position of the children' household is captured by a wealth index, constructed using the principal component analysis (see Filmer and Pritchett, 2001).

5.1 High secondary track

Preferences for high secondary school tracks are modelled with a discrete choice model. Students in middle secondary can consider four mutually exclusive alternatives: academically-oriented, technical, vocational, other. Our variable EI is likely to be endogenous. Indeed, unobservables could simultaneously affect the variable ethnic identity and the choice of high secondary school. Specifications which ignore the correlation of unobservable variables will lead to biased and inconsistent estimate of EI on y . Moreover, the problem of omitted variable bias is even larger in a cross-sectional setting. We use an instrumental variable approach to correct the potential bias of EI . Our empirical strategy relies on the identification of a valid instrument correlated with the ethnic identity index but not directly affecting the educational outcome. The first-stage regression is

$$EI_{ir} = \alpha + \lambda D_i + \sum_{k=1}^k \beta_k X_{ikr} + \delta_r + \mu_{ir} \quad (2)$$

where λ is the parameter of the instrumental variables chosen and D represents the proximity of foreign languages spoken at home to Italian. The error terms ε and μ are independent of D and X but are correlated with each other.

The instrument D is constructed using the average distance between the first two languages spoken at home. We extract the distance between foreign languages to Italian from eLinguistics.net which provides the relatedness (generic proximity) between 220 languages. The distances are expressed as values from 0 (the nearest distance - so the same language) to 100 (biggest possible distance between two languages). Italian is spoken at home for the 50 per cent of our respondents in middle secondary education. Language closeness makes the assimilation process easier. The linguistic and cultural categories of our ethnic identity index are highly correlated with closeness of languages to Italian. Any effect of languages on immigrants' high secondary choice is captured directly from our ethnic identity index and Italian and math scores which help explain educational choices. We claim that our instrument is valid as long as we control for actual school performance and in particular for the latest score in the Italian subject. Indeed, Italian score and ethnic identity index capture directly language effects leaving our instrument representing a proxy for cultural distance and ties to an ethnic social group. —Cai and Zimmermann use language proximity using Chinese data whereas Islam and Raschky use cultural distance.

Following Vadean et al (2019), we estimate a system of linear probability equations using a 3SLS model, which allows the simultaneous estimation of the coefficients for the entire system and account for the correlation structure in the disturbances across the school outcome and indicator equations, producing consistent estimates (see Zellner and Theil, 1962). We run the 3SLS estimator using the user written command `cmp` in Stata 15.

5.2 University choice

We investigate the choice of pursuing further education through a binary outcome capturing the intention to enrol at university for the sample of immigrant students in high secondary.

In order to solve the potential bias arising from linear probability estimates and to capture the causal impact of ethnic identity, we again use an instrumental variable (IV) approach estimating equations (1) and (2) through a two-stage least squares (2SLS) model. Again, we claim that, our instrument, the linguistic distance to Italian, is strongly correlated with our ethnic identity index which helps to explain the assimilation process but not with the outcome variable. There is no reason to believe that after controlling for individual and household characteristics as well for a number of school related variables the distance of languages spoken by our respondents has a direct impact on the intention to pursue further education. As before, it is crucial to control for the latest Italian and math scores. Any impact of languages on the choice to enrol to university is directly captured by our ethnic identity index and latest scores. In other words, our measure of linguistic distances has not impact on university choice except through the ethnic identity index and school performance.

6. Results

6.1 High secondary track

We first run a linear probability model (LPM) for each alternative outcome as a baseline for the analysis of the effect of ethnic identity on school-track preference. Our ethnic identity index, ranging from 0 to 1, captures a weak sense of Italian identity. If, on the one hand, the value 0 captures a strong sense of belonging to Italy and its culture, on the other hand, the value 1 means no commitment to the Italian culture and society. When the endogeneity of ethnic identity is not taken into account we find a negative impact of our identity index on the preference for academically-oriented (-21%) and technical schools (-9.8%). In other words, having a weak sense of Italian identity disincentives the preference for high secondary schools which offer much better education and employment prospects than vocational schools. When we estimate a 3SLS model we find, on the one hand, that the negative effect of our ethnic identity index is no longer significant for technical schools. On the other hand, the negative effect on academically-oriented tracks persists and the coefficient, after controlling for endogeneity, is slightly larger (-25%). We present the summary of all estimation results in Table 3.

The extended set of results based on the 3SLS estimation of model equations (1) and (2) are reported in Appendix, see Table A1. Columns (2)-(5) present estimates of the school-track preferences, our second-stage regression. As expected, we find that school performance affects the school track. For

example, higher scores in Italian and math, on the one hand, increase the intention to choose an academically-oriented high school. On the other hand, they decrease the preference for vocational tracks or other. Repeating a grade decreases the preference for academically-oriented and technical tracks while the intention to enrol to a vocational school increases of 7.8%. Perceive a discriminatory teacher behaviour decreases of 22% the preference for an academically-oriented track. This is the most challenging track for a student and school performance as well as teacher behaviour influence the attitude toward academic schools. None of the other tracks are affected by how teachers behave in class.

Girls seem more orientated toward academic respect to technical and vocational schools. In fact, being a girl increases the preference for academically-oriented of 13% while the preference for technical and vocational schools decreases respectively of 8.6% and 2.1%. Moreover, while practice a sport affects positively the preference for academically-oriented tracks (+3.9 per cent), working has the opposite effect. It, on the one hand, decreases the preference for academically-oriented school (4.4 per cent), on the other hand, increases the preference for vocational tracks (+1.6%) which prepare students for immediate employment in low-skilled jobs. Parental education is also an important determinant of children educational career. Parents holding high secondary education or a degree increase the children preference for academically-oriented (+5.3% and +10.3% respectively) and technical schools (+7.1% and 6% respectively). Instead, the preference for a vocational school decreases as the level of parental education increases (-2.8% if at least a parent hold high secondary and -6.9% if at least a parent holds a degree). Parents have a large influence on how children perceive education. Education is a means for achieving better economic outcomes and parents with a higher level of education push their children to prefer high secondary school which can facilitate university enrolment. In fact, we find a negative effect of parental education on other preferences (for example, working). Children from wealthier families prefer academically-oriented (+1.1%) and technical schools (1.3%). Finally, children living in a big city prefer academically-oriented (+4.6%) and technical schools (+3.1%) respect to vocational (-4.9%) or other (-1.7%).

Column 1 shows estimates of our first-stage regression in which we regress our ethnic identity index on a set of explanatory variables and the instrument average language distance. F-statistic of the excluded instrument is very high, suggesting the instrument is not weak. The explanatory variables included in the ethnic identity equation show the expected sign. A weak Italian identity is negatively affected by being born in Italy suggesting the development of an attachment to Italy if it is the country of birth. Diversely, the grade in which Italian school has started is a proxy for when students have moved to Italy. Enrolling to Italian schools older does not help the assimilation process and therefore

the formation of Italian identity. Starting Italian school one grade later increases of 21% the weak sense of Italian identity

When we move our attention to some school related variables, we find that grade retention increases of 17% our ethnic identity index and therefore the distance from the Italian sense of belonging. Repeating a grade can be a traumatic experience for a child who might feel different from the rest of its schoolmates. This perception could be larger for immigrant children and consequently affects their Italian sense of identity. We expect that having a good performance at school, instead, has a positive effect on the Italian identity and in fact we find it when we consider the latest score in math. A better score in math decreases of 25% the ethnic identity index and therefore the weak Italian identity. The story does not hold for the Italian score which presents the opposite sign. We believe that the unexpected sign is a consequence of the high correlation of this variable with our instrument. In fact, in the LPM estimates we have exactly the opposite sign as expected.

Family characteristics are also important in shaping identity. Distance to Italian identity decreases as the parental level of education increases. A higher level of education allows migrants to settle in the destination country under better economic conditions that can facilitate the assimilation process of migrants and their families. The same argument applies to the wealth index, constructed using the principal component analysis, which is a proxy for economic conditions. Weak Italian identity decreases of 29% as the wealth index increases. The weak Italian identity is also negatively affected by either mother or father does not work. (think why?)

6.2 University choice

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, the choice 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. When we model our outcome variable with a linear probability model (LPM) we find that a weaker sense of Italian identity decreases the preference for enrolling to the university of 23%. The negative effect gets larger (-31%) when we estimate a 2SLS model using an instrumental variable approach to take care of the potential endogeneity of our ethnic identity index. Table 4 presents the summary of all estimation results.

The estimated effects of the other covariates included are in line with the results expected; see Table A4 in the appendix. The intention to enrol at the university increases with the grade of high secondary

school. In fact, it is during grade 13 that students decide concretely whether to pursue further education. Several university tracks request an exam to access them and students normally start preparing for it during their final high secondary year. As expected, performing well at school contribute in modelling preferences. A higher Italian or math grade increases the intention to enrol at the university (+3.1 and +3.8 respectively) while grade retention has a negative effect (-12%). School teachers also influence students' choices. A discriminatory teacher behaviour decreases the preference for university (-1.9%). Looking at the respondent's characteristics, the preference for university is positively affected by being a girl (+20%) and being born in Italy (+20%). Practicing a sport allows to meet and play with other peers and this interaction influence and shape the formation of preferences also related to the future. We find that practicing a sport increases preference for university (+3.5%). Instead, working, even if not regularly, during school-age decreases the intention to enrol at the university. Living in a large family, in which there are two or more siblings decrease the preference for university (why? see if there is something in the literature). Finally, as expected the intention to pursue further education increases as the highest level of parental education increases.

Moving our attention to the determinants of the ethnic identity index...

6.3 Educational career by gender

When we run separated estimates for immigrant boys and girls, we find that overall, ethnic identity has a larger weight in shaping girl educational aspirations. The baseline analysis for children in middle secondary shows that a weak sense of Italian identity has a larger negative effect on boys (-0.11%) respect to girls (-7.7%). The effect gets null when the endogeneity of the ethnic identity index is taken into account through the 3SLS model. Instead, girls are more affected by being less assimilated when academically-oriented high schools are taken into consideration. The baseline analysis shows that a weak sense of Italian identity decreases the preference for an academically-oriented track by 26.4% for girls and 16% for boys. The negative effect is still larger for girls when we run a 3SLS model. In particular, we find that for girls their preference for academically-oriented school decreases of 27.2% while the negative effect for boys is 23.4%. The level of significance passes from 1% to 5% for girls and from 1 per cent to 10% for boys. In Table 3 we also report the summary of the relevant estimates by gender. We note that.....

Gender differences in aspiration to pursue further education exist when we focus on the sample of immigrant high secondary students. A weak Italian identification touches girls more than boys. The baseline estimates show that the intention to enrol at university decreases more for girls (-26%) than for boys (-0.18%) when there is a low commitment to the Italian culture. This gender difference exist even when we run a 2SLS model, taking care of the potential endogeneity of our ethnic identity index.

The negative coefficients get slightly larger and the effect is stronger for girls (-36%) respect to boys (-24%).....

7. Conclusion

References

- Akerlof, George A. and Rachel E. Kranton. 2002. "Identity and Schooling: Some Lessons for the Economics of Education." *Journal of Economic Literature*. December, 40 (4): 1167–1201
- Akerlof, George A., and Rachel E. Kranton. 2000. Economics and identity. *Quarterly Journal of Economics* 115 (3): 715-753.
- Battu, Harminder, Mwale McDonald, and Yves Zenou. 2007. Oppositional identities and the labor market. *Journal of Population Economics* 20 (3): 643-667.
- Battu, Harminder, and Yves Zenou. 2010. Oppositional identities and employment for ethnic minorities: Evidence from England. *Economic Journal* 120 (542).
- Cai, Shu and Klaus F. Zimmermann. 2020. Social assimilation and labor market outcomes of migrants in China. *Global Labor Organization (GLO) DP 716*.
- Carlana, M., La Ferrara, E. and Pinotti, P. (2021). 'Goals and gaps: educational careers of immigrant children', forthcoming in *Econometrica*.
- Carillo, Maria Rosaria, Vincenzo Lombardo, and Tiziana Venittelli. 2021. Identity and labour market outcomes of immigrants. *Global Labor Organization (GLO) DP 749*.
- Constant, Amelie F., Liliya Gataullina, and Klaus F. Zimmermann. 2009. Ethnosing immigrants. *Journal of Economic Behavior & Organization* 69 (3): 274-287.
- 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.
- Nekby, Lena, and Magnus Rodin. 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? *Int Mig Rev* 43(4):938–973

Patacchini, Eleonora and Zenou, Yves. 2016. Racial identity and education in social networks. *Social Networks* 44: 85-94

Piracha, M., Tani, M., Cheng, Z., Wang, B.Z. (2021). Ethnic Identity and Immigrants' Labour Market Outcomes. IZA Discussion Paper No. 14123

Schuller, Simone (2015). Parental ethnic identity and educational attainment of second-generation immigrants. *Journal of Population Economics* 28: 965–1004.

Table 1 – Descriptive Statistics

VARIABLES	Middle		High secondary	
	mean	sd	mean	sd
Aspiration for education				
preference for academically oriented school	0.347	0.476	-	-
Preference for technical school	0.244	0.429	-	-
Preference for vocational school	0.201	0.401	-	-
Preference for other	0.094	0.292	-	-
Intention to enroll at university			0.459	0.498
Average language distance	34.384	26.393	34.627	25.474
School information				
Attend low secondary school: grade 6	0.330	0.470	-	-
Attend low secondary school: grade 7	0.334	0.472	-	-
Attend low secondary school: grade 8	0.336	0.473	-	-
Attend high secondary school: grade 9	-	-	0.311	0.462
Attend high secondary school: grade 10	-	-	0.218	0.413
Attend high secondary school: grade 11	-	-	0.193	0.395
Attend high secondary school: grade 12	-	-	0.153	0.360
Attend high secondary school: grade 13	-	-	0.125	0.330
Math score	6.502	1.114	6.373	1.091
Italian score	6.412	1.325	6.154	1.479
Grade retention	0.182	0.386	0.299	0.458
Discriminatory teacher behavior	0.180	0.384	0.301	0.459
Grade Italian school has started	1.973	1.938	3.022	2.991
Pupils characteristics				
Be a girl	0.463	0.499	0.544	0.498
Born in Italy	0.397	0.489	0.197	0.398
Practice a sport	0.553	0.497	0.462	0.499
Work	0.308	0.462	0.319	0.466
Household characteristics				
siblings_none	0.207	0.405	0.257	0.437
Have a sibling	0.428	0.495	0.426	0.494
Have two or more siblings	0.365	0.482	0.317	0.465
Mother does not work	0.357	0.479	0.344	0.475
Father does not work	0.132	0.338	0.152	0.359
Parental education: no education	0.079	0.269	0.041	0.198
Parental education: primary	0.026	0.158	0.018	0.132
Parental education: low secondary	0.152	0.359	0.159	0.366
Parental education: high secondary	0.258	0.438	0.448	0.497
Parental education: degree	0.162	0.368	0.221	0.415
Family's wealth index	-0.021	0.967	0.055	0.943
Location				
Living in a big city	0.202	0.401	0.317	0.465
Region: Piemonte	0.070	0.255	0.077	0.267
Region:Valle D'Aosta	0.013	0.112	0.008	0.090
Region: Lombardia	0.074	0.263	0.090	0.287
Region: Veneto	0.084	0.278	0.100	0.300
Region: Friuli-Venezia Giulia	0.051	0.220	0.053	0.224
Region: Liguria	0.054	0.227	0.070	0.255
Region: Emilia Romagna	0.080	0.271	0.119	0.324
Region: Toscana	0.069	0.254	0.067	0.250

Region: Umbria	0.046	0.208	0.043	0.203
Region: Marche	0.048	0.214	0.039	0.193
Region: Lazio	0.056	0.230	0.064	0.244
Region: Abruzzo	0.048	0.213	0.037	0.188
Region: Molise	0.009	0.095	0.008	0.088
Region: Campania	0.070	0.256	0.041	0.198
Region: Puglia	0.033	0.178	0.037	0.189
Region: Basilicata	0.011	0.105	0.009	0.096
Region: Calabria	0.038	0.192	0.035	0.184
Region: Sicilia	0.053	0.224	0.042	0.199
Region: Sardegna	0.017	0.129	0.014	0.118
Region: Provincia autonoma di Bolzano	0.035	0.183	0.017	0.131
Region: Provincia autonoma di Trento	0.040	0.196	0.029	0.168
Observations	12945		14790	

Table 2 – Ethnic identity index

	Low secondary		High secondary	
	mean	sd	mean	sd
Language	0.161	0.177	0.114	0.156
How well do you understand Italian? (very well=0)	0.114	0.185	0.073	0.153
How well do you speak Italian? (very well=0)	0.150	0.198	0.111	0.173
How well do you read Italian? (very well=0)	0.180	0.213	0.112	0.178
How well do you write Italian? (very well=0)	0.201	0.218	0.160	0.197
Culture	0.295	0.336	0.321	0.336
Italian used for thinking (yes=0)	0.349	0.477	0.337	0.473
watching Italian TV (yes=0)	0.231	0.320	0.300	0.341
Ethnic self-identification				
Do you identify yourself as Italian? (yes=0)	0.563	0.496	0.639	0.480
Migration history				
Wish to live in Italy when you get old (yes=0)	0.642	0.479	0.720	0.449
Social Network				
Meeting Italian schoolmates outside school (yes=0)	0.265	0.378	0.263	0.373
Meeting Italian friends outside school* (yes=0)	0.298	0.457	0.316	0.465
Ethnic identity index	0.232	0.422	0.211	0.408
	0.296	0.187	0.298	0.181

Notes: * schoolmates are excluded

Table 3 –summary table for aspiration to high secondary school

<i>Etnosizer</i>	Academically oriented	Technical	Vocational	Other
FULL SAMPLE				
LPM	-0.211*** (0.023)	-0.098*** (0.022)	0.002 (0.020)	-0.010 (0.015)
3SLS	-0.249*** (0.088)	-0.006 (0.085)	-0.077 (0.078)	-0.088 (0.056)
BOYS				
LPM	-0.159*** (0.029)	-0.117*** (0.031)	-0.004 (0.028)	-0.013 (0.021)
3SLS	-0.234* (0.130)	0.034 (0.138)	-0.025 (0.123)	-0.220** (0.086)
GIRLS				
LPM	-0.264*** (0.035)	-0.077*** (0.030)	0.005 (0.029)	-0.010 (0.022)
3SLS	-0.272** (0.120)	-0.038 (0.101)	-0.119 (0.098)	0.031 (0.074)

Table 4 –summary table for aspiration for university

Model	FULL SAMPLE	BOYS	GIRLS
LPM	-0.232*** (0.022)	-0.180*** (0.032)	-0.265*** (0.031)
3SLS	-0.313** (0.088)	-0.247* (0.130)	-0.367*** (0.123)

APPENDIX

Table A1 – Aspiration for secondary school (3SLS)

VARIABLES	Ethnic identity index	Academically oriented	Technical	Vocational	Other
Ethnic identity index		-0.249*** (0.088)	-0.006 (0.085)	-0.077 (0.078)	-0.088 (0.056)
Attend secondary school: grade 7	-0.011*** (0.004)	-0.004 (0.010)	0.031*** (0.009)	0.033*** (0.008)	-0.002 (0.007)
Attend secondary school: grade 8	-0.014*** (0.004)	-0.132*** (0.010)	0.140*** (0.010)	0.173*** (0.009)	-0.062*** (0.006)
Math score	-0.025*** (0.002)	0.054*** (0.005)	-0.013*** (0.005)	-0.016*** (0.005)	-0.007** (0.003)
Italian score	0.002* (0.001)	0.040*** (0.004)	0.008** (0.004)	-0.033*** (0.003)	-0.008*** (0.002)
Grade retention	0.017*** (0.004)	-0.076*** (0.010)	-0.037*** (0.010)	0.078*** (0.011)	0.007 (0.007)
Discriminatory teacher behavior	0.005 (0.004)	-0.022** (0.010)	-0.008 (0.010)	0.009 (0.009)	0.002 (0.007)
Grade Italian school has started	0.021*** (0.001)	0.003 (0.003)	-0.003 (0.003)	0.004 (0.003)	0.002 (0.002)
Be a girl	0.017*** (0.003)	0.136*** (0.008)	-0.086*** (0.008)	-0.021*** (0.007)	0.005 (0.006)
Born in Italy	-0.028*** (0.003)	0.046*** (0.009)	-0.015* (0.009)	-0.025*** (0.008)	-0.007 (0.006)
Practice a sport	-0.022*** (0.003)	0.039*** (0.008)	0.005 (0.008)	-0.007 (0.008)	-0.010* (0.006)
Work	0.020*** (0.003)	-0.044*** (0.009)	0.005 (0.008)	0.016** (0.008)	0.005 (0.006)
Have a sibling	0.001 (0.004)	0.010 (0.011)	0.003 (0.010)	0.012 (0.009)	-0.014** (0.007)
Have two or more siblings	0.001 (0.004)	-0.018 (0.011)	-0.005 (0.011)	0.023** (0.010)	-0.002 (0.008)
Mother does not work	-0.010*** (0.003)	-0.000 (0.008)	0.012 (0.008)	-0.010 (0.008)	-0.001 (0.006)
Father does not work	-0.013*** (0.004)	-0.018 (0.011)	0.009 (0.011)	-0.006 (0.010)	0.003 (0.008)
Parental education: primary	0.017 (0.011)	-0.013 (0.025)	0.052** (0.026)	0.011 (0.027)	-0.038** (0.018)
Parental education: low secondary	-0.021*** (0.007)	0.016 (0.016)	0.049*** (0.016)	0.015 (0.016)	-0.031** (0.012)
Parental education: high secondary	-0.026*** (0.006)	0.053*** (0.016)	0.071*** (0.015)	-0.028* (0.015)	-0.032*** (0.012)
Parental education: degree	-0.033*** (0.007)	0.103*** (0.018)	0.060*** (0.017)	-0.069*** (0.016)	-0.029** (0.012)
Family's wealth index	-0.029*** (0.002)	0.011** (0.005)	0.013*** (0.005)	-0.009* (0.005)	-0.005 (0.004)
Average language distance	0.002*** (0.000)				
Living in a big city	0.025*** (0.004)	0.046*** (0.011)	0.031*** (0.011)	-0.049*** (0.010)	-0.017** (0.007)
Region	-0.060*** (0.012)	0.057 (0.038)	-0.004 (0.035)	-0.036 (0.033)	0.015 (0.028)
Constant	0.385*** (0.014)	-0.282*** (0.053)	0.204*** (0.050)	0.497*** (0.046)	0.285*** (0.035)
Observations	12,945	12,945	12,945	12,945	12,945

Table A2 – Aspiration for secondary school (3SLS) - BOYS

VARIABLES	Ethnic identity index	Academically oriented school	Technical school	Vocational school	Other
Ethnic identity index		-0.234*	0.034	-0.025	-0.220**
		(0.130)	(0.138)	(0.123)	(0.086)
Attend secondary school: grade 7	-0.016***	-0.014	0.045***	0.040***	-0.004
	(0.005)	(0.013)	(0.013)	(0.011)	(0.009)
Attend secondary school: grade 8	-0.023***	-0.160***	0.166***	0.190***	-0.061***
	(0.005)	(0.013)	(0.014)	(0.013)	(0.009)
Math score	-0.022***	0.043***	-0.011	-0.003	-0.008
	(0.002)	(0.007)	(0.007)	(0.006)	(0.005)
Italian score	0.002	0.043***	0.011**	-0.037***	-0.009***
	(0.002)	(0.005)	(0.005)	(0.004)	(0.003)
Grade retention	0.018***	-0.063***	-0.047***	0.067***	0.015
	(0.005)	(0.011)	(0.014)	(0.013)	(0.009)
Discriminatory teacher behavior	0.009*	-0.009	-0.016	0.012	-0.009
	(0.005)	(0.013)	(0.014)	(0.013)	(0.009)
Grade Italian school has started	0.020***	0.005	-0.008*	0.001	0.006**
	(0.001)	(0.004)	(0.004)	(0.004)	(0.003)
Born in Italy	-0.025***	0.048***	-0.024*	-0.022**	-0.003
	(0.005)	(0.012)	(0.013)	(0.011)	(0.008)
Practice a sport	-0.017***	0.051***	0.006	0.001	-0.019**
	(0.004)	(0.011)	(0.012)	(0.011)	(0.008)
Work	0.017***	-0.040***	0.003	0.017	0.006
	(0.004)	(0.011)	(0.011)	(0.010)	(0.008)
Have a sibling	0.002	0.010	0.008	0.004	-0.016*
	(0.005)	(0.014)	(0.014)	(0.013)	(0.009)
Have two or more siblings	-0.001	-0.011	0.001	0.015	-0.011
	(0.006)	(0.014)	(0.015)	(0.014)	(0.010)
Mother does not work	-0.013***	-0.012	0.015	0.004	-0.000
	(0.004)	(0.011)	(0.012)	(0.011)	(0.008)
Father does not work	-0.015**	0.000	0.014	-0.011	-0.012
	(0.006)	(0.015)	(0.016)	(0.015)	(0.011)
Parental education: primary	0.023	-0.027	0.044	-0.004	-0.016
	(0.015)	(0.031)	(0.036)	(0.035)	(0.023)
Parental education: low secondary	-0.011	-0.013	0.046**	0.023	-0.011
	(0.009)	(0.021)	(0.022)	(0.022)	(0.016)
Parental education: high secondary	-0.015*	0.033*	0.076***	-0.012	-0.021
	(0.008)	(0.020)	(0.021)	(0.020)	(0.015)
Parental education: degree	-0.025***	0.086***	0.069***	-0.066***	-0.020
	(0.009)	(0.023)	(0.023)	(0.021)	(0.016)
Family's wealth index	-0.031***	-0.003	0.021***	-0.001	-0.014**
	(0.002)	(0.007)	(0.007)	(0.007)	(0.005)
Average language distance	0.002***				
	(0.000)				
Living in a big city	0.024***	0.042***	0.041***	-0.053***	-0.016
	(0.006)	(0.015)	(0.016)	(0.013)	(0.010)
Region	yes	yes	yes	yes	yes
Constant	0.377***	-0.230***	0.138*	0.427***	0.320***

	(0.019)	(0.072)	(0.075)	(0.068)	(0.050)
Observations	6,948	6,948	6,948	6,948	6,948

Table A3 - Aspiration for secondary school (3SLS) – GIRLS

VARIABLES	Ethnic identity index	Academically oriented school	Technical school	Vocational school	Other
Ethnic identity index		-0.272** (0.120)	-0.038 (0.101)	-0.119 (0.098)	0.031 (0.074)
Attend secondary school: grade 7	-0.005 (0.005)	0.009 (0.015)	0.016 (0.012)	0.022** (0.011)	-0.001 (0.010)
Attend secondary school: grade 8	-0.004 (0.005)	-0.101*** (0.015)	0.110*** (0.013)	0.154*** (0.013)	-0.066*** (0.009)
Math score	-0.029*** (0.003)	0.064*** (0.008)	-0.014** (0.007)	-0.030*** (0.007)	-0.006 (0.005)
Italian score	0.003 (0.002)	0.036*** (0.006)	0.004 (0.005)	-0.026*** (0.005)	-0.008** (0.004)
Grade retention	0.018** (0.007)	-0.103*** (0.018)	-0.020 (0.016)	0.099*** (0.018)	-0.005 (0.012)
Discriminatory teacher behavior	-0.001 (0.006)	-0.038** (0.016)	0.003 (0.014)	0.004 (0.013)	0.017* (0.011)
Grate Italian school has started	0.021*** (0.001)	0.002 (0.004)	0.002 (0.004)	0.006 (0.004)	-0.002 (0.003)
Born in Italy	-0.031*** (0.005)	0.043*** (0.014)	-0.003 (0.012)	-0.028*** (0.011)	-0.012 (0.009)
Practice a sport	-0.024*** (0.004)	0.024* (0.013)	0.004 (0.011)	-0.016 (0.010)	0.001 (0.008)
Work	0.024*** (0.005)	-0.050*** (0.014)	0.007 (0.012)	0.015 (0.012)	0.004 (0.009)
Have a sibling	-0.002 (0.006)	0.004 (0.016)	-0.000 (0.014)	0.023* (0.012)	-0.011 (0.010)
Have two or more siblings	0.000 (0.006)	-0.033* (0.018)	-0.009 (0.015)	0.033** (0.014)	0.009 (0.011)
Mother does not work	-0.007 (0.005)	0.014 (0.013)	0.009 (0.011)	-0.026** (0.011)	-0.003 (0.008)
Father does not work	-0.011* (0.006)	-0.037** (0.017)	0.002 (0.015)	0.001 (0.015)	0.019* (0.012)
Parental education: primary	0.010 (0.017)	0.001 (0.042)	0.066* (0.037)	0.031 (0.041)	-0.066** (0.029)
Parental education: low secondary	-0.036*** (0.010)	0.056** (0.026)	0.051** (0.022)	0.005 (0.024)	-0.056*** (0.020)
Parental education: high secondary	-0.040*** (0.009)	0.083*** (0.026)	0.066*** (0.021)	-0.046** (0.023)	-0.047** (0.019)
Parental education: degree	-0.043*** (0.010)	0.129*** (0.028)	0.048** (0.023)	-0.072*** (0.023)	-0.043** (0.020)
Family's wealth index	-0.026*** (0.003)	0.028*** (0.007)	0.005 (0.006)	-0.018** (0.007)	0.003 (0.005)
Average language distance	0.002*** (0.000)				
Living in a big city	0.027*** (0.006)	0.054*** (0.017)	0.018 (0.015)	-0.046*** (0.013)	-0.018* (0.010)
Region	yes	yes	yes	yes	yes

Constant	0.411*** (0.020)	-0.192** (0.079)	0.181*** (0.067)	0.544*** (0.065)	0.262*** (0.050)
Observations	5,997	5,997	5,997	5,997	5,997

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A4 – Aspiration for university

VARIABLES	LPM	IV APPROACH	
		1st stage	2nd stage
Ethnic identity index	-0.232*** (0.022)		-0.313*** (0.088)
Attend secondary school: grade 10	0.018* (0.010)	-0.001 (0.004)	0.018* (0.010)
Attend secondary school: grade 11	0.029*** (0.011)	-0.007* (0.004)	0.029*** (0.011)
Attend secondary school: grade 12	0.060*** (0.012)	-0.026*** (0.004)	0.058*** (0.012)
Attend secondary school: grade 13	0.076*** (0.013)	-0.031*** (0.005)	0.074*** (0.013)
Math score	0.041*** (0.004)	-0.026*** (0.001)	0.038*** (0.005)
Italian score	0.031*** (0.003)	0.004*** (0.001)	0.031*** (0.003)
Grade retention	-0.127*** (0.008)	0.001 (0.003)	-0.127*** (0.008)
Discriminatory teacher behavior	-0.019** (0.008)	0.002 (0.003)	-0.019** (0.008)
Grate Italian school has started	0.001 (0.001)	0.015*** (0.000)	0.002 (0.002)
Be a girl	0.198*** (0.008)	0.013*** (0.003)	0.199*** (0.008)
Born in Italy	0.021** (0.010)	-0.026*** (0.004)	0.020* (0.010)
Practice a sport	0.036*** (0.008)	-0.014*** (0.003)	0.035*** (0.008)
Work	-0.065*** (0.008)	0.014*** (0.003)	-0.064*** (0.008)
Have a sibling	-0.015 (0.010)	-0.005 (0.003)	-0.016 (0.010)
Have two or more siblings	-0.044*** (0.011)	-0.002 (0.004)	-0.043*** (0.011)
Mother does not work	-0.002 (0.008)	-0.004 (0.003)	-0.001 (0.008)
Father does not work	-0.030*** (0.011)	-0.003 (0.004)	-0.031*** (0.011)
Parental education: primary	0.061* (0.032)	-0.003 (0.012)	0.061* (0.032)
Parental education: low secondary	0.055*** (0.019)	-0.033*** (0.008)	0.052*** (0.019)
Parental education: high secondary	0.137*** (0.018)	-0.047*** (0.008)	0.132*** (0.019)
Parental education: degree	0.239*** (0.019)	-0.037*** (0.008)	0.234*** (0.020)
Family's wealth index	0.006 (0.004)	-0.022*** (0.002)	0.004 (0.005)
Living in a big city	0.014 (0.010)	0.023*** (0.003)	0.016 (0.010)
Region	yes	yes	yes
Average language distance		0.002***	

Constant	-0.136***	(0.000)	0.375***	-0.100*
	(0.036)		(0.013)	(0.052)
Observations	14,790		14,790	14,790
R-squared	0.176			0.175

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A5 - Aspiration for university by gender

VARIABLES	LPM		IV APPROACH			
	Boy	Girl	1st stage	2nd stage	1st stage	2nd stage
			Boy	Girl	Boy	Girl
Ethnic identity index	-0.180***	-0.265***		-0.247*		-0.367***
	(0.032)	(0.031)		(0.130)		(0.123)
Attend secondary school: grade 10	0.015	0.020	-0.007	0.014	0.003	0.020
	(0.014)	(0.015)	(0.005)	(0.014)	(0.005)	(0.015)
Attend secondary school: grade 11	0.064***	0.000	-0.013**	0.063***	-0.003	0.000
	(0.016)	(0.015)	(0.006)	(0.016)	(0.005)	(0.015)
Attend secondary school: grade 12	0.098***	0.030*	-0.026***	0.096***	-0.026***	0.028
	(0.018)	(0.017)	(0.006)	(0.018)	(0.006)	(0.017)
Attend secondary school: grade 13	0.095***	0.061***	-0.031***	0.093***	-0.032***	0.058***
	(0.019)	(0.018)	(0.007)	(0.019)	(0.006)	(0.018)
Math score	0.030***	0.049***	-0.021***	0.028***	-0.030***	0.045***
	(0.006)	(0.005)	(0.002)	(0.006)	(0.002)	(0.007)
Italian score	0.035***	0.027***	0.001	0.036***	0.007***	0.028***
	(0.004)	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)
Grade retention	-0.121***	-0.131***	-0.001	-0.121***	0.003	-0.130***
	(0.011)	(0.013)	(0.004)	(0.011)	(0.004)	(0.013)
Discriminatory teacher behavior	-0.020	-0.019*	-0.005	-0.020*	0.009**	-0.019
	(0.012)	(0.011)	(0.004)	(0.012)	(0.004)	(0.012)
Grate Italian school has started	0.003	0.000	0.015***	0.004	0.014***	0.001
	(0.002)	(0.002)	(0.001)	(0.003)	(0.001)	(0.003)
Born in Italy	0.065***	-0.019	-0.030***	0.063***	-0.023***	-0.020
	(0.015)	(0.014)	(0.005)	(0.015)	(0.005)	(0.014)
Practice a sport	0.002	0.068***	-0.012***	0.001	-0.016***	0.065***
	(0.011)	(0.011)	(0.004)	(0.011)	(0.004)	(0.012)
Work	-0.054***	-0.071***	0.010**	-0.054***	0.018***	-0.069***
	(0.011)	(0.012)	(0.004)	(0.011)	(0.004)	(0.013)
Have a sibling	-0.022	-0.009	-0.007	-0.022	-0.003	-0.009
	(0.014)	(0.013)	(0.005)	(0.014)	(0.005)	(0.013)
Have two or more siblings	-0.050***	-0.034**	-0.010*	-0.050***	0.004	-0.032**
	(0.015)	(0.015)	(0.005)	(0.015)	(0.005)	(0.015)
Mother does not work	-0.012	0.008	-0.005	-0.012	-0.004	0.008
	(0.012)	(0.012)	(0.004)	(0.012)	(0.004)	(0.012)
Father does not work	-0.028*	-0.033**	-0.002	-0.029*	-0.004	-0.034**
	(0.015)	(0.015)	(0.006)	(0.015)	(0.005)	(0.015)
Parental education: primary	0.069	0.066	-0.017	0.068	-0.001	0.067
	(0.044)	(0.046)	(0.019)	(0.044)	(0.016)	(0.046)
Parental education: low secondary	0.042*	0.069**	-0.031***	0.039*	-0.040***	0.064**
	(0.023)	(0.031)	(0.011)	(0.023)	(0.012)	(0.032)
Parental education: high secondary	0.115***	0.157***	-0.045***	0.111***	-0.053***	0.149***
	(0.022)	(0.030)	(0.010)	(0.023)	(0.012)	(0.031)
Parental education: degree	0.239***	0.236***	-0.035***	0.235***	-0.043***	0.229***
	(0.024)	(0.031)	(0.011)	(0.025)	(0.012)	(0.032)
Family's wealth index	0.004	0.009	-0.018***	0.002	-0.027***	0.006
	(0.005)	(0.007)	(0.002)	(0.006)	(0.003)	(0.008)
Living in a big city	0.049***	-0.006	0.010*	0.050***	0.034***	-0.003
	(0.015)	(0.013)	(0.005)	(0.015)	(0.005)	(0.014)
Region	0.055	0.147***				
	(0.064)	(0.057)				
lang_dist_av			0.002***		0.002***	
			(0.000)		(0.000)	

Constant	-0.094*	0.028	0.378***	-0.065	0.389***	0.074
	(0.050)	(0.053)	(0.019)	(0.073)	(0.019)	(0.075)
Observations	6,740	8,050	6,740	6,740	8,050	8,050
R-squared	0.138	0.133		0.137		0.132

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1