I Z A

Speaking in Numbers:
The Effect of Reading Performance on
Math Performance among Immigrants

Ingo E. Isphording
Marc Piopiunik
Núria Rodríguez-Planas

Speaking in Numbers:
The Effect of Reading Performance on
Math Performance among Immigrants

Ingo E. Isphording
IZA

Marc Piopiunik
Ifo Institute at the University of Munich

Núria Rodríguez-Planas
*City University of New York (CUNY), Queens College
and IZA*

ABSTRACT

Speaking in Numbers: The Effect of Reading Performance on Math Performance among Immigrants[*]

This paper is the first to estimate a causal effect of immigrant students' reading performance on their math performance. To overcome endogeneity issues due to unobserved ability, we apply an IV approach exploiting variation in age-at-arrival and the linguistic distance between origin and destination country languages. Using four PISA waves, we find a strong influence of reading performance on math performance, highlighting the importance of early language support for immigrants for their educational career.

Corresponding author:

Ingo E. Isphording
Institute for the Study of Labor (IZA)
Schaumburg-Lippe-Str. 5-9
53113 Bonn
Germany
E-mail: isphording@iza.org

---

1 Introduction

With international migration on the rise, the growing number of school-aged immigrants has raised concerns about their well-documented school performance disadvantages and the subsequent implications for future labor market assimilation around the world. Proficiency in the destination country language deserves special attention when searching for the roots of these disadvantages because language proficiency is not only an important target of the educational process, but also key in the acquisition of further skills (Dustmann and Glitz, 2011; Akresh and Akresh, 2011). Against this background, this paper is the first to estimate the causal effect of reading performance on math performance among immigrants. Assessing the influence of reading on math skills provides important insights into the reasons behind the poor math performance of immigrant students (Dustmann and Glitz, 2011) and offers support for the hypothesis that language skills are a requisite for acquiring other types of skills.

Naïve correlations between reading and math test scores are likely driven to some extent by common unobservable factors such as ability or motivation. To identify a causal effect of reading on math performance, we exploit quasi-experimental variation by comparing immigrants with differential age-at-arrival and linguistic origin. We assume that differences in age-at-arrival patterns by linguistic background affect math performance only because of differences in the costs of acquiring reading skills. Immigrant children learn a new language easier and faster the linguistically closer their mother tongue is and the younger their age is at migration. This provides an exogenous source of variation in the destination country language skills, which is uncorrelated with unobservable math determinants such as ability and motivation.

Specifically, we use the interaction between age-at-arrival and linguistic distance between origin and destination country languages as an instrument for the reading skills of 15-year-old first-generation immigrants surveyed in the Programme for International Student Assessment (PISA). Our identification strategy builds upon the approach by Bleakley and Chin (2004), enhancing it by using multiple destination countries and a continuous measure of linguistic distance.

The strong effect of reading on math skills that we uncover in our analysis indicates significant gains and externalities of supporting the language acquisition of immigrant children. Our results suggest that providing language training for immigrant children immediately after arrival significantly facilitates the acquisition of other skills that are equally important for labor-market success (Hanushek et al., 2015).

2   Data

To assess the importance of reading skills for acquiring math skills, we combine internationally comparable student performance data for first-generation migrants with a unique cardinal measure of the linguistic distance between the migrant's origin and destination country languages.

Data on students' math and reading performance, socioeconomic background, and school characteristics come from the PISA assessments in 2003, 2006, 2009, and 2012, conducted by the OECD. Our sample contains all first-generation migrants, given that the migrant's origin country is observed. To ensure that our empirical strategy works properly, we restrict our sample to (1) origin countries for which we observe emigrants in at least two different destination countries, (2) destination countries in which we observe immigrants from at least two different origin countries, and (3) origin country-destination country cells with at least 10 students. Our estimation sample includes 11,582 first-generation migrants from 47 different origin countries, living in 16 different destination countries.[1]

We combine the PISA data with a unique measure of linguistic distance. This measure is based on an algorithm comparing the pronunciations of common words and has been shown to be an excellent summary statistic of language differences in vocabularies, phonetic inventories, grammars, and scripts (Petroni and Serva, 2010). The closest linguistic distance (apart from native speakers such as Britons in Australia)

---

[1]Destination countries are Australia, Austria, Belgium, China, the Czech Republic, Denmark, Finland, Germany, Great Britain, Ireland, Israel, Latvia, Luxembourg, the Netherlands, New Zealand, and Switzerland.

is experienced by Belgian and German migrants in the Netherlands and by Ukrainian migrants in the Czech Republic. The largest linguistic difference in our sample is faced by Vietnamese migrants in Australia and by Turkish migrants in Denmark.

As linguistic differences between populations may not be orthogonal to other dimensions of international differences, we isolate the effect of linguistic differences on students' reading skills by additionally controlling for cultural differences (using information on genetic distances between populations following Spolaore and Wacziarg (2009)) and for the geographic distance between the capitals of the origin and destination country. Summary statistics of the key variables are reported in Table A-1 in the web appendix. Table A-2 reports summary statistics by destination country.

3 Empirical Strategy

Naïve (partial) correlations between reading and math test scores do not provide insights into how reading performance affects math performance as both types of skills are driven by unobservable common factors such as ability and motivation. We achieve identification of a causal effect of reading performance on math performance by using a quasi-experiment, comparing migrants with different ages-at-arrival and linguistic distance between their origin and destination country languages. Immigrating at later ages increases the costs (in terms of effort) of becoming proficient in the destination country language, with a distinctive structural break in early adolescence commonly referred to as the "critical period" (Ohinata and van Ours, 2012). This effect of age-at-arrival on the language acquisition strongly differs with the linguistic distance between the mother tongue of the migrant and the language in the destination country: the larger the linguistic distance the more detrimental is a late arrival to acquiring the new language (Isphording, 2014). Hence, we estimate a two-stage least squares model. In the first stage, we estimate the following OLS equation:

$$reading_{ijkt} = \gamma_0 + \gamma_1 AAA_{ijkt} + \gamma_2 LD_{ijkt} + \gamma_3 \{AAA_{ijkt} \times LD_{ijkt}\} + \mathbf{X}_{ijkt}\rho + \tau_j + \tau_k + \tau_t + \omega_{ijkt}$$

where $reading_{ijkt}$ is the reading performance of immigrant i from origin country j living in destination country k in PISA wave t. AAA_{ijkt} is the migrant's age-at-arrival to the destination country k, LD_{ijkt} is the linguistic distance between the migrant's origin and destination country languages, $\{AAA_{ijkt} \times LD_{ijkt}\}$ is the interaction between both variables, $\mathbf{X}_{ijkt}$ is a vector of sociodemographic controls (explained below), and τ_j, τ_k, and τ_t are origin country, destination country, and PISA wave fixed effects.

The coefficient on the interaction between age-at-arrival and linguistic distance on immigrants' reading performance, γ_3, captures the variation in immigrants' reading competencies that can be attributed to differences in the age-at-immigration across different linguistic distances. These differences are unrelated to differences in unobserved factors such as ability, motivation, or family background that are known to affect math performance. Thus, our key identification assumption is that differences in the age-at-arrival effect on the reading performance can solely be attributed to differences in the linguistic distance and have no direct effect on the math performance.[2]

It is important to note that neither age-at-arrival nor linguistic distance are used as instruments for migrants' reading performance since both variables might have a direct effect on migrants' math performance (through earlier socialization or different migrant selection). However, exploiting the *interaction* between age-at-arrival and linguistic distance as our identifying variable allows us to isolate variation that affects language acquisition but likely has no direct impact on migrants' math performance. Given that this identification assumption holds, the interaction term $\{AAA_{ijkt} \times LD_{ijkt}\}$ is uncorrelated with unobservable determinants of math performance. Therefore, we can use the predicted reading performance in the second stage to obtain an unbiased estimate of β_1, the causal effect of reading performance on math performance, as follows:

---

[2]Conditional means of the socio-economic background by language background and age-at-arrival (Table A-3 in the web appendix) support this assumption since there are no systematic differences in the age-at-arrival gradients of the socio-economic background between native-speaking and non-native-speaking migrants.

$$math_{ijkt} = \beta_0 + \beta_1 \widehat{reading}_{ijkt} + \beta_2 AAA_{ijkt} + \beta_3 LD_{ijkt} + \mathbf{X}_{ijkt}\gamma + \tau_j + \tau_k + \tau_t + \mu_{ijkt}$$

In the estimations, we successively add three sets of control variables contained in the vector $\mathbf{X}_{ijkt}$. First, we include bilateral measures of geographic and cultural distance to isolate the effect of linguistic distance from cultural differences. Second, we control for students' characteristics and family background, including age, gender, highest parental education, parents' occupational status, and number of books at home. Third, we control for school characteristics such as location, enrollment, private versus public ownership, teacher shortage, math instruction time, and measures of school autonomy.

4 Results

Figure 1 displays conditional means of reading and math performance by age-at-arrival cohorts for native-speaking and non-native-speaking immigrants. It suggests a close relationship between reading and math performance. Both skills decrease with age-at-arrival. This pattern is, however, much more pronounced for non-native-speaking immigrants, especially for immigrants arriving between ages 13 and 16, coinciding with the "'critical period"' in early adolescence. While differences in reading performance are likely driven by higher costs of language acquisition for non-native-speaking immigrants, the analogous pattern for math skills suggests an in-depth relationship between the two skills.

OLS regressions in Table 1 indicate very strong partial correlations between reading and math performance of a magnitude of 0.78, which are robust to different sets of control variables. As argued above, these partial correlations likely reflect both the causal effect of reading performance on math performance as well as the influence of unobservable factors such as ability and motivation.

To isolate the effect of reading performance, we turn to the IV estimates in Table 2,

using the interaction between linguistic distance and age-at-arrival as the identifying instrument. As expected, the IV coefficients are distinctively smaller. Increasing the reading performance by one standard deviation raises the math performance by 0.57 standard deviations in our preferred specification with full controls. Thus, the results indicate a severe upward ability bias of the OLS coefficient, which exceeds the IV coefficient by about 36%. To put it differently: About three quarters of the partial correlation between reading and math performance can be attributed to a causal effect of the former on the latter. The IV results are robust to different sets of parental and school controls, increasing the confidence that the variation in linguistic distance times age-at-arrival is indeed as good as random.

5 Conclusion

The steadily growing number of school-aged immigrants raises concerns about the impact of language hurdles on their educational performance and, hence, their future assimilation in the destination countries. Simple conditional correlations between immigrants' reading and math performance indicate a strong positive relationship. Using four PISA waves, we exploit quasi-experimental variation in age-at-arrival and linguistic background to disentangle the causal effect of reading performance on the math performance of 15-year-old first-generation migrants from common unobserved influences such as ability and motivation.

We find that raising a migrant's reading performance by one standard deviation improves her math performance by 0.57 standard deviations. Naïve partial correlations exceed this causal effect by about 36%, indicating a substantial upward bias due to unobserved ability and motivation. These findings highlight the importance of reading skills in the educational process and point to important gains and externalities of early language support for immigrants.

References

Akresh, R. and Akresh, R. (2011). Using achievement tests to measure language assimilation and language bias among the children of immigrants. *Journal of Human Resources* 46: 647–667.

Bleakley, H. and Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *The Review of Economics and Statistics* 86: 481–496.

Dustmann, C. and Glitz, A. (2011). Migration and education. In Hanushek, E. A., Machin, S. and Woessmann, L. (eds), *Handbook of the Economics of Education*, *4*, chap. 4, 327–441.

Hanushek, E. A., Schwerdt, G., Wiederhold, S. and Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review* 73: 103–130.

Isphording, I. E. (2014). Disadvantages of linguistic origin - Evidence from immigrant literacy scores. *Economics Letters* 123: 236–239.

Ohinata, A. and Ours, J. C. van (2012). Young immigrant children and their educational attainment. *Economics Letters* 116: 288–290.

Petroni, F. and Serva, M. (2010). Measures of lexical distance between languages. *Physica A: Statistical Mechanics and its Applications* 389: 2280–2283.

Spolaore, E. and Wacziarg, R. (2009). The diffusion of development. *The Quarterly Journal of Economics* 124: 469–529.

Figures and Tables

Figure 1: Reading Performance and Math Performance
by Age-at-Arrival and Language Background
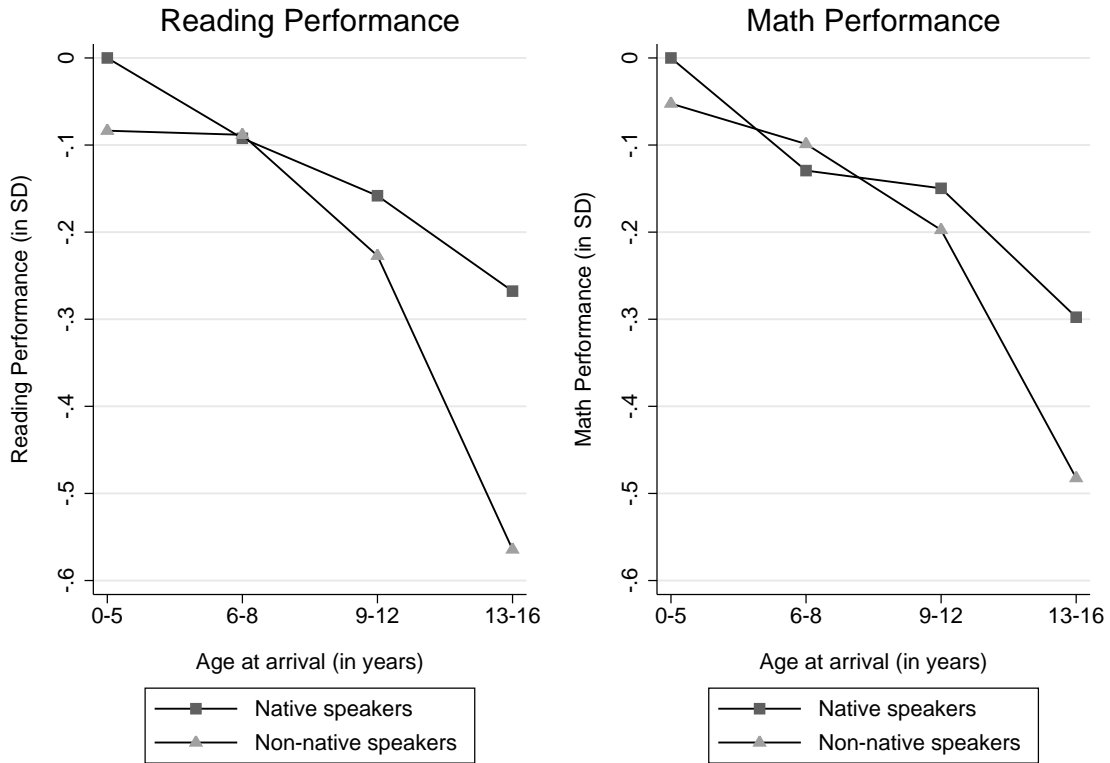
Notes: This figure displays average reading and math scores by age-at-arrival and language background (native speaker yes/no) of first-generation immigrants, conditional on origin and destination country fixed effects.
Data sources: PISA 2003, 2006, 2009, and 2012.

Table 1: Relationship Between Reading Performance and Math Performance (OLS Model)

	Math Performance		
	(1)	(2)	(3)
Reading performance	0.771***	0.781***	0.776***
	(0.010)	(0.009)	(0.009)
Linguistic distance	0.007	0.008	0.006
	(0.021)	(0.018)	(0.019)
Arrived at age 6 – 8	−0.041***	−0.031**	−0.032**
	(0.015)	(0.014)	(0.014)
Arrived at age 9 – 12	−0.024	−0.004	−0.004
	(0.016)	(0.015)	(0.015)
Arrived at age 13 – 16	−0.067***	−0.033*	−0.029
	(0.021)	(0.018)	(0.018)
Country fixed effects	Yes	Yes	Yes
Geographical and cultural distance	Yes	Yes	Yes
Individual characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Students	11,582	11,582	11,582
Clusters (origin country × destination country)	295	295	295
Adj. R2	0.73	0.78	0.78

Notes: Dependent variable: student PISA test score in math. *Linguistic distance* between the PISA test language (in the destination country) and the majority language of the student's country of birth. *Country fixed effects* include origin country fixed effects and destination country fixed effects. *Geographical distance* is the distance of the capitals between origin and destination country. *Cultural distance* is based on genetic proximity measure between populations (see Spolaore and Wacziarg, 2009). *Individual characteristics* include student age and gender as well as highest education of parents, highest occupational status of parents, and number of books at home. *School characteristics* include school location, indicator for private school, weekly math instruction time, enrollment, shortage of math and language teachers, and three autonomy measures (content, personnel, and budget). Robust standard errors (in parentheses) are clustered at the interaction of origin country times destination country. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.
Data sources: PISA 2003, 2006, 2009, and 2012.

Table 2: The Effect of Reading Performance on Math Performance
(IV Model)

Second stage	Math Performance		
	(1)	(2)	(3)
Reading performance	0.563***	0.541***	0.570***
	(0.147)	(0.151)	(0.147)
Linguistic distance	−0.015	−0.016	−0.011
	(0.028)	(0.027)	(0.025)
Arrived at age 6 − 8	−0.055**	−0.045**	−0.045**
	(0.022)	(0.019)	(0.018)
Arrived at age 9 − 12	−0.061	−0.038	−0.034
	(0.037)	(0.027)	(0.028)
Arrived at age 13 − 16	−0.151**	−0.113**	−0.095**
	(0.071)	(0.052)	(0.048)
First stage	Reading Performance		
Identifying instrument			
Linguistic distance × age-at-arrival	−0.087***	−0.080***	−0.078***
	(0.028)	(0.025)	(0.024)
Linguistic distance	−0.051	−0.051*	−0.035
	(0.035)	(0.028)	(0.025)
Arrived at age 6 − 8	−0.042	−0.015	−0.017
	(0.042)	(0.036)	(0.034)
Arrived at age 9 − 12	−0.123**	−0.061	−0.070*
	(0.050)	(0.040)	(0.038)
Arrived at age 13 − 16	−0.331***	−0.227***	−0.217***
	(0.067)	(0.053)	(0.046)
Country fixed effects	Yes	Yes	Yes
Geographical and cultural distance	Yes	Yes	Yes
Individual characteristics	No	Yes	Yes
School characteristics	No	No	Yes
Instrument F statistic	9.3	9.9	10.7
Students	11,582	11,582	11,582
Clusters (origin country × destination country)	295	295	295

Notes: See notes to Table 1. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.
Data sources: PISA 2003, 2006, 2009, and 2012.

Appendix

Table A-1: Summary Statistics

Variable	Mean	SD	Min	Max	Description
Reading performance	-0.13	1.1	-4.7	3.2	PISA reading score; z-standardized across all students in each PISA wave
Math performance	-0.11	1.1	-3.5	3.2	PISA math score; z-standardized across all students in each PISA wave
Age-at-arrival	6.35	4.7	0.0	16.0	Age-at-arrival in destination country
Linguistic distance	54.70	45.8	0.0	104.1	Linguistic distance b/w test language and predominant language of origin country
Age (in years)	15.79	0.3	15.2	16.3	Age (in years)
Male	0.50	0.5	0.0	1.0	Male student
Geographical distance	4933.47	5947.6	173.0	19147.1	Geographic distance between capitals of test country and country of birth
Cultural distance	360.21	565.4	0.0	2288.0	Cultural distance; see Spolaore and Wacziarg (2009) for details
Number of books at home					
0-10 books	0.16		0	1	
11-25 books	0.17		0	1	
26-100 books	0.28		0	1	
101-200 books	0.16		0	1	
201-500 books	0.14		0	1	
More than 500 books	0.10		0	1	
Highest educational degree of parents					
ISCED 0	0.04		0	1	
ISCED 1	0.07		0	1	
ISCED 2	0.09		0	1	
ISCED 3B,C	0.06		0	1	
ISCED 3A,4	0.19		0	1	
ISCED 5B	0.15		0	1	
ISCED 5A,6	0.40		0	1	
Highest occupational status of parents					
Blue collar-low skilled	0.10		0	1	
Blue collar-high skilled	0.15		0	1	
White collar-low skilled	0.19		0	1	
White collar-high skilled	0.55		0	1	
Number of observations	11,582				

Notes: First-generation migrants only. *Data sources:* PISA 2003, 2006, 2009, and 2012.

Table A-2: Summary Statistics for Student Characteristics and Distance Measures by Destination Country

Destination country	AUS	AUT	BEL	CHE	CHN	CZE	DEU	DNK
Reading performance	0.31	-0.58	-0.22	-0.43	-0.33	-0.01	-0.65	-0.96
	(0.99)	(1.10)	(1.09)	(1.01)	(0.93)	(0.84)	(1.08)	(0.84)
Math performance	0.32	-0.50	-0.25	-0.27	-0.20	0.10	-0.62	-1.04
	(0.98)	(0.99)	(1.08)	(1.05)	(1.15)	(0.77)	(0.97)	(0.77)
Age	15.79	15.82	15.83	15.81	15.73	15.77	15.83	15.76
	(0.28)	(0.29)	(0.29)	(0.28)	(0.26)	(0.30)	(0.28)	(0.29)
Male	0.51	0.50	0.53	0.51	0.44	0.53	0.47	0.45
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Age-at-arrival	7.66	5.73	5.99	6.38	6.26	6.59	4.46	5.11
	(4.78)	(4.50)	(5.29)	(4.89)	(4.97)	(4.00)	(3.99)	(3.14)
Linguistic distance	59.24	68.14	17.19	52.02	78.67	73.42	95.68	98.36
	(49.10)	(43.54)	(35.52)	(41.76)	(42.27)	(18.69)	(3.71)	(3.46)
Geographical distance	12115.82	899.56	625.46	1173.54	5034.42	3637.72	1117.87	4306.42
	(4220.31)	(519.83)	(583.79)	(543.53)	(3216.17)	(3347.70)	(648.07)	(1453.77)
Cultural distance	668.19	320.67	78.23	123.92	609.72	535.67	268.14	455.12
	(712.04)	(346.45)	(230.76)	(217.49)	(433.99)	(659.28)	(335.10)	(298.13)
Number of students	2,267	666	1,015	1,479	50	93	279	313

Destination country	FIN	GBR	IRL	ISR	LUX	LVA	NLD	NZL
Reading performance	-0.66	-0.16	0.27	-0.18	-0.49	-0.02	-0.09	0.48
	(1.10)	(1.14)	(0.89)	(1.10)	(1.16)	(0.90)	(1.07)	(1.00)
Math performance	-0.58	0.09	0.03	-0.23	-0.42	-0.10	0.02	0.51
	(0.99)	(1.12)	(0.86)	(0.99)	(1.08)	(0.83)	(1.03)	(0.94)
Age	15.66	15.77	15.69	15.82	15.83	15.83	15.75	15.76
	(0.30)	(0.27)	(0.29)	(0.29)	(0.29)	(0.29)	(0.31)	(0.29)
Male	0.49	0.56	0.48	0.48	0.49	0.51	0.50	0.54
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Age-at-arrival	6.57	8.36	5.25	5.07	5.75	3.13	3.91	8.12
	(4.26)	(4.76)	(4.11)	(3.88)	(4.37)	(3.56)	(3.69)	(4.44)
Linguistic distance	96.81	99.57	0.00	98.49	61.03	25.43	75.06	57.35
	(0.57)	(1.79)	(0.00)	(6.79)	(34.97)	(36.54)	(25.00)	(49.42)
Geographical distance	2288.64	6994.17	554.53	4300.03	1210.19	845.91	1715.77	14483.33
	(2231.48)	(964.12)	(679.03)	(2852.66)	(642.09)	(3.44)	(1452.60)	(4050.00)
Cultural distance	1141.17	619.93	0.00	217.45	35.33	1066.00	246.17	813.95
	(232.98)	(428.94)	(0.00)	(13.25)	(22.16)	(0.00)	(323.14)	(837.68)
Number of students	406	59	1,060	471	1,915	173	103	1,233

Table A-3: Relative Conditional Means by Age-at-Arrival and Language Background

Age-at-arrival	Native Speakers		Non-native Speakers		$\Delta_{\text{native speakers}} - \Delta_{\text{non-native speakers}}$
	0-8 years	9-16 years	0-8 years	9-16 years	
Reading performance	0	-0.18	-0.08	-0.39	-0.14
		(0.03)	(0.05)	(0.05)	(0.04)
Math performance	0	-0.18	-0.04	-0.30	-0.09
		(0.03)	(0.05)	(0.05)	(0.04)
Parental education	0	-0.03	0.12	0.05	-0.04
		(0.04)	(0.07)	(0.08)	(0.06)
Parental occupation	0	-0.07	0.01	0.02	0.07
		(0.03)	(0.05)	(0.05)	(0.04)
Number of books at home	0	-0.28	-0.14	-0.50	-0.08
		(0.04)	(0.07)	(0.07)	(0.05)
Students	2,927	1,725	4,594	2,336	11,582

Notes: Means of dependent and key control variables by age-at-arrival and language background (native speaker yes/no), relative to base category of native-speaking immigrants who arrived at age 0-8 years. All means conditional on origin and destination country fixed effects. The last column reports the difference-in-differences of the conditional means. Standard errors are reported in parentheses.

Table A-4: Relationship Between Reading Performance and Math Performance (OLS Model): Results on All Covariates

	Math Performance		
	(1)	(2)	(3)
Reading performance	0.771***	0.781***	0.776***
	(0.010)	(0.009)	(0.009)
Linguistic distance	0.007	0.008	0.006
	(0.021)	(0.018)	(0.019)
Arrived at age 6 – 8	−0.041***	−0.031**	−0.032**
	(0.015)	(0.014)	(0.014)
Arrived at age 9 – 12	−0.024	−0.004	−0.004
	(0.016)	(0.015)	(0.015)
Arrived at age 13 – 16	−0.067***	−0.033*	−0.029
	(0.021)	(0.018)	(0.018)
Geographical distance	−0.000**	−0.000*	−0.000
	(0.000)	(0.000)	(0.000)
Cultural distance	0.000**	0.000	0.000
	(0.000)	(0.000)	(0.000)
Age		0.039**	0.042**
		(0.017)	(0.017)
Male		0.422***	0.420***
		(0.010)	(0.010)
Family background			
ISCED 1		0.047*	0.046*
		(0.027)	(0.027)
ISCED 2		0.030	0.026
		(0.031)	(0.031)
ISCED 3B,C		−0.018	−0.017
		(0.032)	(0.032)
ISCED 3A,4		−0.015	−0.015
		(0.025)	(0.025)
ISCED 5B		0.010	0.010
		(0.024)	(0.025)
ISCED 5A,6		0.062**	0.058**
		(0.028)	(0.028)
11-25 books		−0.018	−0.016
		(0.018)	(0.018)
26-100 books		0.047**	0.047**
		(0.019)	(0.019)
101-200 books		0.098***	0.097***
		(0.022)	(0.022)
201-500 books		0.135***	0.132***
		(0.025)	(0.025)
More than 500 books		0.188***	0.188***
		(0.026)	(0.026)
White collar-low skilled		−0.060***	−0.056***
		(0.014)	(0.014)
Blue collar-high skilled		−0.053***	−0.048***
		(0.017)	(0.017)
Blue collar-low skilled		−0.046***	−0.042**
		(0.017)	(0.017)

(continued on next page)

	Math Performance		
	(1)	(2)	(3)
School characteristics			
Small town			0.024
			(0.021)
Town			0.013
			(0.024)
City			0.021
			(0.027)
Large city			0.041
			(0.028)
Private school			0.038**
			(0.017)
School enrollment			0.000***
			(0.000)
Instructional time math			0.032***
			(0.011)
Shortage math teacher: very little			0.011
			(0.016)
Shortage math teacher: to some extent			0.003
			(0.019)
Shortage math teacher: a lot			−0.003
			(0.027)
Shortage language teacher: very little			−0.021
			(0.015)
Shortage language teacher: to some extent			−0.006
			(0.020)
Shortage language teacher: a lot			−0.030
			(0.039)
Content autonomy			0.032
			(0.020)
Personnel autonomy			−0.046**
			(0.022)
Budget autonomy			−0.014
			(0.016)
Constant	−0.125	−1.022***	−1.290***
	(0.101)	(0.291)	(0.316)
Country fixed effects	Yes	Yes	Yes
Students	11,582	11,582	11,582
Clusters (origin country × destination country)	295	295	295
Adj. R2	0.73	0.78	0.78

Notes: The table reports results on all covariates corresponding to the OLS estimations in Table 1. Omitted categories of family background and school characteristics: *parents have no educational degree*; *0-10 books*; *white collar-high skilled*; *village*; *shortage of math (language) teachers: not at all*. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.
Data sources: PISA 2003, 2006, 2009, and 2012.