Parents, Schools and Human Capital Differences across Countries

Marta De Philippis* and Federico Rossi†

June 2017

ONLINE APPENDIX

A Data Appendix

A.1 Data Construction

Given that individual host countries have great flexibility in choosing how to report parents' countries of birth, some aggregation is necessary to get a set of countries consistently defined over time. For what concerns countries included in the PISA sample, we make the following adjustments: we code *Yugoslavia* and similar labels as Serbia and Montenegro, *USSR* and similar labels as Russia, *Albania or Kosovo* as Albania, *France or Belgium* as France, *Germany or Austria* as Germany, *China (including Hong Kong)* as China. Moreover, for the purpose of estimating (4), we group countries of origin not belonging to the PISA sample in several categories (introducing a fixed effect for each of those): in particular, we create dummies for individual countries when possible (Belarus, Bolivia, Bosnia, Pakistan, Paraguay, Philippines, Ukraine), aggregate others in broad geographical groups (Africa, Europe, Middle East) and classify any remaining case as Rest of the World. We drop all observations with inconsistent or missing information on students' or parents' countries of birth.

Parents' educational attainment is reported according to the ISCED 1997 classification system. We group levels 0 and 1 into *primary* education, levels 2, 3 and 4 into *secondary* education and levels 5 and 6 into *tertiary* education.

*Bank of Italy, Department of Economics and Statistics, via Nazionale 91, 00184, Rome, Italy; email: marta.dephilippis@bancaditalia.it

†London School of Economics and CFM, Department of Economics, 32 Lincoln's Inn Fields, WC2A 3PH, London, UK; email: f.rossi2@lse.ac.uk.

A.2 Additional Summary Statistics

Table A.1: Average PISA Scores across Regions

	Math	Reading	Science	# Countries
China	1.33	0.96	1.07	1
Other East Asia	0.79	0.58	0.67	6
Canada	0.63	0.66	0.68	1
EU North	0.57	0.53	0.58	15
Oceania	0.54	0.62	0.67	2
US	0.26	0.45	0.43	1
EU South	0.13	0.18	0.21	5
EU East	-0.08	-0.16	-0.06	19
Other Asia	-0.42	-0.38	-0.36	5
Middle East/NA	-0.55	-0.40	-0.43	7
Latin America	-0.58	-0.38	-0.46	11

Notes: The Table shows the average PISA score of native students across countries belonging to each region, for all available waves (for Science, only waves from 2006 onwards are considered, since the scale was established in 2006 and results from 2003 are not fully comparable with the subsequent ones). Country averages are computed using the provided sample weights. Scores are standardized to have mean 0 and (individual-level) standard deviation 1 across the (pooled, equally weighted) 73 countries participating to at least one wave of the test. Countries are assigned to regional groups as follows. *East Asia*: China, Hong Kong, Japan, Macao, Singapore, South Korea, Taiwan. *EU North*: Austria, Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Liechtenstein, Luxembourg, Netherlands, Norway, Sweden, Switzerland, United Kingdom. *Oceania*: Australia, New Zealand. *EU South*: Greece, Italy, Malta, Portugal, Spain. *EU East*: Albania, Azerbaijan, Bulgaria, Croatia, Czechia, Estonia, Georgia, Hungary, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Poland, Romania, Russia, Serbia and Montenegro, Slovakia, Slovenia. *Other Asia*: India, Indonesia, Malaysia, Thailand, Vietnam. *Latin America*: Argentina, Brazil, Chile, Colombia, Costa Rica, Mexico, Panama, Peru, Trinidad and Tobago, Uruguay, Venezuela. *Middle East / North Africa*: Israel, Jordan, Mauritius, Qatar, Tunisia, Turkey, United Arab Emirates.

Table A.2: Second Generation Immigrants by Country of Origin - PISA

Country of Origin	Mothers Number	Mothers # Host Countries	Mothers Top Host Country	Fathers Number	Fathers # Host Countries	Fathers Top Host Country
Albania	330	5	Greece (183)	306	5	Greece (162)
Argentina	69	2	Uruguay (68)	64	1	Uruguay (64)
Australia	150	2	New Zealand (149)	117	1	New Zealand (117)
Austria	239	2	Switzerland (181)	175	2	Switzerland (137)
Belgium	258	3	Luxembourg (244)	239	2	Luxembourg (219)
Brazil	187	4	Uruguay (78)	162	4	Uruguay (72)
Bulgaria	30	1	Turkey (30)	16	1	Turkey (16)
Canada	2	1	Ireland (2)	2	1	Ireland (2)
Chile	58	1	Argentina (58)	48	1	Argentina (48)
China	14161	12	Macao (9149)	13529	11	Macao (8416)
Colombia	5	1	Costa Rica (5)	5	1	Costa Rica (5)
Croatia	190	3	Serbia-Mont. (101)	167	3	Serbia-Mont. (81)
Czech Republic	187	2	Slovakia (179)	176	2	Slovakia (167)
Denmark	75	2	Norway (74)	91	1	Norway (91)
Estonia	81	1	Finland (81)	53	1	Finland (53)
France	1312	7	Switzerland (626)	1110	7	Switzerland (456)
Germany	1364	9	Switzerland (630)	1086	9	Switzerland (461)
Greece	85	2	Australia (68)	133	2	Australia (112)
Hong Kong	236	2	Macao (168)	432	3	Macao (348)
Hungary	16	2	Austria (14)	15	2	Austria (12)
India	218	4	Australia (189)	220	4	Australia (187)
Italy	1510	9	Switzerland (1006)	2606	9	Switzerland (1741)
Jordan	155	1	Qatar (155)	119	1	Qatar (119)
Liechtenstein	38	1	Switzerland (38)	27	1	Switzerland (27)
Macao	138	1	Hong Kong (138)	123	1	Hong Kong (123)
Malaysia	65	4	Australia (53)	55	4	Australia (45)
Netherlands	226	5	Belgium (195)	264	4	Belgium (192)
New Zealand	781	1	Australia (781)	790	1	Australia (790)
Panama	9	1	Costa Rica (9)	15	1	Costa Rica (15)
Poland	275	3	Germany (211)	213	3	Germany (173)
Portugal	2646	4	Luxembourg (1762)	2517	5	Luxembourg (1730)
Romania	49	2	Austria (47)	55	3	Austria (43)
Russia	4216	13	Estonia (1219)	4092	13	Estonia (1225)
Serbia-Mont.	2615	9	Switzerland (1548)	2653	9	Switzerland (1563)
Singapore	8	1	Indonesia (8)	8	2	Indonesia (7)
Slovakia	472	2	Czech Republic (467)	569	2	Czech Republic (564)
Slovenia	11	2	Austria (8)	17	2	Austria (10)
South Korea	42	2	Australia (30)	43	2	Australia (33)
Spain	334	5	Switzerland (317)	409	4	Switzerland (391)
Sweden	362	2	Finland (230)	272	2	Finland (173)
Switzerland	106	1	Liechtenstein (106)	90	1	Liechtenstein (90)
Taiwan	26	1	Hong Kong (26)	9	2	Hong Kong (6)
Thailand	13	1	Finland (13)	2	1	Finland (2)
Turkey	2411	8	Denmark (535)	2648	8	Switzerland (591)
United Kingdom	3514	5	Australia (2142)	3659	5	Australia (2313)
United States	407	5	Mexico (198)	532	5	Mexico (326)
Uruguay	79	1	Argentina (79)	72	1	Argentina (72)
Vietnam	304	4	Australia (249)	299	3	Australia (240)
Average	834.69	3.38		839.67	3.31	

Notes: The Table shows summary statistics on second generation immigrants from each country of origin in the PISA sample (with at least one observation per parent). *# Host Countries* is the number of different host countries in which second generation immigrants are observed. *Top Host Country* is the host country where the highest number (reported in brackets) of second generation immigrants are observed.

Table A.3: Second Generation Immigrants by Host Country - PISA

Host Country	Mothers: Number	Mothers: # Countries of Origin	Mothers: Top Country of Origin (in PISA)	Fathers: Number	Fathers: # Countries of Origin	Fathers: Top Country of Origin (in PISA)
Argentina	541	6	Uruguay (79)	497	6	Uruguay (72)
Australia	8403	17	United Kingdom (2142)	8740	17	United Kingdom (2313)
Austria	1623	15	Turkey (370)	1603	15	Turkey (393)
Belgium	2820	7	Turkey (384)	3178	7	Turkey (437)
Costa Rica	423	3	Panama (9)	490	3	Panama (15)
Croatia	2013	4	Serbia-Mont. (347)	1819	4	Serbia-Mont. (329)
Czech Republic	661	6	Slovakia (467)	872	6	Slovakia (564)
Denmark	2343	6	Turkey (535)	2460	6	Turkey (541)
Estonia	1492	2	Russia (1219)	1605	2	Russia (1225)
Finland	1045	10	Sweden (230)	1193	10	Sweden (173)
Georgia	58	2	Russia (40)	47	2	Russia (32)
Germany	1194	10	Turkey (392)	1247	10	Turkey (424)
Greece	1167	3	Russia (198)	706	3	Albania (162)
Hong Kong	5110	4	China (4470)	4982	4	China (4650)
Indonesia	39	5	Singapore (8)	41	3	Singapore (7)
Ireland	1085	16	United Kingdom (872)	971	15	United Kingdom (761)
Israel	2015	5	Russia (531)	2130	5	Russia (520)
Kazakhstan	1106	2	Russia (926)	1056	2	Russia (864)
Kyrgyzstan	443	2	Russia (98)	275	2	Russia (94)
Latvia	1932	4	Russia (811)	2197	4	Russia (919)
Liechtenstein	301	11	Switzerland (106)	258	11	Switzerland (90)
Luxembourg	4116	10	Portugal (1762)	4207	10	Portugal (1730)
Macao	9755	5	China (9149)	9238	7	China (8416)
Mauritius	80	4	China (11)	54	4	China (8)
Mexico	972	4	United States (198)	1256	4	United States (326)
Moldova	174	3	Russia (53)	159	4	Russia (50)
Netherlands	1468	16	Turkey (167)	1547	16	Turkey (191)
New Zealand	1736	8	United Kingdom (471)	1869	8	United Kingdom (548)
Norway	1025	3	Sweden (132)	1013	3	Sweden (99)
Portugal	1450	5	Brazil (58)	1232	5	Brazil (57)
Qatar	4587	4	Jordan (155)	4027	4	Jordan (119)
Serbia-Mont.	2088	4	Croatia (101)	1617	4	Croatia (81)
Slovakia	524	3	Czech Republic (179)	526	3	Czech Republic (167)
Slovenia	1591	3	Italy (7)	1630	3	Italy (9)
South Korea	26	5	China (10)	-	-	-
Switzerland	8036	11	Serbia-Mont. (1548)	7886	11	Italy (1741)
Turkey	211	5	Germany (62)	176	5	Germany (29)
United Kingdom	2025	7	China (23)	2181	7	China (24)
Uruguay	265	4	Brazil (78)	283	4	Brazil (72)
Average	1947.26	6.26		1930.03	6.15	

Notes: The Table shows summary statistics on second generation immigrants observed in each country in the PISA sample, across all available waves. Only host countries with second generation immigrants from at least one country of origin in the PISA sample are included. *# Countries of Origin* is the number of different countries of origin of second generation immigrants in a given host country. *Top Country of Origin (in PISA)* is the country of origin from which the highest number (across all countries in the PISA sample, not considering other countries of origin) of second generation immigrants in a given host country are observed (number reported in brackets).

B Robustness of Baseline Result

B.1 PISA

B.1.1 Results for Second Generation Immigrants on the Father's Side

Table B.1: Main Results for Fathers - PISA

	Dependent Variable: Math Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country *f*	0.676***	0.535***	0.230***	0.225***	0.212***	0.179*
	(0.079)	(0.166)	(0.086)	(0.080)	(0.072)	(0.108)
Female	-0.139***	-0.140***	-0.204***	-0.200***	-0.197***	-0.181***
	(0.028)	(0.024)	(0.028)	(0.028)	(0.028)	(0.035)
Father Sec Edu				-0.003	-0.016	-0.025
				(0.023)	(0.022)	(0.041)
Father Ter Edu				0.080**	0.030	0.016
				(0.034)	(0.033)	(0.046)
Mother Sec Edu				0.025	0.005	0.056
				(0.044)	(0.043)	(0.082)
Mother Ter Edu				0.073	0.025	0.067
				(0.056)	(0.057)	(0.103)
Mother Working × ISEI					0.003***	0.003***
					(0.001)	(0.001)
Father Working × ISEI					0.003***	0.003***
					(0.001)	(0.001)
N	40304	40304	40304	40304	40304	26160
# Country *f*	48	48	48	48	48	42
R Squared	0.16	0.25	0.66	0.66	0.66	0.63
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the father's side. Sample includes only cases where both parents report a country of origin and the country of origin of the father runs a PISA test on natives. *Score Country f* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the father, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for mother's immigrant status; specifications 5-6 additionally control for dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. Observations are weighted according to the provided sample weights. Standard errors are clustered by father's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

B.1.2 Results for Second Generation Immigrants and Natives

Table B.2: Main Results for Mothers and Fathers - PISA

	Dependent Variable: Math Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country *m*	0.307***	0.304***	0.163***	0.156***	0.149***	0.132**
	(0.106)	(0.118)	(0.042)	(0.041)	(0.042)	(0.058)
Score Country *f*	0.400***	0.386***	0.202***	0.195***	0.184***	0.172***
	(0.098)	(0.095)	(0.041)	(0.039)	(0.038)	(0.067)
Score Country *m* * Native Mother	0.219**	0.103	0.026	0.024	0.014	0.050
	(0.087)	(0.071)	(0.030)	(0.033)	(0.034)	(0.049)
Score Country *f* * Native Father	0.093	-0.032	-0.045	-0.053	-0.060*	-0.042
	(0.099)	(0.091)	(0.036)	(0.037)	(0.034)	(0.060)
Female	-0.117***	-0.117***	-0.151***	-0.146***	-0.143***	-0.143***
	(0.011)	(0.011)	(0.012)	(0.013)	(0.012)	(0.013)
Native Mother	-0.038	0.003	0.015	0.006	-0.002	0.031
	(0.060)	(0.054)	(0.024)	(0.036)	(0.038)	(0.031)
Native Father	0.016	0.071	0.057**	0.034	0.026	0.025
	(0.067)	(0.060)	(0.023)	(0.025)	(0.025)	(0.030)
Father Sec Edu				0.015	0.000	0.009
				(0.022)	(0.022)	(0.035)
Father Ter Edu				0.080**	0.024	0.024
				(0.036)	(0.032)	(0.041)
Mother Sec Edu				0.025	0.012	0.056*
				(0.032)	(0.035)	(0.030)
Mother Ter Edu				0.066*	0.014	0.050
				(0.039)	(0.039)	(0.042)
Native Father × Father Sec Edu				0.023	0.023	0.014
				(0.024)	(0.023)	(0.036)
Native Father × Father Ter Edu				0.024	0.020	0.018
				(0.041)	(0.037)	(0.047)
Native Mother × Mother Sec Edu				0.002	0.000	-0.047
				(0.036)	(0.037)	(0.030)
Native Mother × Mother Ter Edu				0.019	0.015	-0.024
				(0.045)	(0.045)	(0.045)
Mother Working × ISEI					0.004***	0.004***
					(0.000)	(0.000)
Father Working × ISEI					0.004***	0.004***
					(0.001)	(0.001)
N	1181347	1181347	1181347	1181347	1181347	1089297
# Country *m*	49	49	49	49	49	42
# Country *f*	48	48	48	48	48	42
R Squared	0.34	0.34	0.61	0.61	0.62	0.60
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants and natives. The sample includes only cases where both parents report a country of origin that runs a PISA test on natives. *Score Country m* and *Score Country f* are the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and father, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for father's immigrant status; specifications 5-6 additionally control for dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's and father's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

B.1.3 Results for Reading and Science

Table B.3: Results for PISA Reading

	Dependent Variable: Reading Test Score					
	[1]	[2]	[3] All	[4]	[5]	[6] No East Asia
Score Read Country *m*	0.724***	0.333*	0.165***	0.166***	0.157***	0.119**
	(0.146)	(0.184)	(0.063)	(0.058)	(0.056)	(0.058)
Female	0.311***	0.301***	0.240***	0.243***	0.246***	0.269***
	(0.038)	(0.035)	(0.023)	(0.024)	(0.024)	(0.017)
Father Sec Edu				0.068	0.055	0.110**
				(0.048)	(0.049)	(0.055)
Father Ter Edu				0.145***	0.100*	0.145**
				(0.056)	(0.059)	(0.068)
Mother Sec Edu				-0.002	-0.017	-0.002
				(0.028)	(0.029)	(0.055)
Mother Ter Edu				0.044	0.002	0.010
				(0.045)	(0.046)	(0.075)
Mother Working × ISEI					0.002***	0.003***
					(0.001)	(0.001)
Father Working × ISEI					0.003***	0.003***
					(0.000)	(0.001)
N	40067	40067	40067	40067	40067	25454
# Country *m*	49	49	49	49	49	42
R Squared	0.14	0.26	0.69	0.69	0.70	0.67
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample includes only cases where both parents report a country of origin and the country of origin of the mother runs a PISA test on natives. *Score Read Country m* is the average reading PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

Table B.4: Results for PISA Science

	Dependent Variable: Science Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Science Country *m*	0.783***	0.537***	0.280***	0.273***	0.262***	0.253***
	(0.113)	(0.196)	(0.087)	(0.081)	(0.075)	(0.096)
Female	-0.049	-0.057*	-0.104***	-0.098***	-0.098***	-0.069***
	(0.036)	(0.031)	(0.027)	(0.028)	(0.028)	(0.017)
Father Sec Edu				0.075**	0.061*	0.105**
				(0.036)	(0.036)	(0.049)
Father Ter Edu				0.180***	0.134**	0.173***
				(0.048)	(0.054)	(0.066)
Mother Sec Edu				0.013	-0.002	0.042
				(0.037)	(0.037)	(0.066)
Mother Ter Edu				0.060	0.013	0.036
				(0.047)	(0.048)	(0.089)
Mother Working × ISEI					0.003***	0.003***
					(0.001)	(0.001)
Father Working × ISEI					0.003***	0.004***
					(0.001)	(0.001)
N	34161	34161	34161	34161	34161	21385
# Country *m*	48	48	48	48	48	41
R Squared	0.14	0.25	0.66	0.67	0.67	0.63
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample includes only cases where both parents report a country of origin and the country of origin of the mother runs a PISA test on natives. *Score Science Country m* is the average science PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

B.1.4 Standard Errors

Throughout the paper, standard errors for the analyses on PISA data are constructed taking into account the fact that student performance is reported through plausible values. Using the average of the five plausible values as a measure of individual performance guarantees unbiased estimates of group-level means and regression coefficients; however, measures of dispersion need to take into account the within-student variability in plausible values.

As recommended in OECD (2009), for the purpose of computing standard errors all regression with individual test scores as dependent variable are estimated five times, using all plausible values in turn. For each regression we employ an estimator for the sampling variance clustered at the level of the mother's country of origin. The final sampling variance, SV, is given by the average of the sampling variances obtained with the five plausible values.

In addition, standard errors are inflated by the imputation variance due to the fact that test scores measure the latent student's skills with error. The imputation variance, IV, is estimated as the average squared deviation between the estimates obtained with each plausible value and the final estimate (obtained using the average of the plausible values), with the appropriate degree of freedom adjustment.

Finally, as shown in Little and Rubin (1987), the final error variance TV can be obtained by combining the sampling and imputation variance in

$$TV = SV + \left(1 + \frac{1}{K}\right) IV$$

where $K = 5$ is the number of plausible values for each student. The final standard errors are given by the squared roots of the final error variances.

As an alternative to estimate SV, OECD (2009) recommends to apply Fay's variant of the Balanced Repeated Replication (BRR) method, which directly takes into account the two-stage stratified sampling design of the PISA test. This is implemented by iterating each regression over the 80 sets of replicate weights provided in the PISA dataset. The sampling variance estimate is then given by the average squared deviation between the replicated estimates and the estimate obtained with final weights, with a degree of freedom correction depending on the Fay coefficient (a parameter that governs the variability between different sets of replicate weights).

Table B.5 shows the resulting standard errors for our baseline specification. For computational convenience, we implemented the "unbiased shortcut" procedure described in OECD (2009), which uses only one set of plausible values to estimate the sampling variance (while the imputation variance is estimated using all five sets, as described above). In all specifications, the standard error on our coefficient of interest is smaller compared to Table II in the main text, suggesting that our clustered sampling variance is rather conservative.

Table B.5: Main results-PISA (BRR Standard Errors)

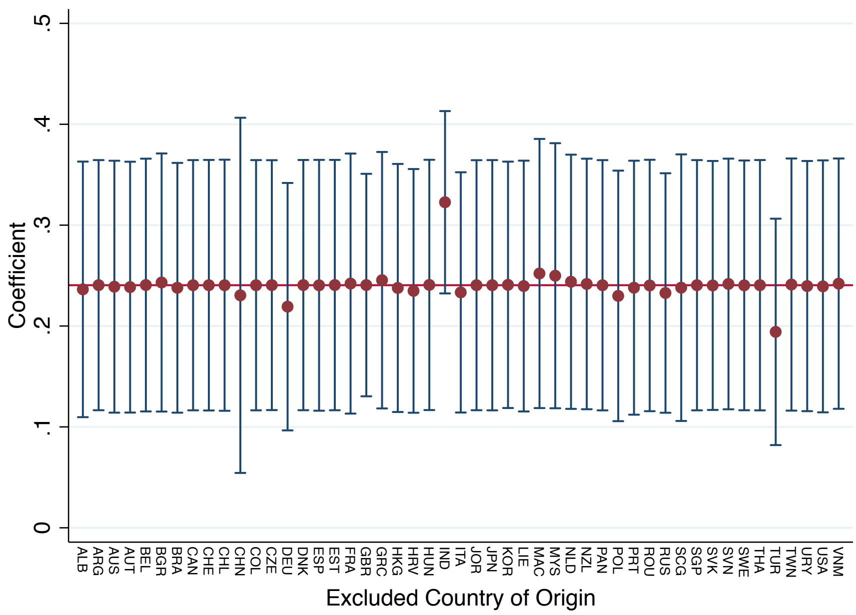
	Dependent Variable: Math Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country *m*	0.662***	0.499***	0.253***	0.249***	0.240***	0.225***
	(0.024)	(0.046)	(0.034)	(0.034)	(0.034)	(0.053)
Female	-0.140***	-0.148***	-0.206***	-0.204***	-0.201***	-0.187***
	(0.025)	(0.022)	(0.019)	(0.018)	(0.018)	(0.023)
Father Sec Edu				0.030	0.014	0.022
				(0.030)	(0.030)	(0.056)
Father Ter Edu				0.099**	0.045	0.049
				(0.039)	(0.042)	(0.063)
Mother Sec Edu				0.001	-0.015	0.027
				(0.027)	(0.026)	(0.046)
Mother Ter Edu				0.032	-0.011	0.023
				(0.034)	(0.034)	(0.050)
Mother Working × ISEI					0.003***	0.003***
					(0.001)	(0.001)
Father Working × ISEI					0.003***	0.003***
					(0.001)	(0.001)
N	40067	40067	40067	40067	40067	25454
# Country m	49	49	49	49	49	42
R Squared	0.16	0.25	0.67	0.67	0.67	0.63
Host Country FE	No	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample and specifications are the same as in Table II in the main text. Standard errors are computed using the provided replicate weights, and inflated by the estimated measurement error in test scores. The sampling variance is estimated through the "unbiased shortcut" procedure described in OECD (2009). * denotes significance at 10%, ** at 5%, *** at 1%.

B.1.5 Excluding Single Countries

In this section we investigate to what extent our results are driven by specific countries of origin or host countries. Figure B.1 shows the estimated coefficient of interest when countries of origin are excluded one by one. The resulting estimates are never significantly different from the baseline, represented by the horizontal line. Even if the difference is insignificant, the coefficient is substantially higher when second generation students from India are excluded; this reflect the fact that these students are outliers since they perform relatively well even though, across natives, India is near the bottom of the international ranking. On the other had, the coefficient becomes somewhat smaller when second generation immigrants from Germany, Poland and Turkey are excluded. Overall, the statistical significance and the rough magnitude of our coefficient of interest is not driven by any specific country of origin.

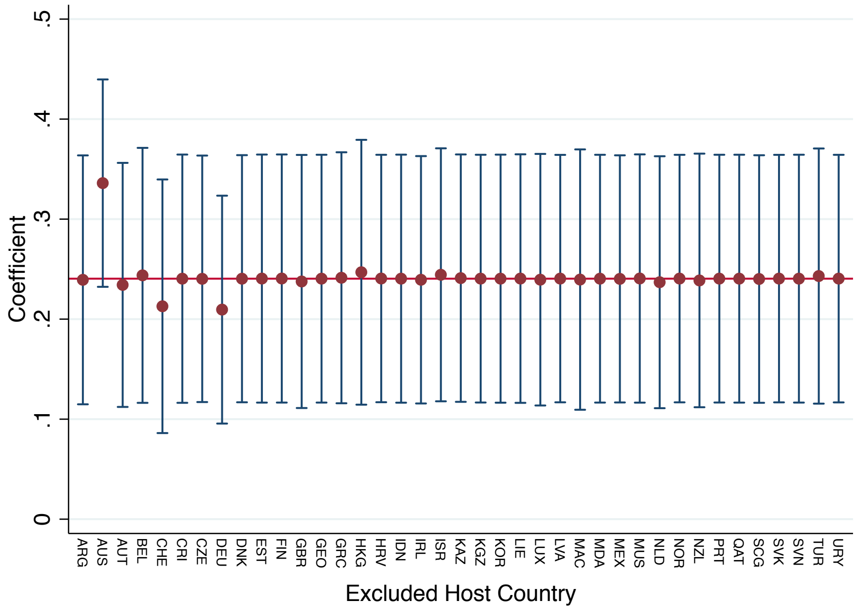
Figure B.1: Reduced Form Coefficient when Excluding Countries of Origin One by One



Notes: The Figure plots the estimated coefficients and 95% confidence intervals on the average PISA score of natives in mother's country of origin, with the dependent variable and other controls being the same as in column 5 of Table II. Each dot corresponds to a different specification, where students with mothers from the indicated country of origin are excluded. Standard errors are clustered by mother's country of origin.

Figure B.2 shows the result from the corresponding exercise on host countries. The coefficient is positive, significant and quite stable across the 39 specifications. As a partial exception, the coefficient is quite a bit higher (even though the difference is not statistically significant) when second generation immigrants in Australia are excluded from the sample. While in principle this might be due to a number of factors, a possible rationalization is the relatively stronger negative selection of East Asian emigrant parents to Australia, given the geographic proximity.

Figure B.2: Reduced Form Coefficient when Excluding Host Countries One by One



Notes: The Figure plots the estimated coefficients and 95% confidence intervals on the average PISA score of natives in mother's country of origin, with the dependent variable and other controls being the same as in column 5 of Table II. Each dot corresponds to a different specification, where students in the indicated host country are excluded. Standard errors are clustered by mother's country of origin.

B.1.6 Alternative Measures of School Performance

Table B.6 shows results from our baseline specification with different measures of school performance available in the PISA data as outcome variables. Columns 1 and 2 consider the grade attended by each student, whose variation within country and within school (once age is controlled for) will be driven by either grade repetition or grade skipping. Columns 3 and 4 consider a measure of truancy, that is the number of school days each student has skipped in the previous two weeks; this measure is available for the 2012 wave only. According to both measures, second generation immigrants from high PISA countries perform better. Within the same school, an extra individual-level standard deviation in the average performance of native students in the mother's country of origin corresponds to 0.07 extra grades (9% of a standard deviation) and 0.041 less school days skipped (13% of a standard deviation). The gap is marginally larger if we do not include school fixed effects.

Table B.6: Alternative Measures of School Performance

	Grade		Truancy	
	[1]	[2]	[3]	[4]
Score Country *m*	0.085*	0.069*	-0.046***	-0.041*
	(0.050)	(0.039)	(0.017)	(0.025)
Female	0.119***	0.092***	-0.014	-0.010
	(0.020)	(0.015)	(0.019)	(0.015)
Father Sec Edu	0.049*	0.017	-0.050	-0.018
	(0.026)	(0.040)	(0.038)	(0.039)
Father Ter Edu	0.047	0.016	-0.045	-0.022
	(0.033)	(0.042)	(0.045)	(0.039)
Mother Sec Edu	0.067**	0.000	-0.004	0.008
	(0.033)	(0.021)	(0.039)	(0.039)
Mother Ter Edu	0.096**	-0.003	-0.007	-0.005
	(0.043)	(0.030)	(0.037)	(0.031)
Mother Working $\times$ ISEI	0.002***	-0.000	0.000	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)
Father Working $\times$ ISEI	0.001*	0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)
N	40017	40017	12368	12368
# Country *m*	49	49	46	46
Mean Dep Var	9.38	9.38	0.11	0.11
SD Dep Var	0.75	0.75	0.31	0.31
R Squared	0.44	0.78	0.18	0.60
Host Country FE	Yes	Yes	Yes	Yes
School FE	No	Yes	No	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The dependent variable is grade attended in columns 1 and 2 and the number of school days skipped in the previous two weeks in columns 3 and 4. The controls are the same as in column 5 of Table II in the main text. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

B.1.7 Alternative Measures of Socio-Economic Status

Table B.7 considers alternative measures of parental socio-economic status available from the PISA questionnaires. In Column 2 we control for an index of family wealth, based on the presence and the number of various items in students' homes, including computers, cars, cellular phones, televisions and rooms with bath or shower. Column 3 includes an index of home possessions, which is based on all elements in the wealth index and additionally considers books, various educational resources and pieces of classical culture. Column 4 considers the broadest measure available in PISA, an index of Economic, Social and Cultural Status (ESCS) which combines home possessions with information on parents' education and occupational status. All indexes are standardized to take mean 0 and (individual-level) standard deviation 1 across the 73 countries (pooled, equally weighted) participating to the test.

The results are very similar compared to the baseline specification, reported in column 1. The magnitude of our coefficient of interest varies little across specifications, even when (in Column 5) we introduce all indexes of socio-economic status in the same regression. Home possessions and the ESCS index are positively related to students' performance, while wealth is not.[1] Overall, the results suggest the controlling further for observable measures of socio-economic background does not affect affect the magnitude of our estimated parental component.

[1]Much of the variation in wealth seems to be absorbed by the school fixed effect, since this index enters positively and significantly in a specification with host country fixed effects (results not shown, available upon request).

Table B.7: Alternative Measures of Socio-economic Status

	Dependent Variable: Math Test Score				
	[1]	[2]	[3]	[4]	[5]
Score Country *m*	0.240***	0.272***	0.248***	0.236***	0.261***
	(0.065)	(0.071)	(0.068)	(0.065)	(0.068)
Female	-0.201***	-0.190***	-0.206***	-0.204***	-0.198***
	(0.022)	(0.027)	(0.022)	(0.023)	(0.026)
Father Sec Edu	0.014	0.045*	0.018		
	(0.022)	(0.026)	(0.021)		
Father Ter Edu	0.045	0.112***	0.076**		
	(0.034)	(0.037)	(0.031)		
Mother Sec Edu	-0.015	-0.005	-0.007		
	(0.037)	(0.039)	(0.039)		
Mother Ter Edu	-0.011	0.042	0.008		
	(0.042)	(0.044)	(0.041)		
Mother Working $\times$ ISEI	0.003***				
	(0.001)				
Father Working $\times$ ISEI	0.003***				
	(0.001)				
Wealth		-0.004			-0.211***
		(0.023)			(0.035)
Home Possessions			0.087***		0.180***
			(0.019)		(0.039)
ESCS				0.111***	0.096***
				(0.029)	(0.032)
N	40067	34134	40049	40060	34134
# Country m	49	48	49	49	48
R Squared	0.67	0.68	0.67	0.67	0.69
Host Country FE	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample includes only cases where both parents report a country of origin and the country of origin of the mother runs a PISA test on natives. *Score Country m* refers to the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across the OECD countries participating in 2003) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' exact age (in months), wave fixed effect and a dummy for father immigrant status; specification 1 additionally controls for dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. *Wealth*, *Home Possessions* and *ESCS* are indexes of socio-economic status, discussed in the text. Observations weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%.

B.2 US Census

B.2.1 Results for Second Generation Immigrants on the Father's Side

Table B.8: Main Results for Fathers - US CENSUS

	Dependent variable: 1 = Never repeated a grade					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country f	0.102***	0.063***	0.038***	0.033***	0.029***	0.023*
	(0.037)	(0.020)	(0.012)	(0.011)	(0.011)	(0.013)
Female	0.072***	0.071***	0.072***	0.072***	0.072***	0.073***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Mother Sec Edu			0.066***	0.065***	0.059***	0.060***
			(0.017)	(0.017)	(0.015)	(0.016)
Mother Ter Edu			0.086***	0.085***	0.071***	0.073***
			(0.013)	(0.014)	(0.012)	(0.013)
Father Sec Edu			0.035***	0.031***	0.027***	0.029***
			(0.007)	(0.007)	(0.007)	(0.007)
Father Ter Edu			0.057***	0.053***	0.041***	0.044***
			(0.009)	(0.009)	(0.007)	(0.007)
Log Family Income					0.034***	0.035***
					(0.005)	(0.005)
N	46410	46410	46410	46410	46410	43909
# Country f	61	61	61	61	61	54
R Squared	0.07	0.10	0.11	0.12	0.12	0.12
Comm Zone FE	No	Yes	Yes	Yes	Yes	Yes
Years Since Migr Father	No	No	No	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the father's side. *Score Country f* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the father, across all available waves. All specifications control for intercept, child age dummies, parents' age, number of siblings, year fixed effect, (year-specific) quarter of birth fixed effect and mother's immigrant status. Observations weighted according to the provided sample weights. Standard errors are clustered by father's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

B.2.2 Results for Second Generation Immigrants and Natives

Table B.9: Main Results for Mothers and Fathers - US CENSUS

	Dependent variable: 1 = Never repeated a grade					
	[1]	[2]	[3]	[4]	[5]	[6]
			All			No East Asia
Score Country *m*	0.050**	0.036***	0.018*	0.015*	0.013*	0.010
	(0.022)	(0.013)	(0.009)	(0.008)	(0.008)	(0.009)
Score Country *f*	0.075**	0.057***	0.032***	0.026***	0.023***	0.020*
	(0.030)	(0.020)	(0.010)	(0.009)	(0.008)	(0.010)
Native Mother	-0.005	-0.005	-0.071***	0.000	-0.071***	-0.071***
	(0.007)	(0.005)	(0.016)	(0.000)	(0.018)	(0.019)
Native Father	-0.002	0.002	-0.059***	-0.066***	-0.061***	-0.058***
	(0.009)	(0.006)	(0.015)	(0.016)	(0.015)	(0.015)
Female	0.084***	0.084***	0.085***	0.085***	0.085***	0.085***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
Mother Sec Edu			0.060***	0.058***	0.055***	0.055***
			(0.015)	(0.015)	(0.015)	(0.016)
Mother Ter Edu			0.077***	0.075***	0.065***	0.062***
			(0.015)	(0.015)	(0.015)	(0.016)
Father Sec Edu			0.040***	0.038***	0.033***	0.035***
			(0.010)	(0.010)	(0.010)	(0.009)
Father Ter Edu			0.061***	0.059***	0.045***	0.048***
			(0.014)	(0.013)	(0.013)	(0.013)
Native Mother × Mother Sec Edu			0.059***	0.061***	0.055***	0.055***
			(0.015)	(0.015)	(0.015)	(0.016)
Native Mother × Mother Ter Edu			0.068***	0.069***	0.063***	0.066***
			(0.015)	(0.015)	(0.015)	(0.016)
Native Father × Father Sec Edu			0.043***	0.045***	0.041***	0.039***
			(0.010)	(0.010)	(0.010)	(0.009)
Native Father × Father Ter Edu			0.058***	0.059***	0.056***	0.052***
			(0.014)	(0.014)	(0.013)	(0.013)
Log Family Income					0.035***	0.035***
					(0.000)	(0.000)
N	1299079	1299079	1299079	1299079	1292410	1288059
# Country *m*	61	61	61	61	61	54
# Country *f*	61	61	61	61	61	54
R Squared	0.04	0.05	0.07	0.07	0.08	0.08
Comm Zone FE	No	Yes	Yes	Yes	Yes	Yes
Years Since Migr Mother	No	No	No	Yes	Yes	Yes
Years Since Migr Father	No	No	No	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants and natives. Sample includes only cases where both parents report a country of origin that runs a PISA test on natives. *Score Country m* and *Score Country f* are the average math PISA score of natives (standardized to have mean 0 and standard deviation 1) in the country of birth of the mother and father, across all available waves. All specifications control for intercept, child age dummies, parents' age, number of siblings, log family income, year fixed effect and (year-specific) quarter of birth fixed effect. Observations are weighted accordind to the provided sample weights. Robust standard errors clustered by mother's and father's country of origin. * denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%.

C Additional Results on Selection

C.1 Selection into Schools

Table C.1 shows regressions with several proxies for school quality (constructed from the information available in the School Questionnaire) as outcome variables. After controlling for country fixed effects and the usual observable characteristics, a higher PISA score in the country of origin of the mother is associated with schools where natives score better in the PISA test, no matter whether we take the raw average (column 1) or clean it from observable characteristics (column 2), where admissions are more likely to be based on academic records, the proportion of teachers with at least some tertiary education is higher and the proportion of students dropping out is lower.

Table C.1: Selection into Schools

	Avg Score School	Estimated School FE	Academic Admission	Share Qual Teachers	Dropout Rate
	[1]	[2]	[3]	[4]	[5]
Score Country *m*	0.167*	0.154**	0.045*	0.026**	-0.016**
	(0.087)	(0.076)	(0.026)	(0.012)	(0.007)
Female	0.046**	0.051***	0.009	0.004	-0.003
	(0.019)	(0.018)	(0.009)	(0.007)	(0.003)
Father Sec Edu	0.074**	0.061*	0.053*	-0.021	0.005
	(0.029)	(0.034)	(0.029)	(0.014)	(0.010)
Father Ter Edu	0.148***	0.127***	0.072**	-0.005	-0.001
	(0.037)	(0.038)	(0.028)	(0.017)	(0.011)
Mother Sec Edu	0.114***	0.100***	0.015	-0.004	-0.034
	(0.024)	(0.022)	(0.018)	(0.010)	(0.028)
Mother Ter Edu	0.146***	0.123***	0.009	-0.010	-0.037
	(0.036)	(0.037)	(0.027)	(0.016)	(0.026)
Mother Working $\times$ ISEI	0.005***	0.004***	0.001***	0.001***	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Father Working $\times$ ISEI	0.005***	0.004***	0.000	0.001***	-0.000**
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
N	39573	39448	40067	29893	9751
# Country *m*	49	49	49	48	41
Mean Dep. Var.	0.46	0.50	0.75	0.80	0.03
St. Dev. Dep. Var.	0.59	0.56	0.43	0.28	0.07
R Squared	0.37	0.39	0.18	0.42	0.17
Host Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes only second generation immigrants on the mother's side. *Avg Score School* is the average math PISA score of native students in the same school, *Estimated School FE* is the estimated school fixed effect in a regression of the math PISA score on gender, age in months, parental education and occupational status (limiting the sample to native students), *Academic Admission* is a dummy that takes value 1 whenever schools report that student's record of academic performance is either *always* or *sometimes* considered for admissions, *Share Qual Teachers* is the share of current teachers with at least the ISCED 5A level of education and *Dropout Rate* is the share of students who leave the school without having obtained the corresponding diploma. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect, a dummy for father's immigrant status and dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

C.2 Selection into Host Countries

An additional concern is that immigrant parents from high PISA countries may be systematically selecting host countries (or schools) where, because of idiosyncratic factors, it is easier for them and their children to integrate and perform well. Of course the quality of the match between parents or children on one side and countries or schools on the other is unobservable, and it is difficult to rule out entirely this possibility. However, we can check whether immigrants from high PISA countries are located in countries which are, according to reasonable proxies, culturally closer to their country of origin. Table C.2 explores this possibility. In column 2 we add to the baseline regression of column 1 a dummy variable that takes value 1 for all students that declare to speak a foreign language at home (which is available only for part of the sample). While the coefficient on this newly added control is, as expected, negative and significant, our main coefficient of interest is virtually unaffected.

In column 4 and 6 we add to the baseline specifications (reported in columns 3 and 5 respectively) controls for linguistic distance (constructed through the softwares provided by the Automated Similarity Judgment Program (Wichmann and Brown, 2016)) and cultural distance (from Spolaore and Wacziarg (2015)); both measures are standardized to have mean 0 and standard deviation 1 across all country pairs in the sample. In both cases the impact on our coefficient of interest is positive and of negligible magnitude.[2]

[2]In recent work, Isphording et al. (2016) argue that linguistic distance impacts immigrant students' mathematics performance through its effect on reading skills. These results are not in contrast with ours given that we are looking at linguistic distance for immigrant parents, while all students in our sample are born in the country where they attend school.

Table C.2: Linguistic and Cultural Distance

	Dependent Variable: Math Test Score					
	[1]	[2]	[3]	[4]	[5]	[6]
Score Country *m*	0.220***	0.217***	0.239***	0.241***	0.216**	0.241***
	(0.062)	(0.060)	(0.067)	(0.065)	(0.088)	(0.088)
Female	-0.192***	-0.194***	-0.201***	-0.200***	-0.202***	-0.201***
	(0.026)	(0.027)	(0.022)	(0.022)	(0.023)	(0.023)
Father Sec Edu	0.012	0.010	0.015	0.015	-0.015	-0.018
	(0.022)	(0.023)	(0.021)	(0.021)	(0.057)	(0.059)
Father Ter Edu	0.034	0.032	0.041	0.042	0.014	0.009
	(0.035)	(0.035)	(0.031)	(0.030)	(0.069)	(0.072)
Mother Sec Edu	-0.030	-0.035	-0.013	-0.012	0.007	0.012
	(0.033)	(0.033)	(0.039)	(0.039)	(0.101)	(0.108)
Mother Ter Edu	-0.012	-0.017	-0.012	-0.010	-0.050	-0.044
	(0.038)	(0.039)	(0.043)	(0.044)	(0.122)	(0.130)
Mother Working × Mother ISEI	0.002***	0.002***	0.003***	0.003***	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Father Working × Father ISEI	0.003***	0.003***	0.003***	0.003***	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Foreign Language at Home		-0.063**				
		(0.028)				
Mother Linguistic Distance				-0.001		
				(0.013)		
Father Linguistic Distance				0.012		
				(0.009)		
Mother Cultural Distance						0.070
						(0.075)
Father Cultural Distance						-0.053
						(0.082)
N	37827	37827	38487	38487	10309	10309
# Country m	49	49	49	49	35	35
R Squared	0.67	0.67	0.67	0.67	0.68	0.69
Host Country FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side, augmented for controls for linguistic and cultural distance. Sample includes only cases where both parents report a country of origin and the country of origin of the mother participates to PISA. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, students' age (in months), wave fixed effect, a dummy for father's immigrant status and dummies for parents' employment status (full-time employed, part-time employed, not working). *Working* refers to either full-time or part-time employed. *Linguistic Distance* and *Cultural Distance* vary at the country-pair level, and are standardized to take mean 0 and standard deviation 1 across all country pairs in the sample (sources are discussed in the main text). Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

C.3 Insights from the Migration Literature

The migration literature has extensively debated the country-level determinants of emigrants' self-selection in terms of observable and unobservable skills. While, to our knowledge, the PISA score itself has not been explicitly considered in this literature, this variable is correlated with several others that have been advocated as measuring direct determinants of selection. In Figure C.1 we plot some of these variables against the PISA score of native students in the country of origin; since in the PISA sample we do not know the exact date of migration, we use data on selection determinants in 1985 or the closest available data, which should plausibly approximate the pre-migration conditions for the average migrant in our sample.[3]

First, the seminal contribution of Borjas (1987) gives a central role to the difference in income inequality between the origin and destination countries, predicting positive selection if the wage structure of the host country is such that skills are rewarded more compared to the country of origin, and negative selection in the opposite case. Panels (a) and (b) of Figure C.1 show that on average emigrant parents from high PISA countries do emigrate to countries more unequal (as measured by the Gini coefficient and the estimated return to education) than their countries of origin, implying that they would be more positively selected according to Borjas' theory.[4] However, this theory has received mixed support (Chiquiar and Hanson, 2005; Belot and Hatton, 2012), and in particular Grogger and Hanson (2011) argue that the absolute (as opposed to the relative) wage gap between high and low earners provides a better rationalization of the patterns of selection observed in the data. Panel (c) shows that, according to the preferred measure in Grogger and Hanson (2011), emigrants from high PISA countries (if anything) face a relatively lower absolute earning spread in their host countries, implying that they would be more negatively selected.[5]

Another strand of the literature emphasizes the importance of liquidity constraints (Chiswick, 2000; Belot and Hatton, 2012). These papers suggest that emigrants' self-selection should be more negative from richer countries, where facing emigration costs is affordable for a larger share of the population. Since the average PISA score is positively correlated with real GDP in 1985 (Panel d), we should expect negative differential selection according to this mechanism as well. Panel (e) shows instead the extent to which emigrants choose countries with a large pre-existing community from the same country of origin, since McKenzie and Rapoport (2010), among others, argue that stronger social networks act to reduce the effective cost of migration inducing negative selection.[6] China is an outlier in this dimension, since many Chinese parents are observed in Macao and Hong Kong, where Chinese-born represented respectively the 37% and 36% of the population in 1980; therefore, this "chain migration" view would predict negative selection for China, and no systematic pattern of differential selection for the other countries.[7] Finally, Panel (f) shows that emigrants from high PISA countries are not systematically located in a country closer or farther from their country of origin.[8] This is relevant since geographical distance has been shown to be associated with negative selection (Grogger and Hanson, 2011; Belot and Hatton, 2012), most likely due to its effect on the cost of migrating.

Recent work by Albornoz et al. (2012) examines theoretically the determinants of selection in terms of parental motivation for their children's education, which might be only partially correlated with parents' skills. Among other channels, the authors stress the importance of the relative quality of

[3]In the US Census, where we observe years since migration, the average mother of a US-born 15-year-old student migrated 20 years earlier.

[4]We take the Gini Index from the cross-country dataset constructed in Brueckner and Lederman (2015), and we use the 1985 observation when available and 1990 or 1995 when not. The Mincerian coefficients come from Psacharopoulos and Patrinos (2004), who collect estimates from a large set of papers; most observations refer to the 1980s.

[5]Grogger and Hanson (2011) combine information from the Luxembourg Income Study and the WIDER dataset to construct an estimate of the absolute income gap (in thousands of 2000 US dollars) between the 80^{th} and 20^{th} percentiles of the income distribution in each country.

[6]We construct a matrix of bilateral migration shares in 1980 from the Global Bilateral Migration Database, discussed in Ozden et al. (2011). Each entry of this matrix gives us the share of the resident population in country i that was born in country j.

[7]The results of the paper are robust to the omission of Macao and Hong Kong as host countries, and to their aggregation to China as well. If anything, the relative over-performance of Chinese second-generation immigrants compared to other countries of origin is weaker in these two countries, perhaps due to the patterns of selection discussed in this section.

[8]The geographical distance data comes from the CEPII's GeoDist dataset (Mayer and Zignago, 2011). We use the simple distance between the most populated cities, expressed in kilometers.

the school systems in the host and source countries, since highly motivated parents are more likely to migrate to countries with better educational prospects for their children.[9] Under the presumption that high PISA countries have better schools on average, parents emigrating from these countries should be, ceteris paribus, relatively more negatively selected.

All in all, given the determinants of self-selection considered in the literature, we conclude that a pattern of (weakly) negative differential selection should be expected.

C.4 Selection Analysis for the Census Data

In this Appendix we provide a discussion of the patterns of differential selection in the US Census data. While the analysis parallels the one in the main text on the PISA sample, the information on years since migration available in the Census allows us to implement additional checks.

In order to benchmark emigrant parents against non-emigrants in their country of origin, we use school attainment data from Barro and Lee (2013), combined with information on the duration of primary and secondary school in each country from the World Development Indicators, to construct estimates for the average and the standard deviation of years of education in the across countries of origin.[10] Differently from the PISA data, we cannot build these measures for parents of school-age children only; we can, however, restrict attention to adults between 35 and 45 years of age. At the individual level, our proxy for selection is therefore years of education standardized by the (gender-specific) average and standard deviation in the country of origin. At the country level, we simply take the average of this measure.

In Figure C.2 we plot these country of origin-level averages against the PISA score of native students in those countries. Similarly to the PISA sample, we find a weakly negative pattern, suggesting that parents from high PISA countries are somewhat more negatively selected. In Table C.3 we check whether this pattern arises also when we include commuting zone fixed effects: for both mothers and fathers, the coefficients are negative and not statistically different from 0.

[9]Other determinants of selection considered in Albornoz et al. (2012) are the absolute skill premia in host and source countries and migration costs. As discussed above, the available evidence on these dimensions suggests that, if anything, we should expect parents emigrating from high PISA countries to be relatively negatively selected.

[10]Following Barro and Lee (2013), we impute a duration of 4 years for tertiary education in all countries.

Table C.3: Selection - US Census

	Dependent Variable: Standardized Years of Education	
	[1] Mothers	[2] Fathers
Score Country *m*	-0.282	
	(0.478)	
Score Country *f*		-0.009
		(0.393)
N	52875	46254
R Squared	0.06	0.08
Year FE	Yes	Yes
Comm Zone FE	Yes	Yes

Notes: The Table shows results for emigrant mothers in specification (1) and emigrant fathers in specification (2). The dependent variable is years of education standardized by the average and standard deviation of mothers' (specification 1) and fathers' (specification 2) years of education in the country of origin. *Score Country m* and *Score Country f* are the average math PISA scores of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and the father, across all available waves. All specifications control for intercept and wave fixed effect. Standard errors are clustered by mother's country of origin in specification (1) and by father's country of origin in specifications (2). * denotes significance at 10%, ** at 5%, *** at 1%.

One of the concerns highlighted in Section 5 was that years of education are not necessarily pre-determined with respect to migration, and parents might have acquired more or less education as a consequence of their migration decision (and, importantly, for our purposes, might have done so differentially from different countries of origin). We can make some progress in testing this hypothesis by analyzig selection patterns for parents that completed their education in their home country, since for those individuals the relative quality of the US school system should have not played any role in their education choices (and therefore education is more likely to represent a good proxy of pre-determined skills). Figure C.3 and Table C.4 are the counterparts of Figure C.2 and Table C.3 when the sample is restricted only to parents more likely to have completed their education before migrating to the US (see Section 7 for a description of how these parents are identified based on the available information). For both mothers and fathers, the pattern of differential selection is weakly negative with respect to the average PISA score, and not very different from the one obtained in the full sample.

Table C.4: Selection - US Census (Parents Entirely Educated in Home Country)

	Dependent Variable: Standardized Years of Education	
	[1] Mothers	[2] Fathers
Score Country *m*	-0.182	
	(0.357)	
Score Country *f*		-0.111
		(0.299)
N	29851	27070
R Squared	0.07	0.09
Year FE	Yes	Yes
Comm Zone FE	Yes	Yes

Notes: The Table shows results for emigrant mothers in specification (1) and emigrant fathers in specification (2). In all specifications, the sample includes only cases where the parent was entirely educated in his or her home country. The dependent variable is years of education standardized by the average and standard deviation of mothers' (specification 1) and fathers' (specification 2) years of education in the country of origin. *Score Country m* and *Score Country f* are the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and the father, across all available waves. All specifications control for intercept and wave fixed effect. Standard errors are clustered by mother's country of origin in specification (1) and by father's country of origin in specification (2). * denotes significance at 10%, ** at 5%, *** at 1%.

Still, parents might have based their educational choices based on their future relocation to the US, and perhaps this might bias the pattern of selection differentially across countries. It would be worrying if differential selection turned out to be negative only for those parents for whom migration is likely to have played a bigger role in their educational choices, and perhaps positive for the rest of the sample. To check for this possibility, in Table C.5 we present results from specifications where we interact the average PISA score in the country of origin with the number of years between education completion and migration (still restricting the sample to parents entirely educated in their home country). The underlying idea is that the more time has passed between education completion and migration, the less is likely that educational choices were made taking future relocation into account, and the closer we get to the ideal situation where education truly reflects skills pre-determined with respect to migration. For both mothers and fathers, the coefficient on the interaction term is positive but not statistically significant, and its magnitude is so small that the pattern of differential selection would not be positive and significant for any gap between education completion and migration observed in the sample. This result gives us some further confidence that our findings on selection are not driven by a differential effect of migration on parental education.

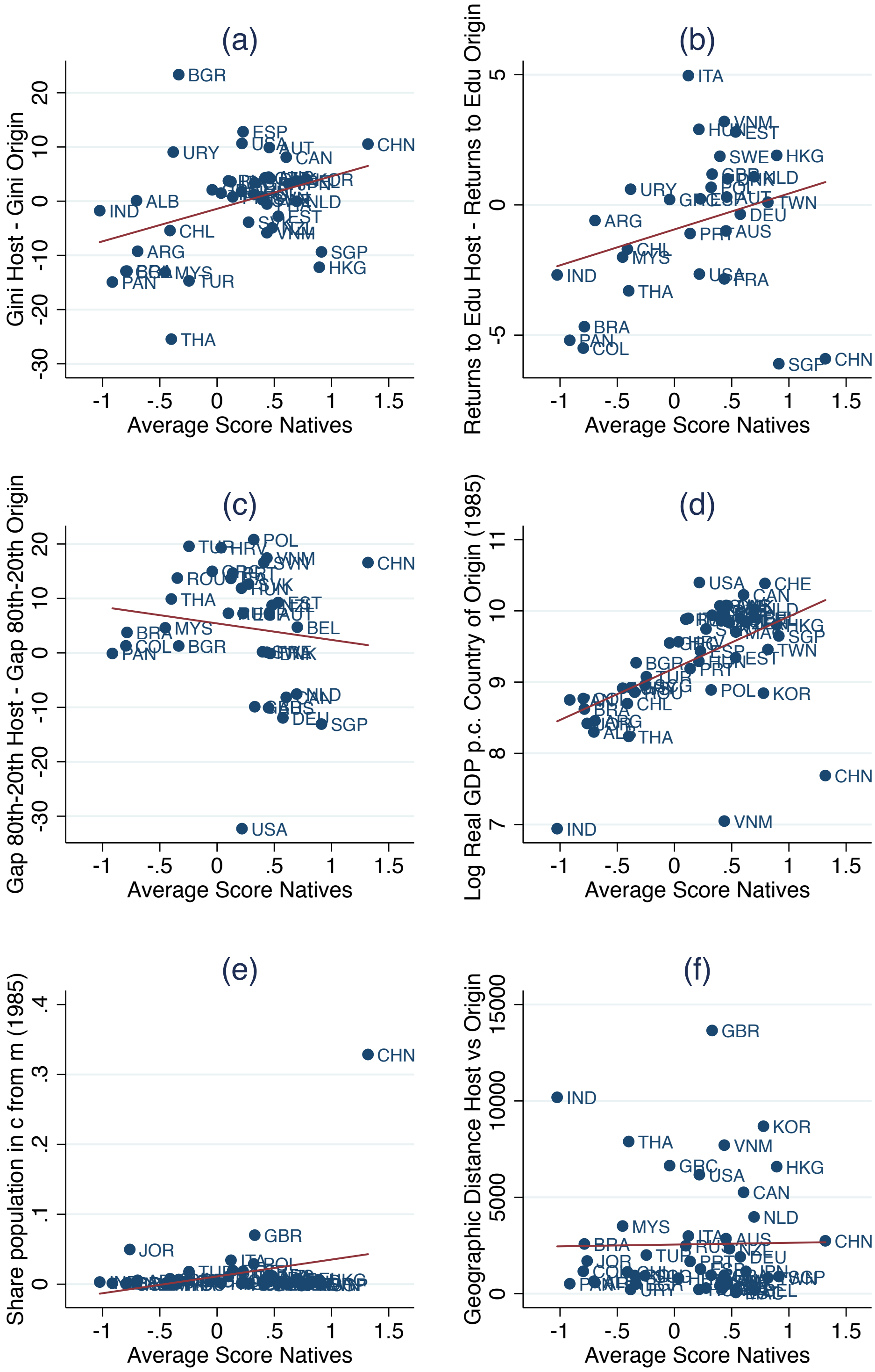
Table C.5: Selection - Heterogeneity with respect to Education Completion and Migration Dates

	Dependent Variable: Standardized Years of Education	
	[1] Mothers	[2] Fathers
Score Country *m*	-0.330	
	(0.497)	
Score Country *m* × Years betw Edu and Migration Mother	0.006	
	(0.018)	
Years betw Edu and Migration Mother	-0.056***	
	(0.006)	
Score Country *f*		-0.248
		(0.381)
Score Country *f* × Years betw Edu and Migration Father		0.004
		(0.012)
Years betw Edu and Migration Father		-0.053***
		(0.005)
N	29851	27070
R Squared	0.16	0.22
Year FE	Yes	Yes
Comm Zone FE	Yes	Yes

Notes: The Table shows results for emigrant mothers in specification (1) and emigrant fathers in specification (2). In all specifications, the sample includes only cases where the parent was entirely educated in his or her home country. The dependent variable is years of education standardized by the average and standard deviation of mothers' (specification 1) and fathers' (specification 2) years of education in the country of origin. *Score Country m* and *Score Country f* are the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and the father, across all available waves. *Years betw Edu and Migration* refers to the number of years occurred between education completion (imputed from the educational attainment) and migration to the US. All specifications control for intercept and wave fixed effect. Standard errors are clustered by mother's country of origin in specification (1) and by father's country of origin in specification (2). * denotes significance at 10%, ** at 5%, *** at 1%.

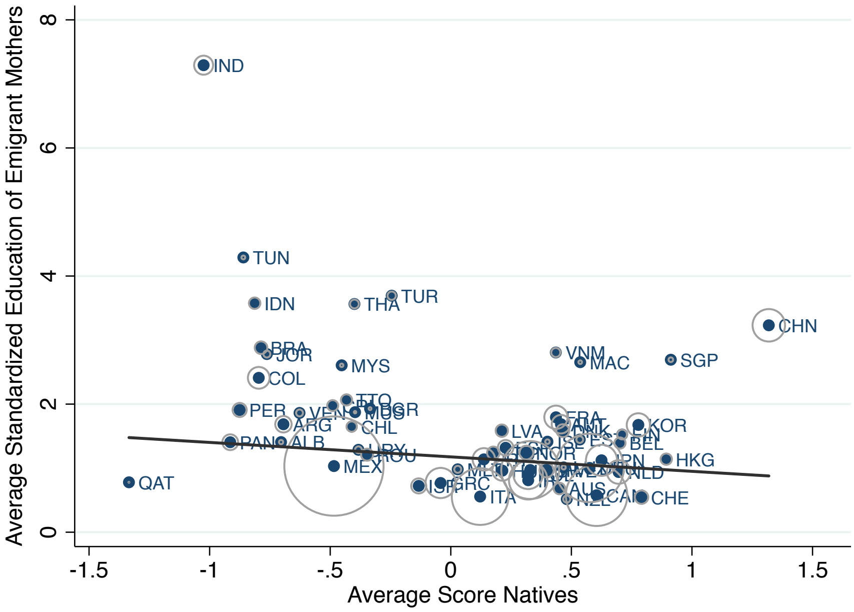
Overall, patterns of selection are on average remarkably similar between the PISA and the Census data. China, however, represents an interesting exception: while Chinese parents appear to be the most negatively selected in the PISA data, they turn out to be above the average country in terms of positive selection in the Census. This latter result is consistent with Feliciano (2005), which shows that Chinese immigrants in the US are among the most positively selected in terms of education across countries of origin. The discrepancy between the two datasets is easily explained by the fact that the native population from China in PISA only includes people from Shanghai, which are substantially more educated compared to the rest of China. Therefore, while Chinese emigrants are positively selected compared to Chinese stayers as a whole, they are still negatively selected compared to stayers observed in the PISA sample.

Figure C.1: Possible Determinants of Emigrants' Selection



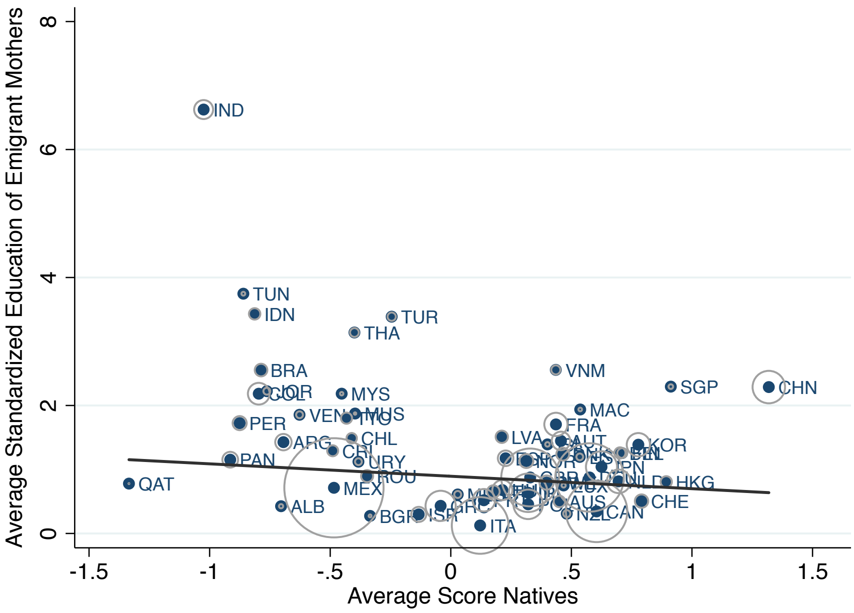
Notes: Each Panel plots the relationship between the average score among natives and a possible determinat of emigrants' selection. Panel (a) plots the difference between the average Gini Index faced by emigrants from country m in their respective host countries and the Gini Index in country m. Similarly, Panels (b) and (c) plot the difference between the average value faced by emigrants from country m and country m's value for the estimated return to education and the absolute income gap between the 80^{th} and the 20^{th} percentiles (in thousands of 2000 US dollars). Panel (d) plots the log real GDP per capita in 1985. Panel (e) and (f) plot the average across emigrants from m of the share of host country population born in country m and of the geographic distance between the host country and country m (in kilometers). The lines show the best linear fits.

Figure C.2: Selection on Parental Education



Notes: The Figure plots the average years of schooling of emigrant mothers from country m standardized by the average and the standard deviation of years of schooling of non-emigrant mothers in country m (y-axis) against the average PISA score of native students in country m (x-axis). The line shows the best (weighted) linear fit.

Figure C.3: Selection on Parental Education (Mothers Entirely Educated in Home Country)



Notes: The Figure plots the average years of schooling of emigrant mothers from country m standardized by the average and the standard deviation of years of schooling of non-emigrant mothers in country m (y-axis) against the average PISA score of native students in country m (x-axis). The sample includes only mothers entirely educated in their home country. The line shows the best (weighted) linear fit.

D Additional Evidence on Mechanisms

D.1 Alternative Measures of Immigrants' Assimilation

In this section we show interaction results using different proxies for parents' integration in their host country. While we focus on years since migration in the main text, immigrants' assimilation is a complex process involving cultural, economic, linguistic and relational transitions, some of which might not be well captured by our proxy. At the same time, some dimensions of parents' assimilation considered here might be direct outcomes of their propensity to human capital accumulation, making the results harder to interpret.

We consider two sets of alternative measures. In Table D.1 we look at intermarriage with natives, which has been widely used as proxy for immigrants' assimilation (Gordon, 1964; Pagnini and Morgan, 1990), and is usually associated with favourable economic outcomes (Furtado and Trejo, 2013). In column 1 we interact a dummy identifying native fathers with T^m, and find that indeed the mother's country-of-origin effect is weaker when the father is a native. Column 2 shows that this patter is robust to the introduction of the other interaction terms explored in the main text. A possible complication arises from the fact that in these specifications we are not considering the father's country of origin, and if mothers from high PISA countries not matched to natives are systematically paired with fathers from high PISA country (and, chiefly, from their own same country), then the omission of a proxy for fathers' influence might explain the negative interaction. To explore this, in columns 3 and 4 we add to the previous specifications the average score from the father's country of origin. While the magnitude of our interaction of interest is unaffected, the coefficients are no longer statistically significant.

Table D.1: Heterogeneity with respect to Intermarriage with Natives

	Dependent Variable: No Grade Repeated			
	[1]	[2]	[3]	[4]
Score Country *m*	0.043***	0.138***	0.043***	0.149***
	(0.008)	(0.029)	(0.015)	(0.026)
Native Father × Score Country *m*	-0.031***	-0.019***	-0.032	-0.023
	(0.008)	(0.006)	(0.020)	(0.018)
Female	0.067***	0.067***	0.067***	0.067***
	(0.003)	(0.003)	(0.003)	(0.003)
Yrs Schooling Father	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Yrs Schooling Mother	0.006***	0.006***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Yrs Since Migr Mother	0.001*	0.002***	0.001*	0.002***
	(0.001)	(0.000)	(0.001)	(0.000)
Log Family Income	0.032***	0.032***	0.029***	0.029***
	(0.008)	(0.008)	(0.008)	(0.008)
Score Country *m* × Yrs Since Migr Mother		-0.002***		-0.002***
		(0.001)		(0.001)
Score Country *m* × Yrs Schooling Mother		-0.006***		-0.006***
		(0.002)		(0.002)
Score Country *f*			-0.000	-0.006
			(0.018)	(0.018)
N	53081	53081	51428	51428
# Country *m*	61	61	61	61
R Squared	0.10	0.11	0.10	0.11
Comm Zone FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. All specifications control for intercept, child age dummies, parents' age, family size, log family income, year fixed effect, (year-specific) quarter of birth fixed effect and father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

In Table D.2 we consider instead the measures of linguistic and cultural distances already discussed in the main text. The idea here is that parents which are linguistically or culturally far from the US norms are less likely to integrate, and perhaps to adapt to the locally prevalent practices and values in terms of children's education. Column 1 shows that indeed the gap between second generation immigrants from high and low PISA countries is larger when parents are linguistically far from the US, and column 2 confirms that this differential effect is robust to the inclusion of our baseline interactions. Results on cultural distance, while of the expected sign, are not statistically different from 0 (columns 3 and 4).

Table D.2: Heterogeneity with respect to Linguistic and Cultural Distance

	Dependent Variable: No Grade Repeated			
	[1]	[2]	[3]	[4]
Score Country *m*	0.033***	0.137***	0.032***	0.143***
	(0.006)	(0.028)	(0.007)	(0.027)
Score Country *m* × Mother Linguistic Distance	0.020***	0.017***		
	(0.006)	(0.006)		
Mother Linguistic Distance	-0.003	-0.002		
	(0.003)	(0.003)		
Father Linguistic Distance	-0.004*	-0.003		
	(0.002)	(0.002)		
Score Country *m* × Mother Cultural Distance			0.002	0.006
			(0.013)	(0.012)
Mother Cultural Distance			0.013	0.009
			(0.008)	(0.007)
Father Cultural Distance			-0.012***	-0.011***
			(0.003)	(0.003)
Female	0.067***	0.067***	0.067***	0.067***
	(0.003)	(0.003)	(0.003)	(0.003)
Yrs Schooling Father	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
Yrs Schooling Mother	0.007***	0.007***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Yrs Since Migr Mother	0.001	0.001***	0.001	0.001***
	(0.001)	(0.000)	(0.001)	(0.000)
Log Family Income	0.030***	0.029***	0.030***	0.030***
	(0.008)	(0.008)	(0.008)	(0.008)
Score Country *m* × Yrs Since Migr Mother		-0.002***		-0.002***
		(0.001)		(0.001)
Score Country *m* × Yrs Schooling Mother		-0.006***		-0.006***
		(0.002)		(0.002)
N	51420	51420	49438	49438
# Country *m*	60	60	47	47
R Squared	0.10	0.11	0.10	0.11
Comm Zone FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. *Linguistic Distance* and *Cultural Distance* are standardized to take mean 0 and standard deviation 1 across all country pairs in the PISA sample (sources are discussed in the paper). All specifications control for intercept, child age dummies, parents' age, family size, log family income, year fixed effect, (year-specific) quarter of birth fixed effect and father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

Overall, the results in this section reinforce the message that the more immigrant parents integrate in the US the more the school performance of their children becomes similar across different countries of origin.

D.2 Immigrants' Ethnic Network

Throughout the paper, we stress the role of parents in the transmission of human capital, and we focus on second generation immigrants in order to fix the characteristics of the local environment. A potential complication arises from the fact that immigrant parents from different nationalities are likely to be differentially exposed to their own ethnic network, even within the same host country or region. The transmission of country-specific skills or attitudes towards education could also take place thorugh this channel, and the objective of this section is to investigate this possibility.

D.2.1 Borjas' Ethnic Capital

In his seminal work, Borjas (1992) uses data from the General Social Survey (GSS) and the National Longitudinal Survey of Youth (NLSY) to argue that the average level of education in the ethnic environment of parents, what he calls "ethnic capital", plays a role in the human capital accumulation process of the following generations in the US. To the extent that second generation immigrants from high-scoring countries are exposed to higher ethnic capital, this could represent a factor behind their superior performance at school additional to any direct interaction with their parents.

We use the Census data to construct a measure of the average years of education of parents of school-age children for each commuting zone and country of origin.[11] In Table D.3, we add this measure of ethnic capital as a control to our baseline specifications, shown in columns 1 and 3. No matter whether commuting zone fixed effects are introduced (column 4) or not (column 2), the coefficient on ethnic capital is positive and significant, consistently with Borjas' result. The coefficient on the average score of natives in the mother's country of origin is somewhat smaller in magnitude, but still positive and significant.

[11] We consider children between 8 and 15 years of age, consistently with the criterion used for our baseline sample. The results are similar when we use the same measure constructed at the state or the country level.

Table D.3: Ethnic Capital

	Dependent Variable: No Grade Repeated			
	[1]	[2]	[3]	[4]
Score Country *m*	0.045***	0.030***	0.028***	0.022**
	(0.013)	(0.008)	(0.009)	(0.009)
Ethnic Capital		0.008***		0.003**
		(0.002)		(0.002)
Female	0.068***	0.068***	0.068***	0.067***
	(0.003)	(0.003)	(0.003)	(0.003)
Mother Sec Edu	0.053***	0.041***	0.047***	0.043***
	(0.012)	(0.012)	(0.011)	(0.013)
Mother Ter Edu	0.057***	0.037***	0.054***	0.046***
	(0.012)	(0.013)	(0.010)	(0.012)
Father Sec Edu	0.042***	0.036***	0.036***	0.034***
	(0.013)	(0.012)	(0.010)	(0.010)
Father Ter Edu	0.061***	0.051***	0.058***	0.054***
	(0.015)	(0.012)	(0.011)	(0.011)
Log Family Income	0.042***	0.039***	0.036***	0.035***
	(0.010)	(0.009)	(0.008)	(0.008)
N	53081	53081	53081	53081
# Country *m*	61	61	61	61
R Squared	0.09	0.09	0.10	0.10
Comm Zone FE	No	No	Yes	Yes
Years Since Migr Mother	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. *Ethnic Capital* is the average years of education of all parents of 8- to 15-year-old children in the same commuting zone and born in country m. All specifications control for intercept, child age dummies, parents' age, family size, year fixed effect, (year-specific) quarter of birth fixed effect and father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

D.2.2 Horizontal and Obliqual Transmission

Parents' ethnic network could also play a role in the transmission of country-specific skills and attitudes towards education. Indeed, the literature on cultural transmission stresses the distinction between the vertical (parents to children), horizontal (children to children) and obliqual (other children's parents to children) transmission of cultural traits, all of which have been shown to be active in different settings (Bisin and Verdier, 2010). In the paper we stress the vertical channel, but to what extent does the performance of second generation immigrants also reflect the horizontal or obliqual ones?

To shed some light on this issue, we exploit the variation across commuting zones in the level of segregation across countries of origin. As discussed by Fernandez and Fogli (2009), local communities with a large share of individuals with the same ancestry might offer more opportunities for the horizontal or obliqual transmission of values through direct interaction, role models and punishments for behaviours not consistent with the social norm. If these channels are important, we would expect a larger country of origin-effect for parents located in such communities, as opposed to more isolated parents.

In Table D.4 we augment the baseline specification (shown in column 1) with interaction terms between T^m and measures of commuting zone-level segregation by country of origin. In particular, we consider the share of all (column 2), 35- to 45-year-old (column 3) and 8- to 15-year-old (column 4) residents born in country m. The coefficients for the interaction terms are positive for all specifications but marginally significant only for the second measure. Moreover, from the coefficient on T^m we can see that in all cases virtually the whole effect persists when the size of the ethnic network approaches zero. The gap in performance is therefore strong even when we focus on rather isolated parents, suggesting that our focus on the vertical channel of transmission might be well-warranted.

The results of Table D.4 should be interpreted with a caveat in mind. Several contributions to the cultural transmission literature argue that the prevalence of given cultural traits in the local context affects the incentives parents face when socializing their children, and that, depending on the setting, vertical and non-vertical (horizontal or obliqual) transmission might be either cultural substitutes or complements (Bisin and Verdier, 2010). Under cultural substitutability, it might be that parents that value education the most play a more active role in shaping human capital accumulation of their children when the horizontal transmission of positive attitudes towards education is muted, to same extent invalidating our interpretation of the results in Table D.4.

Table D.4: Heterogeneity with respect to the Segregation Rate

	Dependent Variable: No Grade Repeated			
	[1]	[2]	[3]	[4]
Score Country *m*	0.028***	0.025**	0.023**	0.027***
	(0.009)	(0.009)	(0.010)	(0.010)
Female	0.068***	0.068***	0.068***	0.068***
	(0.003)	(0.003)	(0.003)	(0.003)
Mother Sec Edu	0.047***	0.048***	0.048***	0.047***
	(0.011)	(0.012)	(0.012)	(0.012)
Mother Ter Edu	0.054***	0.054***	0.055***	0.054***
	(0.010)	(0.010)	(0.010)	(0.010)
Father Sec Edu	0.036***	0.036***	0.037***	0.036***
	(0.010)	(0.011)	(0.011)	(0.011)
Father Ter Edu	0.058***	0.059***	0.060***	0.058***
	(0.011)	(0.013)	(0.013)	(0.013)
Log Family Income	0.036***	0.036***	0.036***	0.036***
	(0.008)	(0.009)	(0.009)	(0.009)
Score Country *m* × Share from *m* (%)		0.006		
		(0.007)		
Share from *m*		0.003		
		(0.004)		
Score Country *m* × Share 35-45 from *m* (%)			0.009*	
			(0.005)	
Share 35-45 from *m*			0.006*	
			(0.003)	
Score Country *m* × Share 8-15 from *m* (%)				0.003
				(0.006)
Share 8-15 from *m*				0.002
				(0.005)
N	53081	53081	53081	53081
# Country *m*	61	61	61	61
R Squared	0.10	0.10	0.10	0.10
Comm Zone FE	Yes	Yes	Yes	Yes
Years Since Migr Mother	Yes	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. *Share from m*, *Share 35-45 from m* and *Share mothers from m* are, within the commuting zone of each student, the shares of, respectively, all residents, residents aged 35 to 45 and residents aged 8 to 15 born in country *m* (in percent). All specifications control for intercept, child age dummies, parents' age, number of siblings, log family income, year fixed effect, (year-specific) quarter of birth fixed effect and father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

In Table D.5 we explore this possibility by turning to the Time Use data, where we can observe a proxy for a parental $input$ for children's human capital accumulation. In particular, we ask whether the gap across nationalities in the time parents spend with their children is smaller when parents live in a community with a larger ethnic network from their country of origin, as it would be implied by cultural substitution. We construct a measure of segregation at the State and country of origin level, given by the share of residents in each State born in each country of origin (results with alternative measures of segregation along the lines of Table D.4 are very similar and available upon request). We then add to our baseline specification an interaction between this measure and T^p, the average score of natives in the country of origin of the interviewed parent. The results show that, if anything, the interaction term is positive, implying that parents spend more time with their children when living in a more segregated State. This finding is not consistent with a cultural substitution story, and provides further support for the fact that vertical transmission plays a key role.

Table D.5: Time Use - Heterogeneity with respect to the Segregation Rate

	Total	Educational	Recreational	Basic
	[1]	[2]	[3]	[4]
Score Country *p*	3.283	1.564	0.153	1.566
	(5.773)	(1.326)	(2.493)	(3.300)
Share from *p*	1.594	-0.007	1.702	-0.101
	(2.100)	(0.490)	(1.157)	(1.156)
Score Country *p* × Share from *p* (%)	4.510	0.359	4.073	0.078
	(4.621)	(0.959)	(2.488)	(2.430)
Parent Sec Edu	-3.009	4.285***	-3.132	-4.162**
	(5.111)	(0.540)	(3.079)	(2.014)
Parent Ter Edu	2.894	3.361***	-3.031	2.564
	(3.044)	(1.102)	(2.006)	(1.865)
Spouse Sec Edu	2.741	-1.911**	6.383**	-1.730
	(3.198)	(0.785)	(2.682)	(1.385)
Spouse Ter Edu	11.895***	2.079	6.890***	2.926
	(3.181)	(1.646)	(2.561)	(2.530)
Log Family Income	5.899***	0.597	-1.515	6.817***
	(2.102)	(0.658)	(0.977)	(1.314)
N	5659	5659	5659	5659
# Country *p*	59	59	59	59
Mean Dep. Var.	89.87	10.53	22.27	57.07
St. Dev. Dep. Var.	119.98	32.30	58.06	88.63
R Squared	0.24	0.06	0.10	0.22
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: The sample includes only immigrant parents of children with at most 18 years. $Parent$ refers to the interviewed parent, $Spouse$ to the other one; $Mother$ is 1 when the interview parent is the mother. $Total$ refers to the total time spent in child care activities, while $Educational$, $Recreational$ and $Basic$ refer to the sub-categories defined in the text. *Score Country p* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the interviewed parent, across all available waves. *Share from p* is the share of residents in the state where each parent lives born in country p (in percent). All specifications control for parents' age, number of children, number of male children, chidren's average age, years since migration, dummies for native spouses and for retired, full time students and disabled parents. Standard errors are clustered by the interviewed parent's country of origin. * denotes significance at 10%, ** at 5%, *** at 1%.

E Complementarities

E.1 Complementarities between School and Parental Influence

Both our reduced form evidence and decomposition exercise are based on a specification where school inputs (as proxied by school fixed effects) and parental influence (as proxied by either the average score in parents' country of origin or country of origin fixed effects) are additively separable. However, a reasonable alternative would be having an interaction between these variables, capturing patterns of complementarity or substitutability between these two kinds of inputs. Were these interactions quantitatively important, the matching pattern between parents and schools of different "qualities" would become potentially important in explaining cross-country differences in the average PISA score, substantially complicating our counterfactual analysis.[12]

To assess the importance of this possibility, we allow for an interaction between the quality of school and parental inputs in our baseline reduced form specification. In particular, we use the average score among students with native parents in a given school as a proxy for school quality, and we ask whether the difference in performance between second generation immigrants from high and low PISA countries varies as a function of school quality.

Table E.1 shows our results. We find that the interaction term is small in magnitude and not significantly different from 0, no matter whether we use the school and country of origin PISA scores as baseline controls (column 2) or whether we absorb those in school and country of origin fixed effects. Moreover, the coefficient on T^m and the R^2 are virtually unaffected by the introduction of the interaction term (columns 2 and 4), suggesting that the linear specification is not missing much in terms of the fitting of the data.

[12]For example, in the case of complementarity between schooling and parental inputs, countries with a more assortative matching between parents and schools would obtain higher average scores.

Table E.1: Complementarities between School and Parental Influence - Reduced Form Results

	Dependent Variable: Math Test Score				
	[1]	[2]	[3]	[4]	[5]
Score Country *m*	0.327***	0.328***	0.243***	0.272***	
	(0.054)	(0.065)	(0.066)	(0.080)	
Score School *s*	0.777***	0.777***			
	(0.023)	(0.023)			
Score Country *m* × Score School *s*		-0.003		-0.041	0.016
		(0.038)		(0.042)	(0.035)
Female	-0.173***	-0.173***	-0.200***	-0.200***	-0.196***
	(0.019)	(0.019)	(0.022)	(0.022)	(0.023)
Mother Sec Edu	0.025	0.025	-0.013	-0.013	-0.029
	(0.035)	(0.035)	(0.039)	(0.039)	(0.034)
Mother Ter Edu	0.057*	0.057*	-0.006	-0.007	-0.030
	(0.036)	(0.036)	(0.042)	(0.042)	(0.040)
Father Sec Edu	0.000	0.000	0.009	0.009	0.007
	(0.031)	(0.031)	(0.022)	(0.022)	(0.022)
Father Ter Edu	0.005	0.005	0.041*	0.041*	0.037
	(0.039)	(0.038)	(0.031)	(0.031)	(0.031)
Mother Working × ISEI	0.004***	0.004***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Father Working × ISEI	0.004***	0.004***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	39573	39573	39573	39573	39573
# Country *m*	49	49	49	49	49
R Squared	0.51	0.51	0.67	0.67	0.67
Country *m* FE	No	No	No	No	Yes
School FE	No	No	Yes	Yes	Yes

Notes: The Table shows results for second generation immigrants on the mother's side. The sample includes only cases where both parents report a country of origin and the country of origin of the mother runs a PISA test on natives. *Score Country m* is the average math PISA score of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother, across all available waves. *Score School s* is the average math PISA score of students with both native parents in school s. All specifications control for intercept, students' age (in months), wave fixed effect and a dummy for father's immigrant status. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

E.2 Complementarities between Maternal and Paternal Influence

Another possible form of complementarity is the one between the influence exterted by mothers and fathers. To explore this issue, in Table E.2 we focus on the sample of second generation immigrants with both parents born abroad, and augment the baseline specifications (reported in columns 1 and 3) with an interaction term between the average score in the mother's and father's country of origin. According to the results in columns 2 and 4, there seems some degree of positive complementarity between maternal and paternal influence, even though the coefficient of the interaction is not statistically significant.

Table E.2: Complementarities between Maternal and Paternal Influence - Reduced Form Results

	Dependent Variable: Math Test Score			
	[1]	[2]	[3]	[4]
Score Country *m*	0.273**	0.253**	0.181**	0.177*
	(0.107)	(0.116)	(0.089)	(0.091)
Score Country *f*	0.229*	0.187	0.017	0.005
	(0.123)	(0.140)	(0.090)	(0.096)
Score Country $m \times$ Score Country f		0.164		0.043
		(0.129)		(0.057)
Female	-0.142***	-0.140***	-0.205***	-0.205***
	(0.038)	(0.038)	(0.026)	(0.026)
Mother Sec Edu	0.079*	0.083*	-0.006	-0.006
	(0.044)	(0.045)	(0.044)	(0.044)
Mother Ter Edu	0.107*	0.112*	-0.032	-0.032
	(0.063)	(0.065)	(0.056)	(0.056)
Father Sec Edu	0.035	0.036	0.002	0.001
	(0.039)	(0.040)	(0.022)	(0.023)
Father Ter Edu	0.112**	0.107**	0.041	0.040
	(0.047)	(0.046)	(0.032)	(0.032)
Mother Working $\times$ ISEI	0.007***	0.007***	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Father Working $\times$ ISEI	0.006***	0.006***	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
N	25534	25534	25534	25534
# Country *m*	47	47	47	47
# Country *f*	47	47	47	47
R Squared	0.31	0.31	0.70	0.70
Host Country FE	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes

Notes: The Table shows results for second generation immigrants on both the mother's and the father's side. The sample includes only cases where both parents report a country of origin that runs a PISA test on natives. *Score Country m* and *Score Country f* are the average math PISA scores of natives (standardized to have mean 0 and standard deviation 1 across all countries participating to the test) in the country of birth of the mother and the father, across all available waves. All specifications control for intercept, students' age (in months) and wave fixed effect. Observations are weighted according to the provided sample weights. Standard errors are clustered by mother's country of origin, and inflated by the estimated measurement error in test scores. * denotes significance at 10%, ** at 5%, *** at 1%.

Ignoring this form of complementarity could lead to an understatement of the quantitative importance of parental influence for cross-country gaps in human capital. We investigate this possibility by replicating our decomposition exercise on the sub-sample of students which have parents of the same nationality. In this case, the country of origin fixed effect captures the combined effect of maternal and parental influence, including any complementarity between the two. Notably, this specification accomodates also the possibility that the degree of complementarity is country-of-origin-specific.

We estimate the regression

$$T^p_{icst} = \beta' ParentsEdu^p_{icst} + \lambda' ParentsOcc^p_{icst} + \gamma^p + \theta^p NatParents^p_{icst} + \\ + \rho' X^p_{icst} + \alpha_{cs} + \alpha_t + u^p_{icst} \tag{1}$$

where the superscript p identifies the country of origin of both parents, γ^p is a parental country of origin fixed effect and $NatParents^p_{icst}$ is a dummy that takes value 1 for native parents. The parental component for country c is defined as

$$Parents^c = \gamma^p + \beta' \overline{ParentsEdu}^p + \lambda' \overline{ParentsOcc}^p \tag{2}$$

where $\overline{ParentsEdu}^p$ and $\overline{ParentsOcc}^p$ are within-country averages across native parents.

Table E.3 displays the resulting $V_{Parents}$ and V_{FE}, as defined in the main text. Consistently with the results in Table E.2, incorporating this form of complementarity amplifies the importance of parental influence. This is especially true for the specification with host country fixed effects, suggesting that the complementarity between mothers and fathers might be operative with particular relevance on the margin of school choice.

Table E.3: Decomposition with Complementarities between Maternal and Paternal Influence - Cross-Country Variance

	$V_{Parents}$ (%)		V_{FE} (%)	
	School FE	Host Country FE	School FE	Host Country FE
Unadjusted	24.24	51.61	23.84	49.87
Adjusted	20.96	48.36	20.58	46.72

Notes: The Table shows the ratio (in percent) between the cross-country variance of either the whole parental component ($V_{Parents}$) or the country specific intercept (V_{FE}) and the cross-country variance of the average math PISA score of natives. The sample only includes students with parents of the same nationality. Columns denoted by *School FE* (*Host Country FE*) refer to specifications that include school fixed effects (host country fixed effects). Adjusted variances are computed by subtracting the average squared standard errors (constructed using the provided replicate weights, and inflated by the estimated measurement error in test scores).

F Development Accounting

In this section we investigate the implications of our results in terms of cross-country differences in output per worker. We start by decomposing the covariance between log GDP per worker and test scores between the contributions of the parental component and residual test scores; that is, we calculate

$$\frac{Cov(\log y_c, Parents^c)}{Cov(\log y_c, T^c)}$$

where y_c, is real GDP per worker, $Parents^c$ is the parental component as defined in the main text and T^c is the average performance of native students in country c. Table F.1 shows the results for both the school and host country fixed effects specifications. In both cases, around 12% of the covariance is driven by the parental component. The fact that this figure is smaller than the relative variance displayed in Table V in the main text reflects that the parental component, besides varying less across countries compared to the residual test score, is also slightly less correlated with other factors of production. In other words, parental influence is more equally distributed across rich and poor countries relative to other forms of physical and human capital.

Table F.1: Decomposition Results - Covariance with GDP per worker

	School FE	Host Country FE
$\frac{Cov(\log y_c, Parents^c)}{Cov(\log y_c, T^c)}$ (%)	11.60	12.79

Notes: The Table shows the ratio (in percent) between the cross-country covariance between log GDP per worker and the parental component and the cross-country covariance between log GDP per worker and average test scores of native students. Columns denoted by *School FE* (*Host Country FE*) refer to specifications that include school fixed effects (host country fixed effects).

To have a sense of the absolute magnitude of the contribution of each component, we implement a simple development accounting exercise. We follow Klenow and Rodríguez-Clare (1997) and much of the literature in postulating an aggregate Cobb-Douglas production function which can be written in per worker terms as:

$$y_c = A_c \left(\frac{k_c}{y_c}\right)^{\frac{\alpha}{1-\alpha}} h_c$$

This formulation allows an additive decomposition of the variance of $\log y_c$ into the contributions of the covariances between $\log y_c$ and the appropriately weighted covariances of the logs of TFP, capital to output ratio and human capital. We are interested in the magnitude and the composition of the latter term,

$$\frac{Cov(\log y_c, \log h_c)}{Var(\log y_c)}$$

which represents our measure of the overall contribution of human capital. We assume that human capital per worker is given by the exponential form

$$h_c = \exp\{\beta_s s_c + \beta_t T^c\}$$

where s_c is average years of schooling in country c. For our baseline, we follow Hanushek and Woessmann (2012) in setting $\beta_s = 0.1$ and $\beta_t = 0.2$, which are picked to match estimates of the returns to schooling and test performance in the labour market.

Table F.2 shows that the baseline measure of human capital accounts for 29% of the variation in GDP per capita. Differences in years of schooling are responsible for 19% of the variance, while test scores account for the remaining 10%. Out of this 10%, a little bit more of a percentage point is to be attributed to the parental component.

Table F.2: Development Accounting - Results

	Baseline	$\beta_t = 0$	$\beta_s = 0$	$\beta_s = 0, T^c = Parents^c$	
				School FE	Host Country FE
$\frac{Cov(\log y_c, \log h_c)}{Var(\log y_c)}$ (%)	29.32	19.39	9.93	1.15	1.27

Notes: The Table shows the ratio (in percent) between the cross-country covariance between log GDP per worker and log human capital per worker and the variance of log GDP per worker. Each column corresponds to a different specification for h_c. Columns denoted by *School FE* (*Host Country FE*) refer to specifications that include school fixed effects (host country fixed effects).

Overall, the results in this section suggest that parental influence does not account for a sizable portion of the cross-country variation in GDP. This is due to the joint effect of two patterns: first, the variation in parental influence is not large enough and, second, parental influence does not covary strongly with other factors of production.

References

Albornoz, Facundo, Antonio Cabrales, and Esther Hauk, "Immigration and the school system," UC3M Working papers 12-03 January 2012.

Barro, Robert J. and Jong Wha Lee, "A New Data Set of Educational Attainment in the World, 1950-2010," *Journal of Development Economics*, 2013, *104* (C), 184–198.

Belot, Michèle V. K. and Timothy J. Hatton, "Immigrant Selection in the OECD," *The Scandinavian Journal of Economics*, 2012, *114* (4), 1105–1128.

Bisin, Alberto and Thierry Verdier, "The Economics of Cultural Transmission and Socialization," NBER Working Papers 16512, National Bureau of Economic Research November 2010.

Borjas, George J., "Self-Selection and the Earnings of Immigrants," *The American Economic Review*, 1987, *77* (4), 531–553.

Borjas, George J, "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics*, 1992, *107* (1), 123–50.

Brueckner, Markus and Daniel Lederman, "Effects of income inequality on aggregate output," Policy Research Working Paper Series 7317, The World Bank June 2015.

Chiquiar, Daniel and Gordon H. Hanson, "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 2005, *113* (2), 239–281.

Chiswick, Barry R., "Are Immigrants Favorably Self-Selected? An Economic Analysis," IZA Discussion Papers 131, Institute for the Study of Labor March 2000.

Feliciano, Cynthia, "Does Selective Migration Matter? Explaining Ethnic Disparities in Educational Attainment among Immigrants' Children," *International Migration Review*, 2005, *39* (4), 841–871.

Fernandez, Raquel and Alessandra Fogli, "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 146–77.

Furtado, Delia and Stephen J. Trejo, "Interethnic marriages and their economic effects," in "International Handbook on the Economics of Migration" Chapters, Edward Elgar Publishing, 2013, chapter 15, pp. 276–292.

Gordon, Milton, *Assimilation in American Life*, Oxford University Press, 1964.

Grogger, Jeffrey and Gordon H. Hanson, "Income maximization and the selection and sorting of international migrants," *Journal of Development Economics*, May 2011, *95* (1), 42–57.

Hanushek, Eric A. and Ludger Woessmann, "Schooling, educational achievement, and the Latin American growth puzzle," *Journal of Development Economics*, 2012, *99* (2), 497–512.

Isphording, Ingo E., Marc Piopiunik, and Núria Rodríguez-Planas, "Speaking in numbers: The effect of reading performance on math performance among immigrants," *Economics Letters*, 2016, *139* (C), 52–56.

Klenow, Peter and Andrés Rodríguez-Clare, "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?," in "NBER Macroeconomics Annual 1997, Volume 12" NBER Chapters, National Bureau of Economic Research, May 1997, pp. 73–114.

Little, Roderick and Donald Rubin, *Statistical Analysis with Missing Data*, Wiley & Sons, 1987.

Mayer, Thierry and Soledad Zignago, "Notes on CEPII's distances measures: The GeoDist database," Working Papers 2011-25, CEPII Research Center December 2011.

McKenzie, David and Hillel Rapoport, "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks," *The Review of Economics and Statistics*, November 2010, *92* (4), 811–821.

OECD, "PISA Data Analysis Manual: SPSS," Technical Report 2009.

Ozden, Cağlar, Christopher R. Parsons, Maurice Schiff, and Terrie L. Walmsley, "Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960–2000," *World Bank Economic Review*, 2011, *25* (1), 12–56.

Pagnini, Deanna and Philip Morgan, "Intermarriage and Social Distance Among U.S. Immigrants at the Turn of the Century," *American Journal of Sociology*, 1990, *96* (2), 405–432.

Psacharopoulos, George and Harry Anthony Patrinos, "Returns to investment in education: a further update," *Education Economics*, 2004, *12* (2), 111–134.

Spolaore, Enrico and Romain Wacziarg, "Ancestry, Language and Culture," CESifo Working Paper Series 5388, CESifo Group Munich 2015.

Søren, Eric W. Holman Wichmann and Cecil H. Brown, "The ASJP Database (version 17)," Technical Report 2016.