Econometrica Supplementary Material

SUPPLEMENT TO "GOALS AND GAPS: EDUCATIONAL CAREERS OF IMMIGRANT CHILDREN"

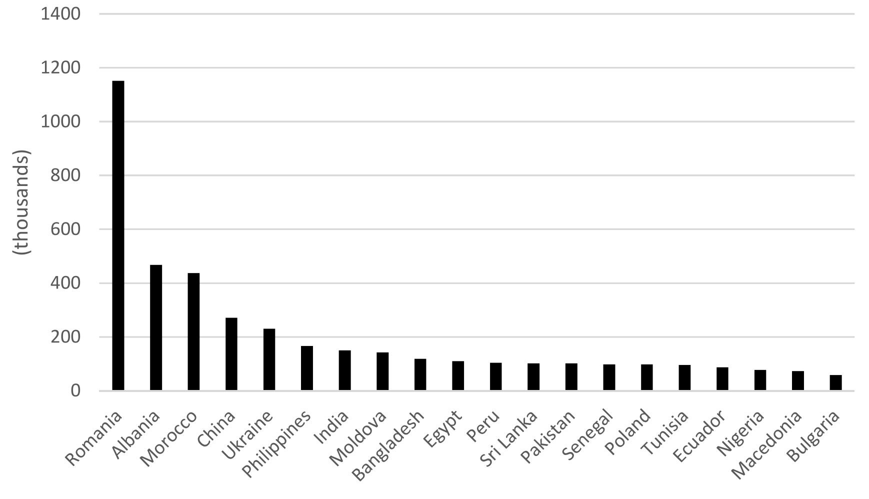
(*Econometrica*, Vol. 90, No. 1, January 2022, 1–29)

MICHELA CARLANA
Harvard Kennedy School, LEAP, IZA, and CEPR

ELIANA LA FERRARA
Bocconi University, IGIER, LEAP, and CEPR

PAOLO PINOTTI
Bocconi University, DONDENA, LEAP, CReAM, and CEPR

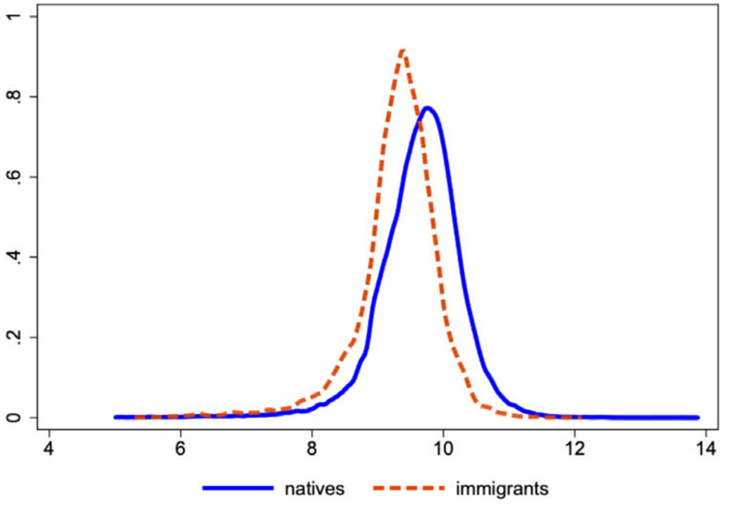
APPENDIX A: APPENDIX FIGURES AND TABLES



Source: ISTAT, "Demografia in Cifre," several years (www.demo.istat.it).

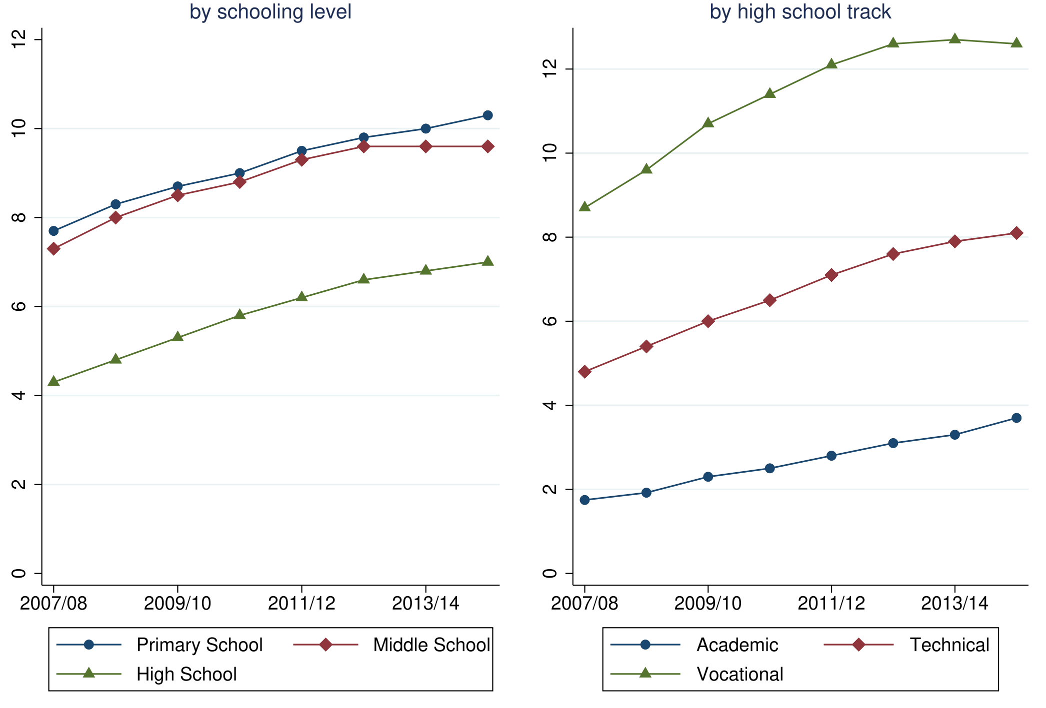
FIGURE A1.—Immigrants in Italy by nationality, 2015.

Michela Carlana: michela_carlana@hks.harvard.edu
Eliana La Ferrara: eliana.laferrara@unibocconi.it
Paolo Pinotti: paolo.pinotti@unibocconi.it



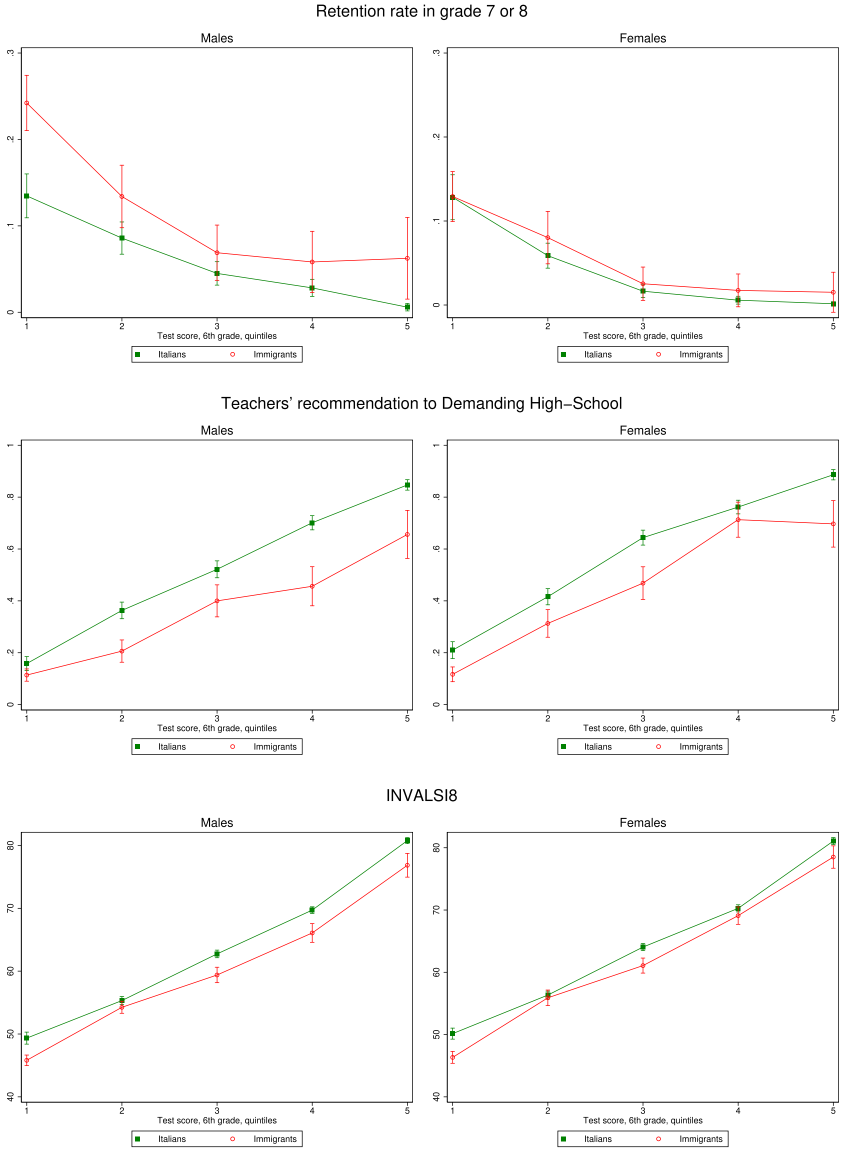
Notes: This graph shows the distribution of (log) disposable income per equivalent adult at constant 2010 prices.
Source: European Union Statistics on Income and Living Conditions (EU-SILC), 2007–2014.

FIGURE A2.—Distribution of (log) income across native and immigrant families in Italy.



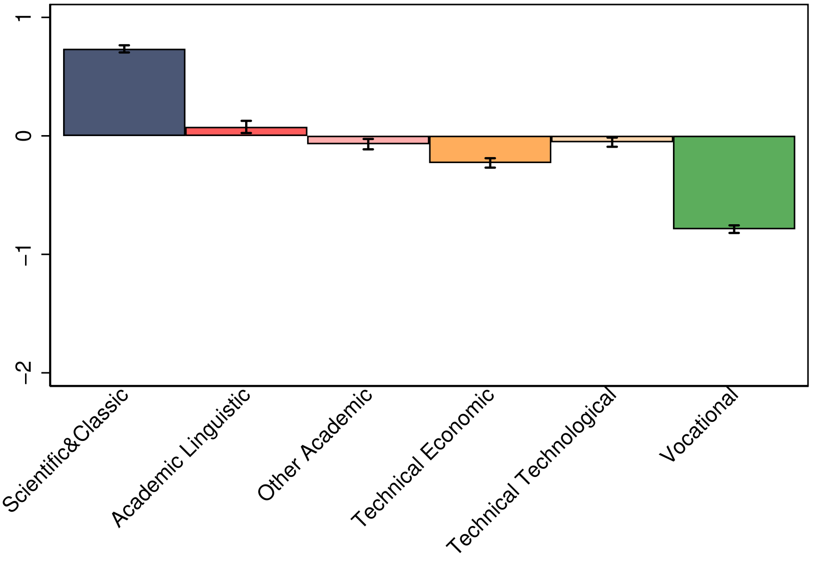
Source: MIUR, "Portale dei dati sulla scuola" (dati.istruzione.it), several years.

FIGURE A3.—Percentage of immigrants over total students in Italy, by schooling level and high school track.



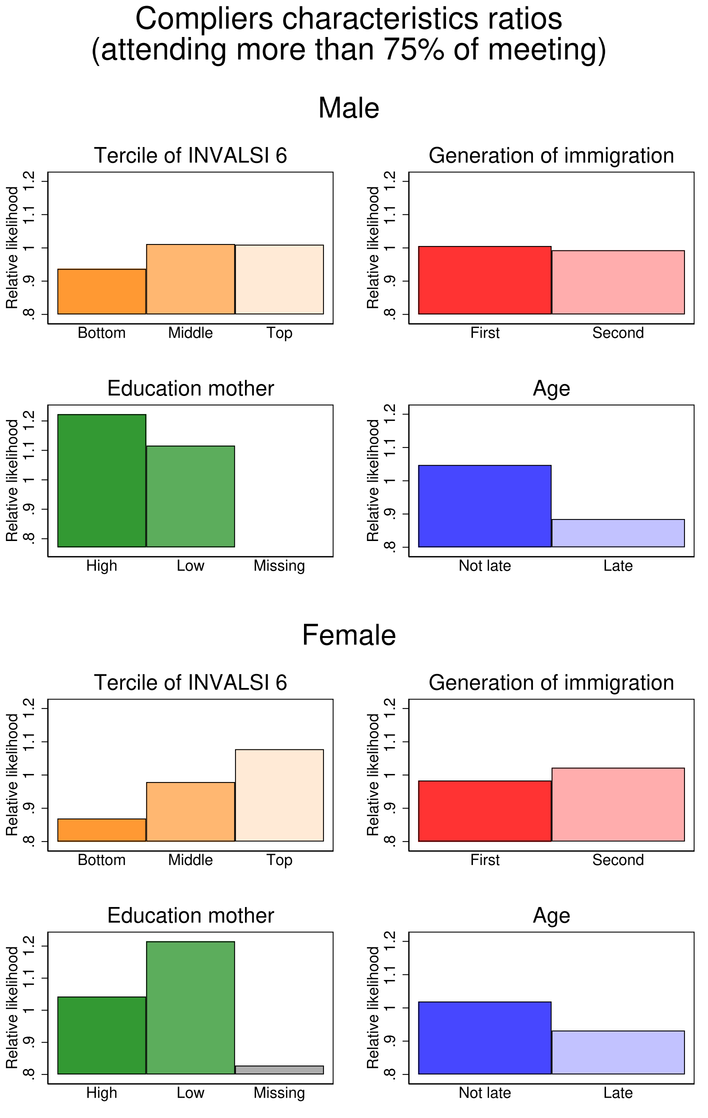
Notes: This figure compares the probability of failing in grade 7 or 8 (Panel A), being recommended by teachers to the high track (Panel B), and the test score in grade 8 (Panel C) between immigrant and native students, by quintiles of performance in the standardized test in grade 6 (INVALSI6). The sample includes all students in the 75 control schools.

FIGURE A4.—Outcomes during middle school, by quintile of standardized test score in grade 6 (INVALS6).



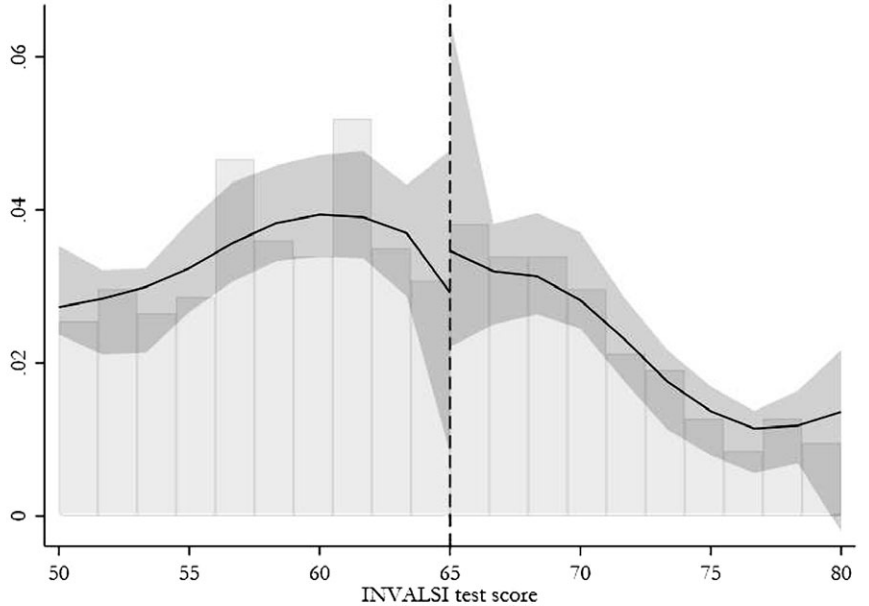
Source: This figure shows the average math test score in grade 8 of students who self-selected into different high school tracks. Other Academic includes social-psychology and arts. It shows a clear ranking in terms of ability between students in the scientific and classical academic track, other academic and technical tracks, and vocational track.

FIGURE A5.—Average test score in math in grade 8, by high school track.



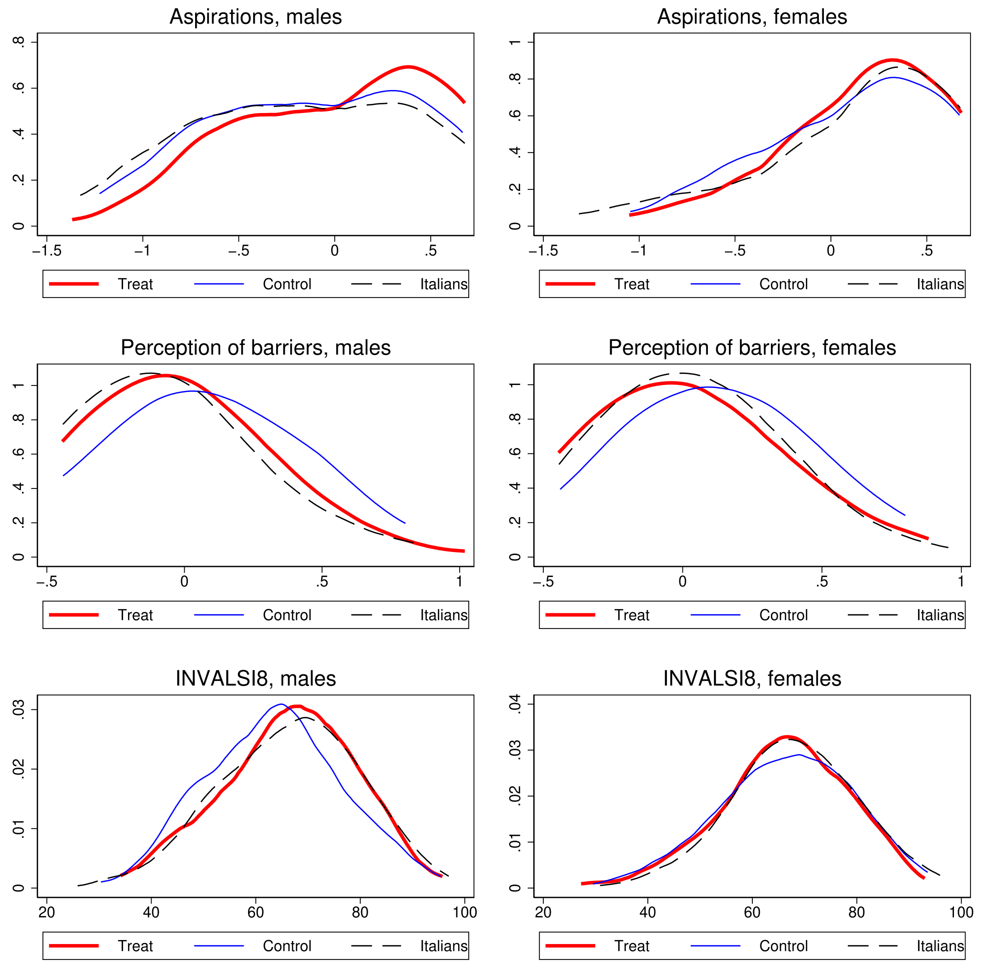
Notes: This figure shows compliers' characteristics ratios, that is, the ratio of the first stage for student of a specific type (e.g., female/male) to the overall first stage. The instrument is the assignment to EOP and the endogenous variable is the probability of attending at least 75% of meetings. The figure illustrates the relative likelihood of compliers' gender, generation of immigration, tercile of INVALSI 6, and age.

FIGURE A6.—Compliers' characteristics.



Notes: This figure shows the McCrary density test to check the potential manipulation of the running variable in the regression discontinuity design (McCrary (2008)).

FIGURE A7.—McCrary test for the manipulation of the running variable in the RDD.



Notes: These graphs show the distribution of aspirations, perception of barriers, and INVALSI8 across treated students, control students, and a group of Italian students that are comparable in terms of schooling ability. Specifically, we match each immigrant student with a native student obtaining exactly the same score in INVALSI6.

FIGURE A8.—Distribution of cognitive and personality skills across treated, controls, and comparable native students.

TABLE AI

EDUCATIONAL AND OCCUPATIONAL OUTCOMES FOUR YEARS AFTER GRADUATION, BY HIGH SCHOOL TRACK.

	All Students		Males		Females	
	High Track	Low Track	High Track	Low Track	High Track	Low Track
			Panel A: Native students			
Percentage of graduates by track	85.5	14.5	84.4	15.6	85.5	13.5
Ever enrolled into university	0.704	0.205	0.650	0.158	0.754	0.256
	(0.003)	(0.004)	(0.006)	(0.006)	(0.004)	(0.007)
Dropout rate in university	0.118	0.306	0.145	0.353	0.097	0.274
	(0.003)	(0.011)	(0.005)	(0.018)	(0.003)	(0.014)
Not in Employment, Education or Training (NEET)	0.199	0.291	0.189	0.264	0.208	0.320
	(0.003)	(0.005)	(0.005)	(0.007)	(0.004)	(0.007)
Regretting high school choice	0.267	0.318	0.266	0.304	0.269	0.333
	(0.003)	(0.005)	(0.005)	(0.007)	(0.004)	(0.008)
			Panel B: Immigrant students			
Percentage of graduates by track	62.8	37.2	57.4	42.6	66.3	33.7
Ever enrolled into university	0.655	0.291	0.686	0.172	0.637	0.390
	(0.03)	(0.022)	(0.049)	(0.027)	(0.037)	(0.032)
Dropout rate in university	0.150	0.257	0.231	0.307	0.100	0.238
	(0.028)	(0.037)	(0.06)	(0.074)	(0.028)	(0.043)
Not in Employment, Education or Training (NEET)	0.264	0.294	0.237	0.238	0.280	0.340
	(0.027)	(0.022)	(0.045)	(0.03)	(0.035)	(0.031)
Regretting about high school choice	0.269	0.331	0.325	0.290	0.238	0.365
	(0.028)	(0.023)	(0.049)	(0.032)	(0.033)	(0.032)

Note: This table shows average educational and occupational outcomes of students graduating from high school in year 2011 by gender and high school track; separate figures for native and immigrant students are presented in Panel A and B, respectively. Standard errors are reported in parentheses.

TABLE AII

IMMIGRANT STUDENTS' PROBABILITY OF CHOOSING THE HIGH TRACK, CONTROLLING FOR SOCIOECONOMIC BACKGROUND.

	Dependent Variable: Choosing the High Track					
	(1)	(2)	(3)	(4)	(5)	(6)
	Males			Females		
Immigrant	−0.087	−0.069	−0.059	−0.046	−0.041	−0.023
	(0.029)	(0.029)	(0.029)	(0.029)	(0.028)	(0.029)
Low-educated Mother		−0.131	−0.119		−0.149	−0.123
		(0.029)	(0.032)		(0.026)	(0.026)
Mid-educated Mother		−0.016	−0.011		−0.051	−0.042
		(0.018)	(0.019)		(0.016)	(0.017)
Low-educated Father		−0.113	−0.094		−0.089	−0.069
		(0.029)	(0.032)		(0.026)	(0.025)
Mid-educated Father		−0.018	−0.012		−0.007	−0.004
		(0.018)	(0.018)		(0.019)	(0.019)
Mother bluecollar			−0.038			−0.051
			(0.022)			(0.025)
Mother home/unemployed			−0.012			−0.026
			(0.017)			(0.018)
Father bluecollar			−0.036			−0.051
			(0.022)			(0.024)
Father home/unemployed			−0.044			−0.110
			(0.032)			(0.038)
Constant	0.660	0.798	0.806	0.722	0.868	0.887
	(0.023)	(0.029)	(0.029)	(0.022)	(0.024)	(0.025)
Observations	3923	3923	3923	3809	3809	3809
R-squared	0.220	0.255	0.264	0.214	0.240	0.247

Note: This table shows how immigrant status influences the probability of choosing the high track. The dependent variable is a dummy equal to 1 for students choosing the high track. The main explanatory variable is a dummy equal to 1 for immigrant students. The sample includes all students in control schools. All regressions control in addition for a second-degree polynomial of test score in grade 6 (INVALSI6), a dummy for first-generation immigrants, and province fixed effects. Standard errors clustered by school are reported in parentheses.

TABLE AIII

MISSING INFORMATION ON PARENTS' EDUCATION AND LOCAL SOCIOECONOMIC CONDITIONS.

	(1)	(2)	(3)	(4)
	Mother's Education Missing		Father's Education Missing	
	Panel A: Individual-level estimates			
Employment rate	−3.526		−2.874	
	(1.512)		(1.417)	
Unemployment rate		5.160		4.347
		(2.391)		(2.241)
Constant	3.428	−0.116	2.844	−0.054
	(1.391)	(0.146)	(1.304)	(0.136)
Observations	1217	1217	1217	1217
R-squared	0.050	0.049	0.054	0.053
	Panel B: Area-level estimates, weighted by area population			
Employment rate	−5.899		−5.151	
	(1.657)		(1.506)	
Unemployment rate		8.871		7.833
		(2.646)		(2.383)
Constant	5.634	−0.315	4.965	−0.235
	(1.526)	(0.161)	(1.387)	(0.144)
Observations	87	87	87	87
R-squared	0.183	0.179	0.194	0.192

Note: This table shows the relationship between accuracy of information on parents' education in INVALSI registries and local socioeconomic conditions. The dependent variables are binary indicators equal to 1 when mother's and/or father's education (columns 1–2 and 3–4, respectively) are not reported in INVALSI registries. The main explanatory variables of interest are the employment and unemployment rates across 87 Census tracts. Regressions in Panel A are estimated across individuals, whereas regressions in Panel B are estimated across Census tracts weighted by population. Province fixed effects are included in all regressions. In the individual-level regressions in Panel A, standard errors are clustered by Census tract.

TABLE AIV

THE EFFECT OF EOP ON EDUCATIONAL CHOICES, DIFFERENCE-IN-DISCONTINUITIES AT THE ELIGIBILITY CUTOFF BETWEEN TREATED AND CONTROL SCHOOLS.

	(1)	(2)	(3)	(4)	(5)	(6)
	Male Students			Female Students		
EOP X top 10	0.144	0.186	0.184	−0.203	−0.202	−0.173
	(0.075)	(0.114)	(0.081)	(0.127)	(0.148)	(0.125)
EOP	−0.041	−0.110		0.213	0.204	
	(0.066)	(0.093)		(0.124)	(0.137)	
top 10	0.151	−0.009	−0.022	0.359	0.218	0.261
	(0.068)	(0.107)	(0.096)	(0.142)	(0.122)	(0.133)
rank	−0.011	−0.003	−0.002	−0.008	0.001	0.001
	(0.004)	(0.004)	(0.005)	(0.006)	(0.005)	(0.007)
top 10 X rank	−0.017	−0.002	−0.002	−0.023	−0.009	−0.014
	(0.007)	(0.011)	(0.011)	(0.007)	(0.010)	(0.008)
EOP X rank	0.001	0.007	0.007	−0.008	−0.007	−0.006
	(0.004)	(0.005)	(0.004)	(0.006)	(0.007)	(0.006)
EOP X top 10 X rank	−0.003	−0.006	−0.008	0.008	0.009	0.009
	(0.009)	(0.015)	(0.010)	(0.008)	(0.013)	(0.008)
Observations	1320	1290	1318	1274	1254	1270
Additional covariates	NO	YES	YES	NO	YES	YES
School FE	NO	NO	YES	NO	NO	YES
R-squared	0.097	0.157	0.255	0.144	0.205	0.286

Note: This table shows the effect of EOP on educational choices exploiting the fact that only the 10 immigrant students with the highest INVALSI6 within each school were eligible for the program. The dependent variable is a dummy equal to 1 for students choosing the high track (academic or technical schools) and equal to zero otherwise. Top 10 is a dummy for students potentially eligible within each school, EOP is a dummy for schools actually participating to the program, and rank is a position in the rank. Therefore, the coefficient of top 10 estimates the average discontinuity in the probability of choosing the high track between eligible and non-eligible students near the cutoff across all schools, and the coefficient of the interaction EOP X top 10 estimates the differential discontinuity in treated schools (i.e., the "difference-in-discontinuities"). Columns (2)–(3) and (5)–(6) control in addition for a squared polynomial in INVALSI6, a dummy for first-generation immigrants, and province fixed effects, and columns (3) and (6) also include school fixed effects. Standard errors clustered by rank in all specifications and, in addition, by school in columns (3) and (6) are reported in parentheses.

TABLE AV

THE EFFECT OF COMPLETING THE QUESTIONNAIRE ON SOFT SKILLS IN CONTROL SCHOOLS.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Demanding High-School		Grade Retention		Std Test Score Grade 8	
School Questionnaire	−0.016	−0.004	0.008	0.007	−0.061	−0.085
	(0.039)	(0.057)	(0.021)	(0.033)	(0.100)	(0.125)
Female X School Questionnaire		−0.029		0.004		0.035
		(0.069)		(0.036)		(0.140)
Female		0.167		−0.058		0.168
		(0.050)		(0.025)		(0.111)
Constant	0.759	0.676	0.052	0.081	0.004	−0.081
	(0.027)	(0.042)	(0.014)	(0.023)	(0.069)	(0.087)
Observations	620	620	620	620	552	552
R-squared	0.000	0.031	0.000	0.015	0.001	0.015

Note: This table tests whether control students in schools selected for the questionnaire differ in their high school choice, grade retention, and performance in the standardized test score from students in the control schools not selected for the questionnaire. The dependent variable is a dummy equal to 1 for students choosing the high track in columns 1 and 2, a dummy equal to 1 if students are retained in grade 7 or 8 in columns 3 and 4, and the standardized value for the test score in columns 5 and 6. Standard errors clustered by school are reported in parentheses.

TABLE AVI

PRINCIPAL COMPONENT ANALYSIS, FACTOR LOADINGS.

	Loadings	Std. Err.
First principal component: Aspirations		
Goal University	1	–
Self-efficacy University	1.649	0.050
Self-efficacy White collar	0.753	0.030
Self-efficacy Manager	0.631	0.030
Second principal component: Perception of barriers		
Barriers economic	1	–
Barriers family ideas	1.339	0.087
Barriers prejudice	0.837	0.063
Barriers family plans and marriage	1.001	0.074
Barriers self-esteem	0.947	0.072

Note: This table shows estimated factor loadings for the two principal components extracted from psychological measures; Satorra–Bentler robust standard errors are also presented. Measurements are categorical variables in a scale from 1 to 4. "Goal University" is the answer to the following question: *Thinking about your future, do you want to achieve a university degree?*. The measurements related to "Self-efficacy" are the answers to the following questions: *Independently from your educational aim but thinking about your abilities, do you think you could get a university degree/white collar job/managerial job?*. The measurements related to "Barriers" are the answers to the following questions: *Do you think the following barriers could be an obstacle in the achievement of your educational aims? Economic resources/The needs and ideas of your family/Racial prejudice/Family plans (children, marriage)/Not feeling good enough*.

TABLE AVII

INITIAL VERSUS WORKING SAMPLE.

	Treated	Controls	Difference	*P*-Value	Std. Difference
Initial sample	700	751			
Fraction missing match MIUR-INVALSI	0.043	0.053	−0.010	[0.72]	−0.049
Number of students with available MIUR-INVALSI	670	711			
Fraction dropped between INVALSI6 and start of EOP	0.109	0.128	−0.019	[0.51]	−0.059
Final sample	597	620			

Note: This table shows the sample size of treated and control students in our sample. *p*-values for difference in means are reported in parentheses. The last column also reports the standardized difference between group averages.

TABLE AVIII

BALANCE TABLE OF THE NATIONALITY OF TREATED AND CONTROL STUDENTS.

	(1)	(2)	(3)	(4)	
Variable	Treated	Control	Diff.	*p*-Value	Std. Difference
Albania	0.090	0.104	0.014	[0.554]	−0.047
Romania	0.192	0.245	0.053	[0.164]	−0.129
Morocco	0.085	0.082	−0.003	[0.890]	0.011
Philippines	0.106	0.069	−0.038	[0.155]	0.131
Peru	0.058	0.057	−0.001	[0.952]	0.004
Ecuador	0.052	0.062	0.010	[0.591]	−0.043
China	0.056	0.054	−0.003	[0.908]	0.009
Observations	620	597			

Note: This table shows the most common nationality of treated and control students in our sample. *p*-values for difference in means are reported in square brackets. The last column also reports the standardized difference between group averages.

TABLE AIX

EOP EFFECT ON TRACK CHOICE USING AN ORDERED PROBIT.

	Dependent Variable: Track Choice					
	(1)	(2)	(3)	(4)	(5)	(6)
	All Immigrants		Male Immigrants		Female Immigrants	
Top Academic	0.029	0.025	0.062	0.053	−0.011	−0.005
	(0.026)	(0.021)	(0.030)	(0.026)	(0.036)	(0.030)
Other academic/Technical	−0.002	−0.001	0.009	0.009	0.003	0.001
	(0.002)	(0.001)	(0.007)	(0.006)	(0.009)	(0.007)
Vocational	−0.028	−0.024	−0.071	−0.062	0.008	0.004
	(0.024)	(0.020)	(0.035)	(0.031)	(0.027)	(0.023)
Mean Control Top Academic	0.245	0.245	0.195	0.195	0.294	0.294
Mean Control Other	0.507	0.507	0.482	0.482	0.530	0.530
Mean Control Vocational	0.248	0.248	0.322	0.322	0.176	0.176
Observations	1217	1217	601	601	616	616
Controls	No	Yes	No	Yes	No	Yes

Note: This table shows marginal effects of an ordered probit, considering tracks split in three groups: top academic (classic and scientific), Other academic/Technical, Vocational. Controls include the second-degree polynomial of test score in grade 6 (INVALSI6), a dummy for first-generation immigrants, and province fixed effects.

TABLE AX

THE EFFECT OF EOP ON EDUCATIONAL CHOICES, HETEROGENEITY.

	Dependent Variable: Choosing the High Track		
	(1)	(2)	(3)
Subgroup	EOP Coeff.	p-Value	MHT p-Value
Panel A: Heterogeneity by gender and mother education			
Boys High Edu Mother	0.017	0.725	0.725
Boys Low Edu Mother	0.160	0.046	0.209
Boys Missing Edu Mother	0.119	0.058	0.211
Girls High Edu Mother	0.050	0.169	0.423
Girls Low Edu Mother	0.185	0.009	0.051
Girls Missing Edu Mother	0.040	0.502	0.750
Panel B: Heterogeneity by gender and terciles of initial test score			
Boys Top INVALSI6	0.106	0.052	0.272
Boys Middle INVALSI6	0.067	0.267	0.468
Boys Bottom INVALSI6	0.088	0.189	0.564
Girls Top INVALSI6	0.016	0.648	0.648
Girls Middle INVALSI6	0.085	0.063	0.272
Girls Bottom INVALSI6	0.075	0.250	0.582
Panel C: Heterogeneity by gender and place of birth			
Boys EU	0.103	0.058	0.580
Boys Not EU	0.085	0.049	0.179
Girls EU	0.058	0.103	0.456
Girls Not EU	0.004	0.895	0.895

Note: This table shows the heterogeneity of the effect of EOP on immigrant students' educational choices at the end of middle school. The dependent variable is a dummy equal to 1 for students choosing the high track (academic or technical schools) and equal to zero otherwise. EOP is a dummy equal to 1 for students in middle schools assigned to the treatment group and equal to zero for schools assigned to the control group. Highly educated mother is a dummy equal to 1 for students whose mother has at least a high school diploma. EU is a dummy equal to 1 for immigrants from EU-member countries.

TABLE AXI

RDD EFFECTS.

	(1)	(2)	(3)	(4)
Panel A: Dep. var.	No. CALP Meetings	No. Counseling Meetings	Choice High Track	Grade Retention
RDD Estimate	−5.418	1.006	−0.024	−0.023
	(2.575)	(0.934)	(0.0758)	(0.0250)
Observations	565	565	565	565
Panel B: Dep. var.	Aspirations	Perception of Barriers	INVALSI8	Teachers' Recom.
RDD Estimate	−0.143	−0.020	−1.900	−0.002
	(0.127)	(0.0877)	(2.775)	(0.111)
Observations	404	404	512	565

Note: This table shows the RDD effect of CALP on several outcomes, as reported also in Figure 7. Squared Order Local Polynomial.

TABLE AXII

TREATMENT EFFECT ON SOFT SKILLS (BY SURVEY QUESTION).

ITT on	Coefficient	p-Value	p-Value FWER
Group 1: Aspirations			
Goal University	0.098	0.012	0.072
Self-efficacy University	0.174	0.008	0.067
Self-efficacy Whitecollar	0.120	0.021	0.078
Self-efficacy Manager	0.056	0.342	0.035
Group 2: Perception of environmental barriers			
Barriers economic	−0.101	0.041	0.019
Barriers family ideas	−0.084	0.068	0.224
Barriers prejudice	0.033	0.513	0.536
Barriers family formation and marriage	−0.087	0.076	0.254
Barriers self esteem	−0.233	0.000	0.001

Note: Robust standard errors clustered at school level. All regressions include generation of immigration, province, and squared test score. p-Values are adjusted for multiple hypothesis testing using the free step-down resampling method (Westfall and Young (1993)) to control the family-wise error rate (FWER). Measurements are categorical variables in a scale from 1 to 4. "Goal University" is the answer to the following question: *Thinking about your future, do you want to achieve a university degree?*. The measurements related to "Self-efficacy" are the answers to the following questions: *Independently from your educational aim but thinking about your abilities, do you think you could get a university degree/white collar job/managerial job?*. The measurements related to "Barriers" are the answers to the following questions: *Do you think the following barriers could be an obstacle in the achievement of your educational aims? Economic resources/The needs and ideas of your family/Racial prejudice/Family plans (children, marriage)/Not feeling good enough*.

TABLE AXIII

SPECIFICATION TEST, MALES.

Outcome: Choosing the High Track	(1) Test Statistic	(1) p-Value	(2) Test Statistic	(2) p-Value	(3) Test Statistic	(3) p-Value
Mediating factors: $H_0 : \alpha_1^j = \alpha_0^j$						
Aspirations	2.24	[0.14]	1.81	[0.18]	2.84	[0.10]
Barriers	1.30	[0.26]	1.08	[0.30]	0.24	[0.63]
INVALSI8			0.02	[0.89]	0.36	[0.55]
Teachers' recommendation					0.36	[0.55]
Controls: $H_0 : \beta_1^i = \beta_0^i$						
INVALSI6	0.52	[0.47]	0.31	[0.58]	0.78	[0.38]
INVALSI6 sq.	2.49	[0.12]	2.22	[0.14]	1.93	[0.17]
First generation immigrant	0.91	[0.34]	0.86	[0.39]	0.54	[0.46]
Prov BS	2.35	[0.13]	1.76	[0.19]	0.21	[0.65]
Prov GE	1.20	[0.28]	1.01	[0.32]	1.74	[0.19]
Prov MI	0.08	[0.77]	0.00	[0.96]	0.27	[0.61]
Prov PD	0.36	[0.55]	0.11	[0.75]	0.03	[0.86]
Prov TO	0.95	[0.33]	0.96	[0.33]	0.10	[0.76]
F-test	1.37	[0.21]	1.28	[0.26]	1.35	[0.23]

Note: The first panel tests whether the treatment group regression coefficients in equation (S2) are the same as the control group coefficients: $H_0 : \alpha_1^i = \alpha_0^i$, for each potential channel θ. The second panel tests whether the treatment group regression coefficients are the same as the control group coefficients: $H_0 : \beta_1^i = \beta_0^i$, for each potential control variable $\mathbf{X}$. In column (1), we consider only two mediating factors, that is, aspirations and barriers, while in column (2), we include also the achievement test scores, and in column (3), teachers' track recommendation. All specifications control in addition for a squared polynomial in INVALSI6, a dummy equal to 1 for first-generation immigrants, and province fixed effects.

TABLE AXIV

DECOMPOSITION OF THE EFFECT OF EOP ON HIGH SCHOOL CHOICE, MALE STUDENTS (GELBACH (2016)).

	(1) Explained	(1) p-Value	(2) Explained	(2) p-Value	(3) Explained	(3) p-Value
Aspiration	0.0354	[0.034]	0.0291	[0.042]	0.0264	[0.046]
Barriers	0.0046	[0.602]	0.0067	[0.446]	0.0042	[0.620]
Cognitive skills			0.0178	[0.085]	0.0105	[0.190]
Teachers' recommendation					0.0385	[0.011]
Total explained	0.0400	[0.032]	0.0537	[0.010]	0.0796	[0.001]
EOP effect on choosing high track	0.091		0.091		0.091	

Note: This table decomposes the effect of EOP between changes in personality skills (aspirations and perception of barriers), increased schooling achievement (as measured by INVALSI8), and teachers' recommendations. The decomposition follows the method devised by Gelbach (2016). All specifications control in addition for a squared polynomial in INVALSI6, a dummy equal to 1 for first-generation immigrants, and province fixed effects.

TABLE AXV

COST-BENEFIT ANALYSIS.

	Scenario 1	Scenario 2	Scenario 3
Parameters			
Discount rate	3%	3%	3%
Tax rate	28%	28%	28%
Higher salary per month (euros)	500	500	650
Lower unemployment probability	4%	4%	6%
Unemployment insurance benefit per month (euros)	1000	1000	1300
Number of beneficiaries	60	125	125
Costs and Benefits			
Total costs (thousand euros)	2177	2177	2177
Higher taxes on wage (thousand euros)	3344	7006	9108
Lower unemployment insurance (thousand euros)	955	2002	3904
Internal Rate of Return	2.8%	6.6%	8.8%

Note: Although the EOP program has potentially strong effects on health and on reduction of crime rates, we present conservative estimates focusing our cost-benefit analysis only on social benefits coming from higher income taxes and public savings on unemployment insurance. In the first scenario, we consider potential benefits only on 10% of students directly treated by EOP. In the second scenario, keeping all other assumptions constant, we consider also the additional spillovers on 5% of classmates of treated students (close to the share who did not fail the school year or decided to attend a more demanding track compared to classmates of control students). In the last scenario, we slightly reduce the unemployment probability; we slightly increase the expected average higher salary per month and the expected unemployment insurance benefit. We use the discount rate of 3% as in the simulation in Heckman, Moon, Pinto, Savelyev, and Yavitz (2010) and a tax rate close to the current one for the income bracket 15,000–28,000 euros.

APPENDIX B: AVERAGE TREATMENT-ON-THE-TREATED EFFECT

While in the main text we focus on intention-to-treat (ITT) effects, in this section we estimate the average treatment-on-the-treated (ATT). Since there is one-sided non-compliance with treatment assignment, the ATT effects on the subset of compliers are larger than the ITT. Appendix Figure B.1 shows that the pattern of meetings attendance is quite heterogeneous, with more than 40 percent of immigrant boys and girls attending at least 87.5 percent of the meetings, another 20 percent attending between 75 and 87.5 percent of the meetings, and the remaining fractions attending less. Interestingly, about 15 percent of the students who were assigned to treatment ended up attending less than 12.5 percent of the meetings. Given this heterogeneity, there is no unambiguous way of defining treatment status. For this reason, in Appendix Table B.I we experiment with three alternative definitions.

In Panel A, we classify as treated all students attending at least one meeting (85 percent of the sample). In Panel B, we restrict the definition to students attending at least 75 percent of the meetings, in accordance with the program guidelines discussed in Section 3. When adopting these definitions, the ATT effects on males range between a 9.4 to 12.5 percentage point increase in enrollment in the high track, and a 4.3 to 5.7 percentage point decrease in grade retention.

In Appendix Figure A6, we characterize compliers with treatment assignment, defining the treatment as attending at least 75 percent of the meetings, by the ratio of the first-stage effect within specific sub-samples to the overall first stage (Angrist, Cohodes, Dynarski, Pathak, and Walters (2016)). Compliers are slightly more likely to be female, equally likely to be first- and second-generation immigrants, and more likely to be in the right grade ('Not late') given their age. The bottom panels of Figure A6 show that while female compliers are more likely to be from the top part of the initial ability distribution, male compliers are more likely to be from the bottom part, thus more in need of support.

In Panel C of Table B.I, we measure treatment 'intensity' by the fraction of meetings attended. The corresponding ATT estimate suggests that one standard deviation increase in the number of meetings attended increases enrollment into the high track by 4.2 percentage points and reduces grade retention by 1.9 percentage points for males. All three approaches in Table B.I recover the ATT effect under strong (and untestable) assumptions about the relationship between number of meetings attended and treatment intensity. For this reason, in the main body of the paper we focus on the ITT.

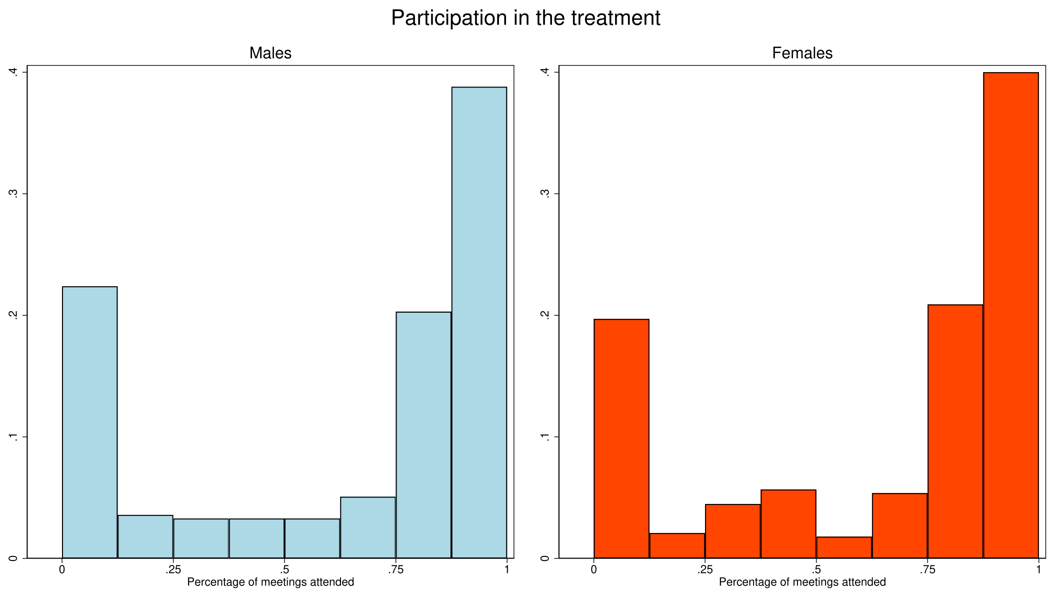


FIGURE B.1.—Meetings attendance of immigrant students assigned to EOP.

TABLE B.I

EFFECTS OF EOP, AVERAGE TREATMENT-ON-THE-TREATED (ATT).

	(1)	(2)	(3)	(4)	(5)	(6)
	Choosing the High Track			Grade Retention		
Dep. Var.:	All	Males	Females	All	Males	Females
Panel A: Treatment = 1 if attended at least one meeting						
ATT	0.051	0.094	0.010	−0.015	−0.043	0.011
	(0.028)	(0.041)	(0.036)	(0.018)	(0.024)	(0.022)
Constant	0.682	0.648	0.720	0.073	0.098	0.047
	(0.038)	(0.049)	(0.049)	(0.028)	(0.028)	(0.043)
Panel B: Treatment = 1 if attended at least 75% of meetings						
ATT	0.067	0.125	0.013	−0.020	−0.057	0.014
	(0.037)	(0.054)	(0.047)	(0.024)	(0.032)	(0.029)
Constant	0.679	0.641	0.720	0.074	0.101	0.047
	(0.038)	(0.050)	(0.049)	(0.027)	(0.028)	(0.043)
Panel C: Treatment = fraction of meetings attended						
ATT	0.064	0.119	0.013	−0.019	−0.055	0.014
	(0.036)	(0.052)	(0.045)	(0.023)	(0.031)	(0.028)
Constant	0.681	0.645	0.720	0.074	0.100	0.048
	(0.038)	(0.049)	(0.049)	(0.027)	(0.028)	(0.043)
Observations	1.217	601	616	1.217	601	616

Note: This table shows the average-treatment-on-the-treated effect of EOP on immigrant students' educational choices (columns 1–3) and grade retention during middle school (columns 4–6). The ATT is computed as the ratio of the reduced form effect of EOP on such outcomes and the first-stage effect on three alternative measures of compliance with treatment assignment: attending at least one meeting (Panel A), attending at least 75% of meetings (Panel B), and fraction of meetings attended (Panel C). All specifications control for a squared polynomial in INVALSI6, a dummy equal to 1 for first-generation immigrants, and province fixed effects. Standard errors clustered by school are reported in parentheses.

APPENDIX C: HETEROGENEOUS TREATMENT EFFECTS USING CAUSAL FOREST

The purpose of the causal forest approach is not to test whether a particular characteristic (e.g., gender) is associated with significantly different treatment effects, but to predict the heterogeneity in the causal treatment effects. The method adapts the classification and regression tree (CART) methodology from the recent literature on supervised machine learning to the problem of predicting treatment effects rather than outcomes. More precisely, the causal forest allows to estimate the Conditional Average Treatment Effect (CATE), defined as $E(Y_{1i} - Y_{0i}|X_i = x)$, where Y is the outcome of interest and X is a vector of observable baseline characteristics.

We build upon the recent literature of machine learning techniques in the context of randomized control trials and heterogeneous treatment effects (Davis and Heller (2017), Bertrand, Crépon, Marguerie, and Premand (2017)). We use the causal forest algorithm (Athey and Imbens (2016), Wager and Athey (2018), Athey et al. (2019)) and the following procedure. First, we train a causal forest by building 100,000 trees and setting the minimum number of treatment and control observations allowed in a leaf to the default value (5).[1] We use the "honest" approach (Athey and Imbens (2016), Wager and Athey (2018)): we split the training sample in two parts, and we use half of the observations for growing the tree and half of the observations for estimating the treatment effect within each leaf of the tree. We include in the causal forest the following baseline characteristics: gender (dummy variable), squared polynomial of baseline test score (continuous variable), generation of immigration (dummy variable), school province (categorical variable with five values: Milan, Turin, Brescia, Padua, Genoa), parents' education (categorical variable with four values: university, high school diploma, less than high school diploma, missing) and occupation (categorical variable with five values: high-level occupation, working class, unemployed, housemaker, missing), area of citizenship (categorical variable with four values: Latin America, East Europe, Africa, Asia), and, crucially given the design of our experiment, we include the clusters at school level. Second, we calculate the out-of-bag predicted CATE and its variance estimates.[2]

We use the predictions on the expected treatment effect on the high school track choice for each individual, given the covariates, to investigate the treatment heterogeneity in our data. We divide the sample in two groups, considering students in the top 50% and bottom 50% of the predictions. Table C.I reports the balancing test for the Conditional Average Treatment Effect (CATE) and provides the p-value adjusted for multiple hypothesis testing. Overall, the results on gender differences are consistent with the main analysis: girls are over-represented among students with lower CATE. However, additional interesting results emerge.

Students whose parents have white collar jobs or high levels of education are more likely to have a low predicted CATE. Furthermore, interestingly, more than 50% of the students with high predicted CATE have missing values for their parental education and

[1]We use the `causal_forest` command in `R` of the package `grf` (generalized random forest). As suggested by Athey et al. (2019), the only substantial difference from the method proposed in Wager and Athey (2018) is that the split is done using a gradient-based loss criterion instead of the exact loss criterion.

[2]Davis and Heller (2017) showed that "honest" approach may lead to overfitting if the CATE is assigned to all observations in the sample, including those used to construct the tree, and they suggested to obtain out of sample predictions by further splitting the sample and running the causal forest of 20% of the observations. Given our small sample and the purpose of our exercise, we provide predictions for an observation in the original data set (at X_i) using only trees that did not use the ith training example. These predictions are not prone to overfitting, as each prediction is only made by learners that did not use the observation for training. In our predictions the 'excess.error' is negligible, with a mean value of 4.21e-07.

occupation, while this share is lower than 20% for students with low predicted CATE. Parental background is asked to parents from schools and then submitted together with the test scores INVALSI. Targeting the children of parents less responsive to school requests and therefore less involved in school activities may increase the effectiveness of EOP. The average test score in grade 6 (INVALSI6) is slightly lower for students with high predicted CATE on track choice (0.263 vs. 0.352 standard deviations), although the difference is not statistically significant when the p-value is adjusted for multiple hypothesis testing. Figure 6 shows the nonlinearities in the CATE considering the deciles of the test score in grade 6 (INVALSI6), gender, and mother education. As suggested by Table C.I, on average boys, students with lower or missing levels of mother education, and students in the central part of the ability distribution[3] benefit the most from participation in EOP. Targeting only boys would have missed a substantially positive impact on these girls.

Finally, Table C.I reports also the descriptive statistics for the school province and the citizenship. There are no substantial differences across provinces in the CATE, suggesting that different psychologists were equally effective in implementing the treatment. However, students from East Europe are more likely to benefit from the intervention (they are 49% of the High Predicted CATE and 34% of the Low Predicted CATE group) and those from Latin America are less likely to benefit from it (they are 20% of the High Predicted CATE and 32% of the Low Predicted CATE group).

[3]The deciles are defined among the 1217 immigrant students in the treatment and control group.

TABLE C.I

CONDITIONAL AVERAGE TREATMENT EFFECT (CATE): CHOICE OF HIGH TRACK.

Variable	(1) High Predicted CATE	(2) Low Predicted CATE	(3) Diff.	(4) MHT *p*-Value
ITT High Track	0.053	0.031	−0.023	
Female	0.447	0.566	0.119	0.001
First Generation	0.586	0.523	−0.063	0.210
INVALSI 6	0.263	0.352	0.089	0.177
Sq. INVALSI 6	0.364	0.742	0.378	0.001
Older age	0.302	0.212	−0.090	0.001
Mother occupation:				
White collar	0.049	0.199	0.150	0.001
Working class	0.151	0.322	0.171	0.001
Homemaker	0.213	0.304	0.091	0.005
Unemplyed	0.062	0.048	−0.015	0.588
Missing	0.524	0.127	−0.397	0.001
Father occupation:				
White collar	0.122	0.275	0.153	0.001
Working class	0.232	0.482	0.250	0.001
Homemaker	0.013	0.012	−0.002	0.805
Unemplyed	0.054	0.071	0.017	0.648
Missing	0.580	0.161	−0.418	0.001
Mother education:				
University	0.053	0.186	0.133	0.001
High school	0.169	0.426	0.257	0.001
Less than High school	0.205	0.243	0.038	0.530
Missing	0.573	0.145	−0.428	0.001
Father education:				
University	0.043	0.156	0.114	0.001
High school	0.149	0.442	0.293	0.001
Less than High school	0.171	0.224	0.053	0.168
Missing	0.637	0.178	−0.459	0.001
Province:				
PD	0.079	0.031	−0.048	0.001
BS	0.164	0.194	0.030	0.674
MI	0.461	0.531	0.070	0.153
TO	0.227	0.179	−0.047	0.257
GE	0.069	0.064	−0.005	0.937
Citizenship:				
Latin America	0.195	0.317	0.122	0.001
Africa	0.166	0.224	0.058	0.104
Asia	0.148	0.123	−0.024	0.684
East Europe	0.489	0.336	−0.154	0.001
Observations	609	608		

Note: The table reports the descriptive statistics of students in the top 50% (column 1) and bottom 50% (column 2) of the predicted Conditional Average Treatment Effect (CATE) on the choice of a demanding high school. The CATE is computed following the procedure explained in Appendix C. Column 3 reports the difference between column 2 and 1. Column 4 shows the p-value of the t-test adjusted for multiple hypothesis testing.

APPENDIX D: METHODOLOGY FOR MEDIATION ANALYSIS

Following Heckman, Pinto, and Savelyev (2013), we decompose the treatment effect on educational choices into experimentally induced changes in the mediating factors listed in Table IV and changes in other (unmeasured) factors. Assume the following linear model for the potential outcome when randomized into the treated ($d = 1$) and into the control group ($d = 0$):

$$Y_d = \tau_d + \sum_{j \in J} \alpha_d^j \theta_d^j + \beta_d \mathbf{X} + \epsilon_d, \quad d \in \{0, 1\}, \tag{S1}$$

where Y is a dummy for choosing the high track, τ is the intercept, $\Theta = (\theta^j : j \in J)$ is the set of observed mediating factors (cognitive skills, personality traits, and teachers' recommendation), $\mathbf{X}$ is a vector of pre-program variables unaffected by the treatment (initial test score INVALSI6, generation of immigration, and province fixed effects), and ϵ_d is an error term. With the exception of $\mathbf{X}$, all variables and coefficients in equation (S1) are allowed to depend on treatment assignment. In particular, τ_d captures the effect of experimentally induced changes in other (unobserved) determinants of Y, in addition to the observed mediating factors in Θ.

Separately identifying the components of the treatment effect attributable to τ_d and Θ, respectively, requires further assumptions, as experimental variation allows us to consistently estimate the effects of EOP on measured factors and final educational decisions, but not the relationship between the former and the latter. Heckman, Pinto, and Savelyev (2013) assumed independence of observed and unobserved factors in the no-treatment state, conditional on the vector $\mathbf{X}$ of pre-treatment characteristics. Maintaining this assumption and imposing the additional testable restriction that coefficients do not vary with treatment assignment (respectively, $\alpha_d^j = \alpha$ for all j and $\beta_d = \beta$) allows us to decompose the effect of EOP as

$$E(Y_1 - Y_0) = \sum_{j \in J} \alpha^j E(\theta_1^j - \theta_0^j) + (\tau_1 - \tau_0), \tag{S2}$$

where $E(Y_1 - Y_0)$ is the average treatment effect; $E(\theta_1^j - \theta_0^j)$ is the average change induced in the jth observed factor, and α^j is the associated effect on educational choices; finally, $(\tau_1 - \tau_0)$ is the effect due to other unmeasured factors. In Appendix Table AXIII, we test and do not reject the structural invariance assumptions on α^j for all j and β.

REFERENCES

ANGRIST, J. D., S. R. COHODES, S. M. DYNARSKI, P. A. PATHAK, AND C. R. WALTERS (2016): "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice," *Journal of Labor Economics*, 34 (2), 275–318. [18]

ATHEY, S., AND G. IMBENS (2016): "Recursive Partitioning for Heterogeneous Causal Effects," *Proceedings of the National Academy of Sciences*, 113 (27), 7353–7360. [20]

ATHEY, S., J. TIBSHIRANI, S. WAGER et al. (2019): "Generalized Random Forests," *The Annals of Statistics*, 47 (2), 1148–1178. [20]

BERTRAND, M., B. CRÉPON, A. MARGUERIE, AND P. PREMAND (2017): "Contemporaneous and Post-Program Impacts of a Public Works Program: Evidence From Côte D'ivoire," Working Paper. [20]

DAVIS, J., AND S. B. HELLER (2017): "Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs," *American Economic Review: Papers and Proceedings*, 107 (5), 546–550. [20]

GELBACH, J. B. (2016): "When Do Covariates Matter? And Which Ones, and How Much?", *Journal of Labor Economics*, 34 (2), 509–543. [16]

HECKMAN, J., R. PINTO, AND P. SAVELYEV (2013): "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review*, 103 (6), 2052–2086. [23]

HECKMAN, J. J., S. H. MOON, R. PINTO, P. A. SAVELYEV, AND A. YAVITZ (2010): "The Rate of Return to the Highscope Perry Preschool Program," *Journal of public Economics*, 94 (1–2), 114–128. [17]

MCCRARY, J. (2008): "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 142 (2), 698–714.

WAGER, S., AND S. ATHEY (2018): "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests," *Journal of the American Statistical Association*, 113 (523), 1228–1242. [20]

WESTFALL, P. H., AND S. S. YOUNG (1993): *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. John Wiley & Sons. [15]

Co-editor Guido Imbens handled this manuscript as an invited Walras–Bowley lecture. The invitation to deliver the Walras–Bowley lecture is also an invitation to publish a suitable version of the lecture in Econometrica.

Manuscript received 4 July, 2019; final version accepted 6 November, 2020; available online 4 August, 2021.