I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17626

The Bilingual Advantage: It's How You Measure It

Ainoa Aparicio Fenoll
Zoë Kuehn

JANUARY 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17626

The Bilingual Advantage: It's How You Measure It

Ainoa Aparicio Fenoll
University of Turin, Collegio Carlo Alberto and IZA

Zoë Kuehn
Universidad Autónoma de Madrid

JANUARY 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

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

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

www.iza.org

ABSTRACT

The Bilingual Advantage: It's How You Measure It*

We use data on Latino children in the United States who have been randomly assigned calculation tests in English or Spanish to check for the so-called bilingual advantage, the notion that knowing more than one language improves individuals' other cognitive skills. After controlling for different characteristics of children and their parents, as well as children's time in the US, we find a bilingual advantage among children who read or write in English and Spanish but not for those who only speak or understand both languages. In particular, bilingual readers or writers perform one fourth to one third of a standard deviation better than monolingual children, equal to learning gains of an additional school year. Applying the Oster test, we find that selection on unobservables would need to be 3-4 times stronger than selection on observables to explain away our results. The bilingual advantage is stronger among children in two-parent households with siblings and for those at the upper end of the ability distribution.

JEL Classification: I2, I24, J15

Keywords: bilingualism, cognitive skills, measurement, selection on observables and unobservables

Corresponding author:
Ainoa Aparicio Fenoll
University of Turin
Corso Unione Sovietica, 218bis
10134 Torino
Italy

E-mail: ainoa.apariciofenoll@unito.it

* The authors are grateful for suggestions made by Daniela del Boca, Vicente Cuñat, Chris Flinn, Libertad González, Giovanni Mastrobuoni, Mario Pagliero, Barbara Petrongolo, Aleksey Tetenov and two anonymous referees. We thank seminar participants at the WOLFE workshop at York, the CHILD internal seminar, the SAEe Bilbao, and the Alp Pop conference at La Thuile. Zoë Kuehn acknowledges financial support from Ministerio de Ciencia e Innovación (PID2020-112739GA-I00, PID2023-149412NB-I00). Competing interests: The authors declare none.

1 Introduction

In the United States, the share of bilingual individuals has more than doubled over the last 40 years. Using data from the US Census, Figure A-1 in the Appendix shows that their share increased from around 6-9% in 1980 to 12-18% in 2019.[1] According to Eurostat, in 2016 around 25% of individuals in Europe reported to be proficient in at least one foreign language, ranging from 11% in Italy to 66% in Luxembourg. More individuals knowing various languages implies better communication skills which have been shown to foster trade (Melitz and Toubal [2014]) and migration (see e.g. Adserá and Pytliková [2016] or Aparicio Fenoll and Kuehn [2016]). Furthermore, if bilingualism improved individuals' cognitive skills, this could help to explain the positive effects of migrants' home country language proficiency on labor market outcomes or social integration (Bleakley and Chin [2004] and [2010]). Bilingualism could then also have an impact on economic growth through improved cognitive skills (Hanushek and Woessmann [2008]). However, the individual effects of bilingualism are unclear. In particular, the question regarding the existence of the so-called bilingual advantage, the idea that knowing more than one language improves individuals' other cognitive skills is unresolved.[2]

We use data on children of Latino immigrants in the Unites States who have been randomly assigned calculation tests in English or Spanish to check whether bilingual children perform better on these tests which are designed to be unaffected by language proficiency (see Aparicio Fenoll [2018]). After controlling for different characteristics of children and their parents, as well as children's time in the US, we find that bilingual readers or writers perform one fourth to one third of a standard deviation better than monolingual children, equal to learning gains of one additional school year.[3] However,

[1] Similarly, for 1980 García [1984] reports around 10% of bilingual individuals in the US, and in 2011 according to the US Census Bureau [2013] 16.1% of US residents who spoke English very well or well, spoke a language other than English at home.

[2] In part this might explain why bilingual education has been a debated school policy for decades. For instance, in 1998 voters in California passed proposition 227 in favor of English only education, abandoning years of bilingual education, only to repeal proposition 227 in 2016.

[3] According to Woessmann [2016], "the rule of thumb [is] that average student learning in a year is equal to about one-quarter to one-third of a standard deviation." (pg. 3)

this finding does not extend to those who report to merely speak or understand both languages. Given that our estimated bilingual advantage for readers and writers could be due to unobservable characteristics such as school or neighborhood quality, we apply the Oster [2019] test, and we show that this is highly unlikely. Selection on unobservables would have to be between 3 to 4 times stronger than selection on observables to explain away our results. Using scores on other cognitive exams which were also randomly assigned in English or Spanish to children in our sample, we do not find any evidence that the bilingual advantage extents to tests which are not language-free. As Bialystok et al [2010] explain, bilingual children tend to know fewer words in each language compared to monolinguals because "some words occur in a context in which they only use one of their languages." (pg. 3)

In the fields of psychology, neuroscience, and education, the idea that bilingualism can shape individuals' brains beyond language use has been debated for decades. Initially the view that additional languages cause confusion dominated, see e.g. Darcy [1953]. As the first to challenge this notion, Peal and Lambert [1962] reported that bilingual individuals outperformed monolinguals on various cognitive tests. A large and growing number of studies have since provided additional evidence in favor of the bilingual advantage; see for instance Bialystok [1988] or Bialystok et al [2004].[4] However, others such as Paap and Greenberg [2013] or Hilchey and Klein [2011] claim that the positive association between bilingualism and cognitive skills is oftentimes the result of flaws in the design of tests, the failure to control for demographic characteristics, differences across age groups considered, or very small samples. Furthermore, no clear consensus exists as to how to measure bilingualism and language proficiency, see e.g. Surrain and Luk [2019] , DeBruin [2019], or Hulstijn [2012] for inconclusive reviews on this topic.[5] Our findings provide two novel explanations for the presence of a bilingual advantage in some and its absence in other

[4]This positive view has become so widespread that there is even evidence of a publication bias, see De Bruin et al [2015].

[5]One attempt to unify the measurement has been made by Anderson et al [2018] proposing the so-called "Language and Social Background Questionnaire" (LSBQ) which provides a measures on the extent of non-English language proficiency, its use at home and socially. However, we did not find any study on differential effects of bilingualism when measured in terms of writing, reading, speaking, and understanding, which makes the direct comparison of our results with those in literature difficult.

contexts: First, we show that it only arises when classifying individuals according to their reading and writing skills, while bilingual speaking and understanding abilities imply no advantage. Second, we find the bilingual advantage to only show up in cognitive tests which are language-free.

Among the few works that like the current paper consider larger samples and control for demographic characteristics are Mouw and Xie [1999] who find "no evidence that bilingualism has a positive effect on academic achievement." Looking at 832 first and second generation Asian Americans from the 1988 National Educational Longitudinal Study, the authors find that speaking a native language with parents has a temporary positive effect on academic achievement which vanishes once immigrant parents achieve a moderate level of English proficiency. Focusing on academic paths, Han [2012] looks at the Early Childhood Longitudinal Study and shows that children who speak both English and their native language well, start out with lower math scores in kindergarten, but that by fifth grade the math gap between bilingual and monolingual children is fully closed. While similar to both papers we also find no significant differences between bilingual and monolingual speakers, we do estimate a bilingual advantage for readers and writers. Hence the current paper is closely related to Golash-Baza [2005] who measures bilingualism as speaking, understanding, reading and writing well in English and in a native language. The author uses data on more than 3,000 individuals from the 1992/93 and 1995/96 Children of Immigrants Longitudinal Study and finds that bilingualism improves performance on the Stanford Math and Reading tests, but only for individuals in communities with low levels of English proficiency and high levels of resources and networks. This is in line with our heterogeneity results which find the bilingual advantage to be stronger among children at the upper end of the ability distribution and those coming from Mexico, the Latino origin-group with the lowest share of migrants reporting to speak English very well (see PEW Hispanic Center [2007]).

Different from these three papers we are able to isolate the role of bilingualism from the influence of English proficiency per se, because (i) the random assignment of tests in English and Spanish allows us to estimate an effect of being bilingual that is not tied to one particular language, and (ii) we differentiate between monolingual individuals in English and Spanish in our estimations. We standardize test scores by age and our estimations include age dummies. As migrant children move to the US and start attending the US school system it becomes more likely that they turn bilingual and improve their cognitive skills at the same time. To address this issue of assimilation which has been shown to have important effects for the performance on cognitive tests (see Akresh and Akresh [2011]), we also control for children's time in the US.

Our findings contribute to the literature in two ways. First, the fact that only bilingual writers and reader seem to have a cognitive advantage but not speakers or those who only understand English and Spanish suggests that merely picking up a second language will not result in better cognitive skills. This speaks to a potential mechanism behind the bilingual advantage which could lie with the additional effort by children, parents, schools, or teachers required to read and write well in two languages. Second, we show that independently of the language of the test, bilingual readers and writers perform better on language-free calculation tests, and that this difference is hardly explained by unobserved characteristics. The remainder of this paper is organized as follows: The next section presents our data. In Section 3, we describe the methodology used, and Section 4 presents our results. Section 5 concludes.

2 Data

For our analysis we use data from the New Immigrant Survey (NIS). The NIS was conducted among US immigrants who were granted permanent residence between May and November of 2003, including both new arrival immigrants as well as so called adjustee immigrants who were already living in the US on temporary non-migrant visas (or, in some cases, undocumented). Surveyed individuals were selected based on administrative records from the US Immigration and Naturalization Service such as for the data to be nationally representative of new arrival and adjustee immigrants. The survey asked questions about migrants' education, language skills, and employment history, as well as family composition and characteristics. The language of the interview (English or the migrant's native language) was chosen by the respondent.

Two aspects of the NIS are key for our analysis. First, the survey administered four cognitive, so-called Woodcock-Johnson (WJ), tests to children of respondents (ages 6-12).[6] Second, for some children of Spanish-speaking immigrants the language of these tests was randomized (English or Spanish). This is important because our estimated effect of being bilingual will thus not be tied to one particular language. For our main analysis, we use children's performance on the numerical reasoning test, with tasks ranging from simple summations to solving equations with unknowns, see Mather and Jaffe [2002]. We also consider children's performance on the other three WJ tests: Passage Comprehension, Letter-Word Identification, and Applied Problems. Apart from these tests, children were also asked to repeat digits forward and backward. Digits were read to children in the language of their choice. We calculate the sum of correct repetitions of digits. Note that digit span tests were originally designed as part of the so-called Wechsler Intelligence Test within the category "Verbal IQ" to test individuals' working memory.[7] We normalize all tests scores by age to have zero mean and standard deviation equal to one.

Children were also asked to assess their Spanish and English proficiency along four dimensions (understanding, speaking, reading, and writing) on the following scale: very well, well, not well, or not at all. Table 1 displays all possible combinations of language skills. We define bilingual children as those who report to perform in both English and Spanish either well or very well. We define as monolingual Spanish (English) individuals who report to only perform well or very well in Spanish (English). Non-performers are those who neither know English nor Spanish well. They make up less than 1% (12%) of

[6] There exists ample evidence that success especially on these type of tests relates positively to later education and labor market outcomes, see e.g. Davis-Kean et al [2021].

[7] Among children, Hale et al [2002] point out that repeating digits backwards is associated with attention and executive function processes, while repeating digits forwards with short-term rote auditory memory processes.

the sample in case of understanding and speaking (reading and writing). While in general there are fewer individuals who read and write both languages compared to understanding and speaking, Figure 1 shows that the four different definitions of bilingualism are not perfect subsets of one another.

Table 1: Combinations of language skills and bilingualism

<table>
<tr><th></th><th>English very well</th><th>English well</th><th>English not well</th><th>English not at all</th></tr>
<tr><td>Spanish very well
Spanish well</td><td colspan="2">bilingual</td><td colspan="2">monolingual Spanish</td></tr>
<tr><td>Spanish not well
Spanish not at all</td><td colspan="2">monolingual English</td><td colspan="2">non-performers</td></tr>
</table>

For certain robustness checks we also use data by US state for 2003 on median household income by state (in current US $) and per student expenditure on public education from the US Census and the U.S. Department of Education, National Center for Education Statistics respectively. We assign this data to parents' current state of residence.[8]

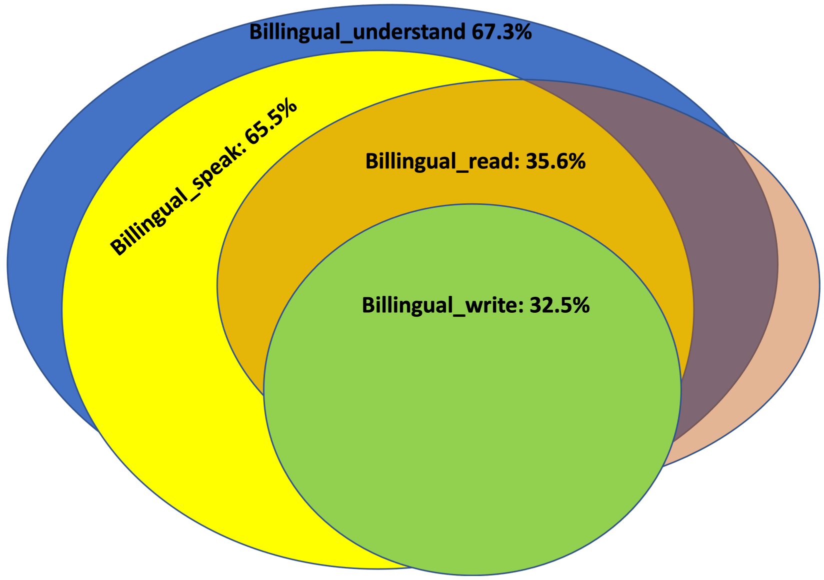
Our sample is composed of all children of Spanish-speaking immigrants aged 6 to 12 who provide information on how well they understand, speak, read, and write English or Spanish, whose language of the test was randomly assigned, and for whom we know their WJ test scores, their correct digits, and for whom we have information on parental characteristics.[9] This leaves us with 652 children.

Table 2 displays the characteristics of children and their parents in our sample. Between 33% and 67% of children in our sample are bilingual and between 15% and 18%

[8]Note that the data only has dis-aggregated information on 6 (large) states: California, Florida, Illinois, New Jersey, New York, and Texas. Individuals in all other states are assigned to US regions (called divisions) of residence. We assign values for these variables by computing population-weighted means (data from US Census) across states belonging to those divisions (excluding the six states mentioned above).

[9]We also exclude from our sample children whose parents interfered with test taking and refused that the child continued with a certain test, because in this case children were automatically assigned a score of 0 which we do not want to confuse with children having obtained a score 0 on the test due to low performance. This restriction reduces initial sample size by around 8%, but once we condition on information regarding language skills, this restriction excludes only 8 observations, less than 1% of the sample, 5 (3) for when the test was administrated in Spanish (English).

Figure 1: Four different definitions of bilingualism



Source: NIS data; sample used in this paper N=652

are monolingual in Spanish, depending on the language dimension used. Around half of them are boys, and they are on average nine years old. Most have been living for approximately seven years in the US. Regarding parental variables, on average mothers and fathers are around 36 and 38 years old, respectively. Mothers and fathers in our sample have received less than 10 years of education on average, and most have acquired their education outside of the US. 86% of parents are married or cohabiting, while around 6-7% are single parents or separated or divorced. 95% of children live with their mother and around 83% with their father.[10] Parents of children in our sample are migrants from 10 different countries, but the majority has arrived from Mexico: 46% in the case of mothers and 36% in the case of fathers.[11] Regarding test scores, children perform best on letter-word identification which is tailored to younger children, applied problems, passage comprehension, and last calculation tests. However, the ranking changes when we standardize test scores by age. Half of the children in our sample were randomly assigned

[10]Sometimes mothers or fathers migrate first with their children and then the other spouse joins later. This is why there is no one-to-one correspondence between parents being married and mother and/or father at home.

[11]The other countries of origin are: Columbia, Cuba, Dominican Republic, El Salvador, Guatemala, Peru, and Puerto Rico.

Table 2: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.
Understands English and Spanish well	0.673	0.469	0	1
Only understands Spanish well	0.187	0.39	0	1
Only understands English well	0.132	0.339	0	1
Bilingual speaker	0.655	0.476	0	1
Only speaks Spanish well	0.196	0.398	0	1
Only speaks English well	0.141	0.348	0	1
Bilingual reader	0.356	0.479	0	1
Only reads Spanish well	0.158	0.365	0	1
Only reads English well	0.373	0.484	0	1
Bilingual writer	0.325	0.469	0	1
Only writes Spanish well	0.152	0.359	0	1
Only writes English well	0.397	0.49	0	1
Boy	0.515	0.5	0	1
Age	8.925	2.026	6	12
Number of siblings	1.785	1.126	0	6
Years in US	7.156	3.563	0	13
Mother's age*	35.52	5.083	22	53
Father's age*	38.203	5.454	23	67
Mother's years of schooling	9.404	3.049	1	21
Father's years of schooling	9.877	2.674	2	30
Mother's years of schooling in US	0.449	1.093	0	11
Father's years of schooling in US	0.513	1.041	0	18
Married Parents	0.807	0.395	0	1
Cohabiting Parents	0.064	0.246	0	1
Separated Parents	0.037	0.188	0	1
Divorced Parents	0.025	0.155	0	1
Single Parent	0.066	0.248	0	1
Widowed Parent	0.002	0.039	0	1
Mother at home	0.948	0.222	0	1
Father at home	0.825	0.38	0	1
Mother Mexican	0.457	0.499	0	1
Father Mexican	0.357	0.48	0	1
WJ: Standardized calculations score	0.079	0.814	-2.352	4.850
WJ: Calculations score	14.96	7.345	0	45
WJ: Standardized passage comprehension score	0.065	0.831	-2.06	2.789
WJ: Passage comprehension score	18.949	9.653	0	47
WJ: Standardized Letter-Word Identification score	0.1	0.895	-2.527	2.738
WJ: Letter-Word Identification score	42.181	20.831	0	76
WJ: Standardized Applied Problems score	0.133	0.823	-2.13	2.64
WJ: Applied Problems score	26.511	11.824	0	57
WJ Tests in Spanish	0.502	0.5	0	1
Standardized correct digits	-0.28	0.845	-2.827	4.65
Correct digits	9.055	3.448	0	28
Median Households income by state/region	46213.96	4620.92	33908.35	56045.00
Per pupil expenditure on public education	9212.44	1534.96	6891.10	13884.09

Notes: NIS data: This table presents averages and standard deviations for the sample used in the estimation. The number of observations is 652. *Data for mothers' and fathers' age (year of birth) is missing for 6.4% and 18.9% of our observations. We replace missing values with mean values by children's ages and include dummy variables indicating missings in our regressions. Data for Median Household income and per public expenditure on education from the US Census and the U.S. Department of Education respectively; for regional data, state data has been weighted by population size of 2003 also from the US Census.

tests in Spanish.[12]

[12]Note that the NIS includes a much larger array of questions, for instance regarding parenting or school environments. However, the large share of missing values in these items make it impossible for us to

Table 3: Differences between Bilingual and Monolingual Readers

Variable	Bilingual	Monolingual	Difference
Standardized calculation score	0.255	0.017	0.238***
Calculation score	17.422	14.341	3.08***
Standardized passage comprehension score	0.275	0.007	0.267***
Passage comprehension score	22.099	18.211	3.888***
Standardized applied problems	0.240	0.087	0.153**
Applied Problems score	29.142	25.680	3.463***
Standardized Letter Word identification score	0.371	0.037	0.334***
Letter Word identification score	49.772	40.763	9.009***
WJ Tests in Spanish	0.461	0.538	-0.076*
Standardized correct digits	-0.185	-0.277	0.092
Correct digits	9.801	9.040	0.761***
Boy	0.483	0.552	-0.069
Age	9.457	8.844	0.613 ***
Number of siblings	1.677	1.864	-0.187*
Years in US	7.858	6.882	0.976***
Mother's age	36.092	35.295	0.797*
Father's age	38.771	38.250	0.520
Mother's years of schooling	9.809	9.251	0.558**
Father's years of schooling	10.018	9.870	0.1408
Mother's years of schooling in US	0.465	0.420	0.044
Father's years of schooling in US	0.533	0.449	0.084
Married Parents	0.815	0.806	0.008
Cohabiting Parents	0.091	0.049	0.041**
Separated Parents	0.026	0.046	-0.020
Divorced Parents	0.034	0.023	0.011
Single Parents	0.030	0.075	-0.045**
Widowed Parents	0.004	0	0.004
Mother at home	0.931	0.960	-0.029
Father at home	0.836	0.815	0.021
Mother is Mexican	0.444	0.448	-0.004
Father is Mexican	0.332	0.358	-0.026
Median Households income by state/region	45972.78	46461.97	- 489.19
Per pupil expenditure on public education	9174.23	9259.86	- 85.63

Notes: This table presents the average observable characteristics of monolingual and bilingual readers who are part of the sample used in the estimation, separately. Here we have excluded non-performers from the sample (11.4%). The last column includes the differences between the two values. *** p<0.01, ** p<0.05, * p<0.1.

Table 3 presents differences between bilingual and monolingual readers in terms of observable characteristics. In line with our main findings, bilingual readers perform significantly better on calculation tests as well as on repeating digits back- and forward; the latter only when not standardized by age. On average, they also perform better on the other three Woodcock-Johnson tests, although this does not hold in our controlled regression analysis. Regarding individual characteristics, bilingual readers are a little older and have lived in the US for longer and have fewer siblings. Their mothers are somewhat older and more educated. Parents of bilingual readers are more likely to be cohabiting

include them here.

and less likely to be single parents. On the other hand, monolingual children are slightly more likely to have taken the test in Spanish compared to bilingual children. We do not observe any differences regarding median household income or public expenditure on education by state of residence.[13]

3 Methodology

We estimate the difference in test scores between bilingual and monolingual children controlling for English and Spanish language proficiency and demographic characteristics of children and their parents as follows:

$$C_i = \beta_0 + \beta_1 B_i + \beta_2 D(L)_i + \beta_3 T_i + \beta_4 B_i \times T_i + \beta_5 D(L)_i \times T_i + \beta_6 D_i + \beta_7 X_i + u_i, \quad (1)$$

where C_i is the test score standardized by age of child i, and B_i is a dummy variable that takes on value one if the child is bilingual. $D(L)_i$ are indicator variables for language proficiency other than bilingualism; i.e. Spanish monolinguals and non-performers. Our reference group are English monolingual children. T_i controls for the language of the test, taking on value one if the test was administered in Spanish. We also control for the interaction of language proficiency (or being bilingual) and the language of the test. Note that the calculation test is understood to be language-free, and hence for our main specification we do not expect coefficients on the test language to be significant. Our controls allow us to separately identify the effect of being more proficient in any language and of speaking two languages. This way we can be sure that our effect of bilingualism is not driven by the fact that those who speak two languages just happen to be more fluent in English or Spanish. In some of our specifications we also control for children's innate ability, including D_i, the standardized correctly repeated digits. Finally, X_i is a vector of children's characteristics (boy, age dummies, number of siblings, time in the US dummies

[13] Tables A1 to A3 in the Appendix provide differences between bilingual and monolingual children using the other three alternative definitions for bilingualism. Across all specifications, bilingual children tend to perform better on most tests, in particular when considering raw test scores. They tend to be older and to have been in the US for longer, and their parents are less likely to be single parents.

and the interaction of age and time in the US dummies), parental characteristics (age, years of schooling, years of schooling in the US, marital status and country of birth dummies). The vector also includes dummies for US state/region of residence.

Note that our coefficient of interest β_1 may be affected by the presence of certain unobservable characteristics which could lead to different cognitive abilities of bilingual individuals but which should not be attributed to bilingualism per se. For instance, bilingual individuals may be those living in neighborhoods or attending schools which are less segregated and also richer allowing them to acquire better English skills and to receive a better education overall. To address this issue, we employ the methodology proposed by Oster [2019]. In particular, we estimate Equation 1 in two different ways. Once, as proposed above and a second time only retaining the controls necessary for the coefficient β_1 to be meaningful. We refer to this last specification as uncontrolled (u) while our main specification is the controlled specification (c). In our uncontrolled specification, we only include test language and language proficiency controls and their interactions. Retaining the estimated coefficients of interest β_{1u} and β_{1c} as well as the R-squared from both estimations (R_u^2 and R_c^2) allows us to compute the following statistic:

$$\delta = \left(\frac{\beta_{1c}}{\beta_{1u} - \beta_{1c}}\right) \times \left(\frac{R_c^2 - R_u^2}{0.3 * R_c^2}\right). \tag{2}$$

The value of δ indicates how many times larger selection on unobservables would have to be compared to selection on observables in order to fully explain away the estimated coefficient.[14]

4 Results

Table 4 shows the results from estimating Equation 1. Column (1) displays our benchmark uncontrolled specification which regresses the standardized calculation test score

[14]Note that this methodology assumes zero correlation between omitted variables and included controls, an assumption which may be too restrictive in our context. For instance, one could image parental education to be correlated with aspects of neighborhood or school quality (see Diegert et al [2023] for a novel method to deal with this issue).

on a bilingual dummy, dummies for language proficiency and a dummy variable for test language as well as its interactions with all language proficiency controls. In column 2, we add children's and parental characteristics including child's years in the US and age, as well as state/region of residence fixed effects. Column 3 includes dummies for each combination of child's years in the US and age and finally, column 4 presents results from our preferred specification that also controls for innate ability using children's standardized correctly repeated digits. Results indicate that bilingual readers perform significantly better on calculation tests than monolinguals. The magnitude of these differences lies between one fourth and one third of a standard deviation. According to Woessmann [2016], this is equal to average learning gains over one school year.

Table 4: Difference in standardized calculation test scores between bilingual and monolingual children

	(1)	(2)	(3)	(4)
Bilingual reader	0.243	0.301	0.29	0.291
	(0.103)**	(0.11)***	(0.116)**	(0.114)**
Individual characteristics	No	Yes	Yes	Yes
Age by Years in US	No	No	Yes	Yes
Innate ability	No	No	No	Yes
Number of observations	652	652	652	652
R-squared	0.044	0.173	0.225	0.251
F statistic	4.232	1.634	1.255	1.436

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. All columns control for language proficiency, test language, and the last two variables interacted. The dependent variable is the standardized calculation test score. Individual characteristics refer to age, years in US, gender, number of siblings, parental age, education, marital status, country of origin and US state/region of residence. Innate ability refers to the age-standardized sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Using results from columns (1) and (4) we calculate parameter δ to be equal to 3.5 This indicates that for the coefficient 0.29 to be explained entirely by selection on unobservables, such selection would need to be more than 3 times larger than the selection on the observables that were added between columns (1) and (4). According to Oster [2019], values above one for δ indicate that unobservable characteristics do not explain the estimated differences.

Tables A4 in the Appendix displays the full estimation results. Apart from being bilingual, mother's years of education, parents being divorced, and higher ability measured by correctly repeated digits also display a significant and positive coefficient. The first result is in line with a large body of literature finding mother's education to matter more for children's academic achievements, see e.g. Rosenzweig and Wolpin [1994] or Carneiro et al [2013]. On the other hand, the result on divorce seems to stand in contrast to existing evidence on a negative effect on educational attainment for children of parental separations or in non-traditional families (see e.g. Garasky [1995], or Astone and Mclanahan [1991] or Haveman and Wolfe [1995] for a review of the literature). Finally regarding innate ability as measured by individuals' sum of correctly repeated digits, there might be some concern that it could be a bad control. The fact that the estimated coefficients hardly differ across columns (4) and (3) which does not include the ability control is reassuring. When we re-calculate Oster's δ using column (3) as our main specification instead, we obtain a similar but somewhat smaller value for δ of 2.7.

We repeat our estimation of column (4) using the three remaining definitions of bilingualism (understanding, speaking and writing in both languages), see Table 5. We find a similar cognitive advantage among bilingual writers but no significant difference in cognitive test scores for bilinguals defined in terms of understanding or speaking both languages.[15]

4.1 Robustness

One important unobservable in our estimation is neighborhood and school quality. While we would ideally control for it, we only know about parental state of residence (and in some cases only US division of residence). Nevertheless, we consider data on median

[15] When bilingualism is defined in terms of writing (reading), values for parameter δ are much higher when the ability control is included 14.1 (3.5), but very similar when not 2.9 (2.7). Tables A5 to A7 in the Appendix display the full results for the three alternative definitions of bilingualism. Note that the negative but insignificant coefficients for bilinguals who understand or speak both languages are most likely due to a peculiar reference group in these regressions, children who having both parents from Spanish-speaking countries but speak or understand only English.

Table 5: Different dimensions of bilingualism and cognitive advantage

	Reads	Understands	Speaks	Writes
Bilingual	0.291	-0.154	-0.063	0.298
	(0.114)**	(0.149)	(0.142)	(0.116)**
Individual characteristics	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes
Number of observations	652	652	652	652
R-squared	0.251	0.234	0.232	0.253
F statistic	1.436	1.312	1.295	1.457

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. All columns control for years in US, age of the child, language proficiency, test language, and the last two variables interacted. The dependent variable is the standardized calculation test score. Individual characteristics refer to age, years in US, gender, number of siblings, parental age, education, marital status, country of origin, and US state/region of residence. Innate ability refers to the age-standardized sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

household income and per pupil spending on public education by these geographical areas (measured in 2003, the year the survey was conducted), and we re-run Table 4 using these controls instead of state/region dummies. Results in Table A8 in the Appendix show that our main coefficients for bilingual reader remain fairly robust.

Note that for ease of interpretation in our benchmark estimation we used the standardized calculation test score as our outcome variable. However, one caveat of this measure is that it has been constructed by age using sample instead of unknown population averages and standard deviations. Hence we test the robustness of our results by using raw test scores. Results in Table 6 are very much in line with those from our main specification. Considering that the standard deviation of the raw calculation test score is 7.3 implies that the estimated coefficients are equal to 0.29 a standard deviation and thus identical to our benchmark results. The same holds true when using our alternative definitions of bilingualism, see Table A9 in the Appendix.

To check that our results do not hinge on the particular functional form imposed by our linear regression we also run a logit model with the outcome variable defined as scoring above the mean of one's age group. Table A10 in the Appendix displays the results. The bilingual advantage among those who understand and speak both languages

Table 6: Difference in raw calculation test scores between bilinguals and monolinguals

	(1)	(2)	(3)	(4)
Bilingual reader	2.954	2.209	2.133	2.131
	(0.901)***	(0.784)***	(0.815)***	(0.804)***
Individual characteristics	No	Yes	Yes	Yes
Age by Years in US	No	No	Yes	Yes
Innate ability	No	No	No	Yes
Number of observations	652	652	652	652
R-squared	0.097	0.486	0.528	0.542
F statistic	9.871	7.375	4.86	5.078

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. All columns control for language proficiency, test language, and the last two interacted. The dependent variable is the raw calculation test score. Individual characteristics refer to age, years in US, gender, number of siblings, parental age, education, marital status, country of origin. Innate ability refers to the raw sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

remains non-existent, among bilingual writers it remains positive and significant while for bilingual readers the sign of the coefficient remains positive but we loose significance, probably due to the reduced sample size caused by perfectly predicted observations being dropped.

Finally, as a type of falsification test, we also check how bilingual children perform on the other three Woodcock-Johnson tests which were administered in this survey and which are not language-free: passage comprehension, applied problems and letter-word identification. Here, we do not expect bilingual children to perform better, because controlling for language of the test, if anything, their language skills are more likely to be lower. Table 7 shows that there are no statistically significant differences between bilingual and monolingual children regarding performance on neither one of those tests; independently of which language dimension is used to define bilingualism (see Table A11 in the Appendix).

4.2 Heterogeneity

Literature has highlighted important gender differences in the performance on math or calculation test scores (e.g. OECD [2019], Guiso et al. [2010] or Anghel, Rodríguez-Planas,

Table 7: Falsification: Difference in scores on language-dependent tests between bilingual and monolingual readers

	(1)	(2)	(3)	(4)
A: Passage Comprehension				
Bilingual reader	0.117	0.114	0.106	0.107
	(0.1)	(0.106)	(0.111)	(0.11)
R-squared	0.135	0.274	0.317	0.325
F statistic	14.36	2.935	2.017	2.07
B: Letter-Word identification				
Bilingual reader	0.214	0.149	0.137	0.139
	(0.104)**	(0.11)	(0.114)	(0.112)
R-squared	0.189	0.317	0.377	0.403
F statistic	21.391	3.626	2.621	2.9
C: Applied Problems				
Bilingual reader	0.068	0.098	0.096	0.097
	(0.102)	(0.109)	(0.113)	(0.113)
R-squared	0.073	0.216	0.273	0.28
F statistic	7.237	2.144	1.627	1.672
Individual characteristics	No	Yes	Yes	Yes
Age by Years in US	No	No	Yes	Yes
Innate ability	No	No	No	Yes
Number of observations	652	652	652	652

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. All columns control for language proficiency, test language, and the last two interacted. The dependent variable in Panel A is the standardized passage comprehension test score, in Panel B the standardized score on the Letter-Word identification test, and in Panel C the standardized score on Applied Problems. Innate ability refers to the standardized sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

and Sanz-de-Galdeano [2020]), as well as regarding the influence of external factors on cognitive development (Rosselli et al [2009]). We hence further explore our results by performing separate regressions for boys and girls. While results in Table 8 show certain gender differences, they are inconclusive. The bilingual advantage among readers is particularly strong and only significant among girls, while among writers the bilingual advantage is only significant and slightly larger for boys. In our sample, similar to most studies, girls are more likely to be bilingual. However, this difference is only significant for the dimensions of speaking and writing. In particular, girls are 9 percentage points more likely to be bilingual writers. Hence, boys who are bilingual writers could simply be more positively selected.

Table 8: Heterogeneity: Bilingualism and cognitive advantage by gender

	Readers		Writers	
	Boy	Girl	Boy	Girl
Bilingual	0.244	0.369	0.302	0.25
	(0.158)	(0.187)**	(0.168)*	(0.181)
Individual characteristics	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes
Number of observations	336	316	336	316
R-squared	0.367	0.406	0.367	0.399
F statistic	1.364	1.534	1.364	1.493

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS separately for boys and girls. All columns control for years in US, age of the child, language proficiency, test language, and the last two interacted. The dependent variable is the standardized calculation test score. Innate ability refers to the age-standardized sum of correctly repeated digits. *** p<0.01, ** p<0.05, * p<0.1.

We also test for differences in the bilingual advantage by parental education, which has been shown to be closely linked to children's educational outcomes (see e.g. Hertz et al [2011]).[16] In our sample, parental education is relatively low, and hence we split the sample into children who have or not at least one parent with more than 10 years of education. Table A13 in the Appendix displays the results. When bilingualism is defined in terms of reading, we only estimate a significant bilingual advantage for children of low educated parents. However, this result is not robust to defining bilingualism along individuals' writing abilities in English and Spanish. In this case, we find significant effects for both, children with high and low educated parents, although the estimated coefficient is slightly larger for the former. Looking at descriptive statistics we observe that children of high educated parents perform significantly better on all tests (when scores are standardized by age), and while they are as likely as children of low educated parents to be bilingual in terms of understanding or speaking, they are 10 percentage points more likely to be bilingual readers and writers. Hence, this suggests that bilingual writers and readers of lower educated parents might be more selected.

[16] Maybe because higher educated parents spend more time with their children (see e.g. Guryan et al [2008]) or children of higher educated parents exert more effort (see e.g Kuehn and Landeras [2014]).

Given that 50% of children in our sample have at least one Mexican parent, we explore whether there are differences in the bilingual advantage between these children and those with parents from other Latin-American countries. In our sample, children with parents from Mexico are more likely to be bilingual speakers compared to those whose parents are not Mexicans. However, there are no statistical differences along the other three language dimensions, see Table A12 in the Appendix. Results in Table 9 show that the bilingual advantage seems to be particularly strong for children with parents from Mexico.

Table 9: Bilingualism and cognitive advantage by parental country of origin

	Bilingual writer		Bilingual reader	
	Mexicans	Non-Mexicans	Mexicans	Non Mexicans
Bilingual	0.342	0.187	0.408	0.095
	(0.151)**	(0.204)	(0.151)***	(0.194)
Individual characteristics	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes
Number of observations	329	323	329	323
R-squared	0.443	0.269	0.44	0.275
F statistic	2.075	0.853	2.045	0.878

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS separately for children with at least one Mexican parent. All columns control for years in US, age of the child, language proficiency, test language, and the last two interacted. The dependent variable is the standardized calculation test score. Innate ability refers to the age-standardized sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Children with Mexican parents have more siblings, they have younger and less educated parents who are more likely to be married and to be living at home and who are less likely to be single. More traditional household structures by Mexican parents could have a positive effect on children's cognitive skills, but also more household members could imply more communication on average. The former notion is in line with findings in literature on the before-mentioned negative effect on educational attainment for children in non-traditional families. To test for this idea we split our sample into children with parents who are married, living together and have at least two children. Table 10 displays the results. The bilingual advantage is particularly strong and significant for children in traditional families.[17]

Table 10: Bilingualism and cognitive advantage by family type

	Bilingual writer		Bilingual reader	
	Traditional	Non-traditional	Traditional	Non-traditional
Bilingual	0.334	0.368	0.296	0.145
	(0.136)**	(0.249)	(0.132)**	(0.246)
Individual characteristics	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes
Number of observations	466	186	466	186
R-squared	0.319	0.563	0.313	0.572
F statistic	1.587	1.621	1.543	1.679

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS separately for children in traditional and non-traditional families, where traditional families are defined as households with married parents, living together and at least two children. All columns control for years in US, age of the child, language proficiency, test language, and the last two interacted. The dependent variable is the standardized calculation test score. Innate ability refers to the age-standardized sum of correctly repeated digits. *** p<0.01, ** p<0.05, * p<0.1.

Another possible explanation for our result of a particularly strong bilingual advantage for children of Mexican parents could be linked to the fact that around the time of the NIS survey individuals from Mexico were least likely to report to be proficient in English among Hispanics in the US (PEW Hispanic Center [2007]). Hence, similar to Golash-Baza [2005], our findings would indicate a stronger bilingual advantage in communities with low levels of English proficiency.

Finally, we run quantile regressions to see if the bilingual advantage changes along children's ability distribution. Results in Table 11 show that the bilingual advantage is larger and only significant for children at the upper end of the ability distribution. In particular, while the coefficient for being bilingual at the lower end of the ability distribution is largest, we only estimate significant coefficients for those at the 50th, 75th and 85th percentile. When defining bilingualism in terms of writing, we already estimate significant coefficient for those at the 25th percentile, see Table A15 in the Appendix. Null results for a bilingual advantage among those who report to only understand or speak

[17] Note that these results are only in part driven by the different sample sizes, see Table A14 in the Appendix for results when we instead define traditional families as married parents living together with at least three children.

both languages also hold along the entire ability distribution.

Table 11: The bilingual advantage: Heterogeneity by ability

	15%	25%	50%	75%	85%
Bilingual reader	0.308	0.203	0.213	0.163	0.226
	(0.228)	(0.124)	(0.113)*	(0.087)*	(0.085)***
Individual characteristics	Yes	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes	Yes
Number of observations	652	652	652	652	652
Pseudo r-squared	0.2986	0.2520	0.1917	0.2201	0.2613

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by quantile regressions. The dependent variable is the standardized reading test score. All columns control for years in US, language proficiency, test language, and age of the child. Innate ability refers to the age-standardized sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

5 Conclusion

Our findings provide support for the existence of a cognitive advantage among Latino migrant children in the US who are bilingual readers or writers but not among those who only speak or understand both languages. We show that the better performance by bilingual readers or writers on cognitive tests, independently of any particular language, is not explained by observables and highly unlikely to be driven by unobserved factors. In particular, for children with parents from Mexico, for those in two-parent households with siblings, and for those at the upper end of the ability distribution, being a bilingual writer or reader is associated with improved cognitive skills.

For potential mechanisms behind the bilingual advantage one typically turns to findings from other disciplines. Neuroscience argues that bilingualism alters the functional involvement of certain brain areas in the performance of tasks (see e.g. Garbin et al [2010], or Rodriguez at al. [2013]). Some studies in sociology sustain that bilingual children have access to positive "cultural capital" in their families and ethnic communities (see e.g. Rumberger [1998]), potentially improving academic performance. Others highlight the importance of speaking the same language as one's family, and how this may enhance the parent-child relationship, which in turn, could have positive implications for children's motivation to succeed academically (e.g. Fuligini [1997] or Tseng and Fuligini [2000]). Our findings provide some suggestive evidence regarding potential mechanisms. The fact that the bilingual advantage is only present among those who read or write both languages but nonexistent among bilingual speakers or those who report to understand English and Spanish, suggests that merely picking up a second language is not enough. In particular, a cognitive advantage might only arise from the additional effort by children, parents, schools, or teachers required to read and write well or very well in two languages.[18]

Unfortunately, our data are not rich enough to further investigate this idea. However, our results suggest that additional efforts to teach migrant children to read and write in their native languages is advisable, and that in particular children in two-parent households with siblings and those at the upper end of the ability distribution could benefit the most from it.[19]

References

Adserà, Alícia and Mariola Pytliková (2016): "Language and migration," in *The Palgrave handbook of economics and language*, edited by Victor Ginsburgh and Shlomo Weber, Palgrave Macmillan, London.

Anderson, John. A.; Mak, Lorinda; Keyvani Chahi, Aram and Ellen Bialystok (2018): "The language and social background questionnaire: Assessing degree of bilingualism in a diverse population," *Behavior research methods*, 50, 250-263.

Brindusa, Anghel; Rodríguez-Planas, Núria and Anna Sanz-de-Galdeano (2010):"Is the math gender gap associated with gender equality? Only in low-income countries," *Economics of Education Review*, 79, 102064.

[18]For instance, currently only six US states (Connecticut, Illinois, New Jersey, New York, Texas, and Wisconsin) require districts to provide bilingual education when a certain number of students (usually 20) from the same language group have enrolled in a district or school, see Rutherford-Quach et al [2021].

[19]Currently some researchers are studying how Covid-19 related lockdown measures which have implied more joint time of children and parents have made it easier to raise a bilingual child, see Harach [2020].

Akresh, Richard and Ilana Redstone Akresh (2011):"Using Achievement Tests to Measure Language Assimilation and Language Bias among the Children of Immigrants," *Journal of Human Resources*, 46(3), 647-667.

Aparicio Fenoll, Ainhoa and Zoë Kuehn (2016): "Does foreign language proficiency foster migration of young individuals within the European Union," in *The economics of language policy*, edited by Bengt-Arne Wickstroem and Michele Gazzola, MIT Press.

Aparicio Fenoll, Ainoa (2018): "English proficiency and mathematics test scores of immigrant children in the US," *Economics of Education Review*, 64, 102-113.

Astone, Nan Marie and Sara S. McLanahan (1991): "'Family Structure, Parental Practices and High School Completion," *American Sociological Review*, 56(3), pp. 309-320.

Bialystok, Ellen; Luk, Gigi; Peets, Kathleen F. and Sujin Yang (2010): "Receptive vocabulary differences in monolingual and bilingual children," *Bilingualism: Language and cognition*, 13(4), 525-531.

Bialystok, Ellen; Craik, Fergus I. M.; Klein, Raymond and Mythili Viswanathan (2004): "Bilingualism, Aging, and Cognitive Control: Evidence From the Simon Task," *Psychology and Aging*, 19(2), 290-303.

Bialystok, Ellen (1988): "Levels of bilingualism and levels of linguistic awareness," *Developmental Psychology*, 24(4), 560-567.

Bleakley, Hoyt and Aimee Chin (2010): "Age at arrival, English proficiency, and social assimilation among US immigrants," *American Economic Journal: Applied Economics*, 2(1), 165-192.

Bleakley, Hoyt and Aimee Chin (2004): "Language skills and earnings: Evidence from childhood immigrants," *Review of Economics and Statistics*, 86(2), 481-496.

Carneiro, Pedro; Meghir, Costas and Matthias Parey (2013): "Maternal Education, Home Environment and the Development of Children and Adolescents, " *Journal of the European Economic Association*, 11, 123-160.

De Bruin, Angela (2019): "Not all bilinguals are the same: A call for more detailed assessments and descriptions of bilingual experiences," *Behavioral Sciences*, 9(3), 33.

De Bruin, Angela; Treccani, Barbara and Sergio Della Sala (2015): "Cognitive advantage in bilingualism: An example of publication bias?," *Psychological Science*, 26(1), 99-107.

Darcy, Natalie T. (1953): "A review of the literature on the effects of bilingualism upon the measurement of intelligence," *The Pedagogical Seminary and Journal of Genetic Psychology*, 82(1), 21-57.

Davis-Kean,Pamela E.; Domina, Thurston; Kuhfeld, Megan;Ellis, Alexa and Elizabeth T. Gershoff (2021): "It matters how you start: Early numeracy mastery predicts high school math course-taking and college attendance," *Infant and Child Development,* 31(2), e2281.

Diegert, Paul; Masten, Matthew A. and Alexandre Poirier (2023): "Assessing Omitted Variable Bias when the Controls are Endogenous," arXiv: 2206.02303v4, econ.EM.

Fuligni, Andrew J. (1997): "The academic achievement of adolescents from immigrant families: The role of family background, attitudes, and behavior," *Child Development,* 68(2), 351-363.

Garasky, Steven (1995): "The Effects of Family Structure on Educational Attainment: Do the Effects Vary by the Age of the Child?," *American Journal of Economics and Sociology,* 54(1), pp. 89-105.

Garbin, Gabrielle; Sanjuan, Ana; Forn, Cristina; Bustamante, Juan Carlos; Rodríguez-Pujadas, Aina; Belloch, Vicente; Hernández, Mireia; Costa, Albert and César Ávila (2010): "Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control," *Neuroimage,* 4, 272-1278.

García, Ofelia (1984): "Bilingualism in the United States: Present Attitudes in the light of past policies," in *The English Language Today: Public Attitudes Toward the English Language,* edited by Sidney Greenbaum, Oxford: Pergamon Press.

Golash-Boza, Tanya (2005): "Assessing the Advantages of Bilingualism for the Children of Immigrants," *International Migration Review,* 39(3), 721-753.

Guiso, Luigi; Monte, Ferdinando; Sapienza, Paola and Luigi Zingales (2010):"Culture, Gender, and Math," *Science,* 320(5880), 1164-1165.

Guryan, Jonathan; Hurst, Erik and Melissa Kearney (2008):"Parental Education and Parental Time with Children," *Journal of Economic Perspectives,* 22(3), 23-46.

Hale, James. B.; Hoeppner, Jo-Ann B. and Catherine A. Fiorello (2002):"Analyzing Digit Span Components for Assessment of Attention Processes,"*Journal of Psychoeducational Assessment,* 20(2), 128-143.

Han, Wen-Jui (2012): "Bilingualism and academic achievement," *Child Development,* 83(1), 300-321.

Hanushek, Eric A. and Ludger Woessmann (2008): "The role of cognitive skills in economic development," *Journal of Economic Literature,* 46(3), 607-68.

Hardach, Sophie (2020): "In Quarantine, Kids Pick Up Parents' Mother Tongues," *New York Times,* September 10, 2020.

Haveman, Robert and Barbara Wolfe (1995): "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature* , 33(4), pp. 1829-1878.

Hertz, Tom; Jayasundera, Tamara; Piraino,Patrizio; Selcuk, Sibel; Smith, Nicole and Alina Verashchagina (2011): "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends," *B.E. Journal of Economic Analysis & Policy,* 7(2).

Hilchey, Matthew D. and Raymond M. Klein (2011): "Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes," *Psychonomic Bulletin & Review,* 18, 625-658.

Hulstijn, Jan H. (2012): "The construct of language proficiency in the study of bilingualism form a cognitive perspective," *Bilingualism: Language and Cognition,* 15(2), 422-433.

Kuehn, Zoë and Pedro Landeras (2014): "The Effect of Family Background on Student Effort," *B.E. Journal of Economic Analysis & Policy,* 14(4), 1337?1403.

Mather Nancy and Lynne E. Jaffe (2002): *Woodcock-Johnson III: Reports, Recommendations, and Strategies,* John Wiley & Sons, New York.

Melitz, Jacques and Farid Toubal (2014): "Native language, spoken language, translation and trade," *Journal of International Economics,* 93(2), 351-363.

Mouw, Ted and Yu Xie (1999): "Bilingualism and the academic achievement of first- and second-generation Asian Americans: accommodation with or without assimilation?," *American Sociological Review,* 64(2), 232-252.

OECD (2019): "PISA 2018 Results (Volume I): What Students Know and Can Do, PISA," OECD Publishing, Paris.

Oster, Emily (2019): "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 37(2), 187-204.

Paap, Kenneth R. and Zachary I. Greenberg (2013): "There is no coherent evidence for a bilingual advantage in executive processing," *Cognitive Psychology,* 66, 232-258.

Peal, Elizabeth and Wallace E. Lambert (1962): "The relation of bilingualism to intelligence," *Psychological Monographs: General and Applied*, 76(27), 1-23.

PEW Hispanic Center (2007): "English Usage among Hispanics in the United States," PEW Research Center, Washington D.C.

Rosselli, Mónica; Ardila, Alfredo; Matute, Esmeralda and Olga Inozemtseva (2009): "Gender differences and cognitive correlates of mathematical skills in school-aged children," *Child Neuropsychology,* 15(3), 216-231.

Rodríguez-Pujadas, Aina; Sanjuán, Ana; Ventura-Campos, Noelia; Román, Patricia; Martin, Clara; Barceló, Francisco; Costa, Albert and César Ávila (2013): "Bilinguals use language-control brain areas more than monolinguals to perform non-linguistic switching tasks," *PLoS One,* 8(9), e73028.

Rosenzweig, Mark R. and Kenneth I. Wolpin (1994): "Are There Increasing Returns to the Intergenerational Production of Human Capital? Maternal Schooling and Child Intellectual Achievement," *Journal of Human Resources,* 29(2), 670-693.

Ruggles, Steven; Flood, Sarah; Foster, Sophia; Goeken, Ronald; Pacas, Jose; Schouweiler, Megan and Matthew Sobek. (2021): IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021.

Rumberger, Russell W. and Katherine A. Larson (1998): "Toward explaining differences in educational achievement among Mexican American language-minority students," *Sociology of Education,* 71(1), 68-92.

Rutherford-Quach, Sara; Torre Gibney, Daniela; Kelly, Hannah;, Ballen Riccards, Jennifer; Garcia, Elisa; Hsiao, Mindy; Pellerin, Emma. and Carrie Parker (2021): "Bilingual education across the United States.," CCNetwork.

Surrain, Sarah and Gigi Luk (2019): "Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005?2015," *Bilingualism: Language and Cognition,* 22(2), pp. 401-415.

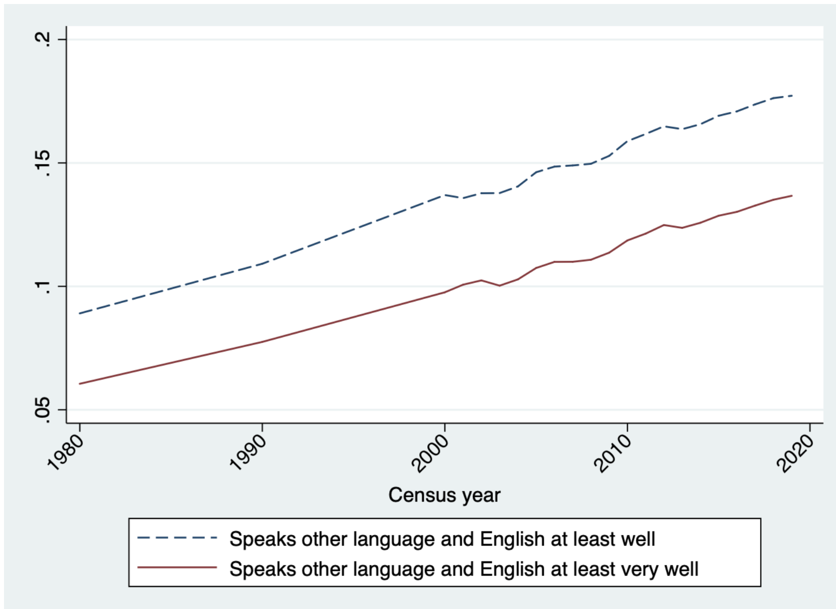
Tseng, Vivian and Andrew J. Fuligni (2000): "Parent-Adolescent language use and relationships among immigrant families with East Asian, Filipino, and Latin American backgrounds," *Journal of Marriage and Family,* 62(2), 465-476.

US Census Bureau (2013): "Language Use in the United States: 2011," American Community Survey Reports, US Census Bureau.

Woessmann, Ludger (2016): "The Importance of School Systems: Evidence from International Differences in Student Achievement," *Journal of Economic Perspectives,* 30(3), 3-32.

A Appendix

Figure A-1: Share of bilingual individuals in the Unites States



Source: US Census, Ruggles et al. [2021], weighted data

Table A1: Differences between Bilingual and Monolingual Individuals (writing)

Variable	Bilingual	Monolingual	Difference
Standardized calculation score	0.225	0.055	0.169**
Calculation score	17.363	14.673	2.690***
Standardized passage comprehension score	0.2875	0.091	0.276***
Passage comprehension score	22.316	18.293	4.023***
Standardized applied problems	0.212	0.121	0.091
Applied Problems score	28.958	26.190	2.768***
Standardized Letter Word identification score	0.391	0.037	0.355***
Letter Word identification score	50.561	40.785	9.776***
WJ tests in Spanish	0.486	0.520	-0.034
Standardized correct digits	-0.259	-0.196	-0.063
Correct digits	9.566	9.355	0.211
Boy	0.448	0.559	-0.111**
Age	9.505	8.874	0.630***
Number of siblings	1.708	1.818	-0.111
Years in US	7.712	6.927	0.785**
Mother's age	36.174	35.342	0.832*
Father's age	39.083	38.064	1.019**
Mother's years of schooling	9.733	9.332	0.401
Father's years of schooling	10.127	9.880	0.247
Mother's years of schooling in US	0.473	0.413	0.061
Father's years of schooling in US	0.545	0.530	0.014
Married Parents	0.830	0.799	0.031
Cohabiting Parents	0.075	0.059	0.017
Separated Parents	0.038	0.034	0.004
Divorced Parents	0.028	0.028	0.0003
Single Parents	0.024	0.081	-0.057***
Widowed Parents	0.005	0	0.005
Mother at home	0.939	0.9543	-0.014
Father at home	0.854	0.810	0.044
Mother is Mexican	0.443	0.458	-0.015
Father is Mexican	0.354	0.349	0.005
Median Households income by state/region	46412.75	46182.8	229.95
Per pupil expenditure on public education	9131.62	9278.66	- 147.04

Notes: This table presents the average observable characteristics of monolingual and bilingual writers who are part of the sample used in the estimation, separately. Here we have excluded non-performers from the sample (12.6%). The last column includes the differences between the two values. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A2: Differences between Bilingual and Monolingual Individuals (understanding)

Variable	Bilingual	Monolingual	Difference
Standardized calculation score	0.087	0.064	0.024
Calculation score	15.442	13.913	1.528**
Standardized passage comprehension score	0.119	-0.024	0.144 **
Passage comprehension score	19.843	17.303	2.540***
Standardized applied problems	0.194	0.025	0.168**
Applied Problems score	27.690	24.226	3.464***
Standardized Letter Word identification score	0.107	0.103	0.005
Letter Word identification score	43.039	40.625	2.414
WJ tests in Spanish	0.474	0.558	-0.084**
Standardized correct digits	-0.236	-0.372	0.136*
Correct digits	9.339	8.442	0.897***
Boy	0.503	0.534	-0.030
Age	9.080	8.567	0.512***
Number of siblings	1.793	1.755	0.038
Years in US	8.057	5.240	2.817***
Mother's age	35.514	35.412	0.102
Father's age	38.349	37.887	0.462
Mother's years of schooling	9.383	9.474	-0.090
Father's years of schooling	9.857	9.989	-0.132
Mother's years of schooling in US	0.451	0.448	0.004
Father's years of schooling in US	0.523	0.498	0.025
Married Parents	0.820	0.774	0.046
Cohabiting Parents	0.071	0.053	0.018
Separated Parents	0.030	0.053	-0.023
Divorced Parents	0.027	0.019	0.008
Single Parents	0.050	0.101	-0.051**
Widowed Parents	0.002	0	0.002
Mother at home	0.948	0.952	0.004
Father at home	0.841	0.782	0.058*
Mother is Mexican	0.467	0.438	0.029
Father is Mexican	0.355	0.361	-0.005
Median Households income by state/region	46529.15	45554.04	975.1117
Per pupil expenditure on public education	9210.079	9207.448	2.630984

Notes: This table presents the average observable characteristics of monolingual and bilingual individuals (in terms of understanding) who are part of the sample used in the estimation, separately. Here we have excluded non-performers from the sample (0.77%). The last column includes the differences between the two values. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A3: Differences between Bilingual and Monolingual Individuals (speaking)

Variable	Bilingual	Monolingual	Difference
Standardized calculation score	0.107	0.027	0.080
Calculation score	15.639	13.700	1.939***
Standardized passage comprehension score	0.165	-0.106	0.271***
Passage comprehension score	20.337	16.523	3.815***
Standardized applied problems	0.198	0.025	0.173**
Applied Problems score	27.817	24.214	3.604***
Standardized Letter Word identification score	0.171	-0.019	0.191**
Letter Word identification score	44.417	38.264	6.153***
WJ tests in Spanish	0.494	0.509	-0.015
Standardized correct digits	-0.229	-0.375	0.146**
Correct digits	9.382	8.441	0.941***
Boy	0.485	0.573	-0.088**
Age	9.103	8.577	0.526***
Number of siblings	1.782	1.773	0.009
Years in US	8.091	5.300	2.791***
Mother's age	35.510	35.511	0.001
Father's age	38.309	38.057	0.253
Mother's years of schooling	9.508	9.241	0.267
Father's years of schooling	9.937	9.804	0.133
Mother's years of schooling in US	0.473	0.409	0.063
Father's years of schooling in US	0.525	0.494	0.031
Married Parents	0.827	0.764	0.063*
Cohabiting Parents	0.070	0.055	0.016
Separated Parents	0.026	0.059	-0.0337**
Divorced Parents	0.030	0.014	0.017
Single Parents	0.044	0.109	-0.065***
Widowed Parents	0.002	0	0.002
Mother at home	0.944	0.955	-0.011
Father at home	0.855	0.764	0.091***
Mother is Mexican	0.471	0.423	0.048
Father is Mexican	0.372	0.318	0.054
Median Households income by state/region	46475.72	45751.59	724.1357*
Per pupil expenditure on public education	9150.276	9349.15	- 198.8746

Notes: This table presents the average observable characteristics of monolingual and bilingual speakers who are part of the sample used in the estimation, separately. Here we have excluded non-performers from the sample (0.77%). The last column includes the differences between the two values. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Full set of results: Bilingual readers

	(1)	(2)	(3)	(4)
Bilingual	0.243 (0.103)**	0.301 (0.11)***	0.29 (0.116)**	0.291 (0.114)**
Monolingual Spanish	-0.101 (0.144)	-0.037 (0.167)	-0.040 (0.178)	0.036 (0.176)
Neither language	-0.295 (0.146)**	-0.209 (0.155)	-0.237 (0.161)	-0.173 (0.159)
WJ tests in Spanish	-0.176 (0.103)*	-0.149 (0.107)	-0.156 (0.112)	-0.156 (0.11)
Test Spanish × Bilingual	-0.057 (0.147)	-0.076 (0.154)	-0.063 (0.163)	-0.090 (0.16)
Test Spanish × Monolingual Spanish	0.128 (0.19)	0.006 (0.199)	-0.013 (0.21)	-0.036 (0.207)
Test Spanish × neither language	0.15 (0.213)	0.033 (0.219)	0.079 (0.231)	0.065 (0.227)
Boy		0.035 (0.066)	0.043 (0.07)	0.029 (0.069)
Number of siblings		-0.020 (0.032)	-0.0004 (0.034)	-0.0004 (0.033)
Mother's years of education		0.027 (0.012)**	0.031 (0.012)**	0.029 (0.012)**
Father's years of education		0.01 (0.014)	0.014 (0.016)	0.009 (0.015)
Mother's years of US education		0.008 (0.032)	0.009 (0.033)	0.014 (0.032)
Father's years of US education		-0.046 (0.033)	-0.054 (0.035)	-0.043 (0.035)
Parents cohabiting		0.402 (0.235)*	0.366 (0.248)	0.337 (0.244)
Parents separated		0.076 (0.268)	0.065 (0.281)	0.067 (0.276)
Parents divorced		0.538 (0.292)*	0.618 (0.306)**	0.597 (0.301)**
Single parent		0.126 (0.249)	0.138 (0.267)	0.103 (0.262)
Widowed parent		-0.688 (0.85)	-1.845 (1.177)	-1.944 (1.158)*
Father at home		0.185 (0.498)	0.331 (0.53)	0.291 (0.522)
Mother's age		0.004 (0.008)	0.004 (0.009)	0.004 (0.009)
Father's age		-0.009 (0.008)	-0.012 (0.008)	-0.013 (0.008)
Age mother missing		0.233 (0.301)	0.231 (0.306)	0.141 (0.302)
Age father missing		0.025 (0.284)	0.047 (0.291)	0.021 (0.286)
Std. sum correctly repeated digits				0.175 (0.041)***
US state of residence		x	x	x
Age		x	x	x
Years in US		x	x	x
Age × years in US			x	x
Number of observations	652	652	652	652
R-squared	0.044	0.173	0.225	0.251
F statistic	4.232	1.634	1.255	1.436

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. The dependent variable is the standardized calculation test score. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A5: Full set of results: Bilingual writers

	(1)	(2)	(3)	(4)
Bilingual	0.235 (0.105)**	0.292 (0.113)***	0.281 (0.118)**	0.298 (0.116)**
Monolingual Spanish	0.023 (0.137)	0.199 (0.159)	0.186 (0.17)	0.226 (0.167)
Neither language	-0.258 (0.141)*	-0.270 (0.148)*	-0.277 (0.154)*	-0.175 (0.153)
WJ tests in Spanish	-0.145 (0.1)	-0.124 (0.103)	-0.151 (0.108)	-0.165 (0.106)
Test Spanish × Bilingual	-0.116 (0.149)	-0.129 (0.154)	-0.080 (0.162)	-0.075 (0.159)
Test Spanish × Monolingual Spanish	0.058 (0.19)	-0.155 (0.197)	-0.133 (0.208)	-0.103 (0.204)
Test Spanish × neither language	0.027 (0.204)	0.006 (0.208)	0.049 (0.22)	0.037 (0.216)
Boy		0.04 (0.066)	0.052 (0.07)	0.039 (0.069)
Number of siblings		-0.021 (0.032)	-0.002 (0.034)	-0.003 (0.033)
Mother's years of education		0.029 (0.012)**	0.032 (0.012)***	0.029 (0.012)**
Father's years of education		0.008 (0.015)	0.012 (0.016)	0.007 (0.015)
Mother's years of US education		0.008 (0.032)	0.011 (0.033)	0.015 (0.032)
Father's years of US education		-0.055 (0.033)*	-0.065 (0.035)*	-0.053 (0.034)
Parents cohabiting		0.414 (0.235)*	0.371 (0.247)	0.339 (0.243)
Parents separated		0.042 (0.267)	0.015 (0.279)	0.02 (0.274)
Parents divorced		0.509 (0.292)*	0.58 (0.305)*	0.566 (0.3)*
Single parent		0.132 (0.248)	0.121 (0.266)	0.081 (0.261)
Widowed parent		-0.681 (0.849)	-1.867 (1.176)	-1.964 (1.156)*
Father at home		0.133 (0.497)	0.256 (0.53)	0.22 (0.521)
Mother's age		0.005 (0.008)	0.005 (0.009)	0.004 (0.009)
Father's age		-0.010 (0.008)	-0.013 (0.008)	-0.013 (0.008)*
Age mother missing		0.179 (0.301)	0.172 (0.306)	0.091 (0.301)
Age father missing		0.024 (0.283)	0.035 (0.291)	-0.0004 (0.286)
Std. sum correctly repeated digits				0.181 (0.041)***
US state of residence		x	x	x
Age		x	x	x
Years in US		x	x	x
Age × years in US			x	x
Number of observations	652	652	652	652
R-squared	0.038	0.174	0.226	0.253
F statistic	3.677	1.647	1.263	1.457

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. The dependent variable is the standardized calculation test score. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A6: Full set of results: Bilingual speakers

	(1)	(2)	(3)	(4)
Bilingual	-0.045 (0.135)	-0.048 (0.141)	-0.065 (0.145)	-0.063 (0.142)
Monolingual Spanish	-0.111 (0.159)	0.001 (0.182)	0.013 (0.191)	0.059 (0.187)
Neither language	0.595 (0.821)	0.706 (0.819)	0.704 (0.829)	0.507 (0.814)
WJ tests in Spanish	-0.262 (0.17)	-0.202 (0.175)	-0.203 (0.182)	-0.265 (0.179)
Test Spanish × Bilingual	0.138 (0.187)	0.07 (0.192)	0.078 (0.201)	0.122 (0.197)
Test Spanish × Monolingual Spanish	0.039 (0.222)	-0.130 (0.228)	-0.120 (0.24)	-0.022 (0.236)
Test Spanish × neither language	-0.903 (0.923)	-1.020 (0.92)	-1.019 (0.932)	-0.698 (0.917)
Boy		0.026 (0.067)	0.033 (0.071)	0.02 (0.069)
Number of siblings		-0.034 (0.033)	-0.016 (0.035)	-0.015 (0.034)
Mother's years of education		0.031 (0.012)***	0.035 (0.013)***	0.032 (0.012)***
Father's years of education		0.015 (0.015)	0.018 (0.016)	0.011 (0.016)
Mother's years of US education		0.014 (0.032)	0.016 (0.033)	0.02 (0.033)
Father's year of US education		-0.057 (0.034)*	-0.066 (0.035)*	-0.053 (0.035)
Parents cohabiting		0.437 (0.239)*	0.385 (0.252)	0.341 (0.247)
Parents separated		0.048 (0.272)	0.014 (0.285)	0.018 (0.28)
Parents divorced		0.561 (0.297)*	0.63 (0.311)**	0.605 (0.305)**
Single parent		0.08 (0.251)	0.07 (0.27)	0.031 (0.265)
Widowed Parent		-0.562 (0.861)	-1.882 (1.195)	-1.985 (1.172)*
Father at home		0.151 (0.504)	0.276 (0.538)	0.215 (0.528)
Mother's age		0.005 (0.009)	0.004 (0.009)	0.004 (0.009)
Father's age		-0.008 (0.008)	-0.012 (0.008)	-0.011 (0.008)
Age mother missing		0.263 (0.306)	0.244 (0.31)	0.16 (0.305)
Age father missing		0.062 (0.288)	0.081 (0.296)	0.039 (0.29)
Std. sum correctly repeated digits				0.192 (0.041)***
US state of residence		x	x	x
Age		x	x	x
Years in US		x	x	x
Age × years in US			x	x
Number of observations	652	652	652	652
R-squared	0.017	0.148	0.2	0.232
F statistic	1.597	1.35	1.087	1.295

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. The dependent variable is the standardized calculation test score. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Full set of results: Bilingual individuals who understand both languages

	(1)	(2)	(3)	(4)
Bilingual	-0.151	-0.148	-0.158	-0.154
	(0.144)	(0.148)	(0.152)	(0.149)
Monolingual Spanish	-0.219	-0.137	-0.143	-0.096
	(0.172)	(0.194)	(0.202)	(0.199)
Neither language	0.337	0.461	0.678	0.501
	(0.588)	(0.612)	(0.628)	(0.618)
WJ tests in Spanish	-0.215	-0.202	-0.215	-0.242
	(0.177)	(0.18)	(0.187)	(0.184)
Test Spanish × Bilingual	0.073	0.061	0.081	0.088
	(0.193)	(0.198)	(0.206)	(0.202)
Test Spanish × Monolingual Spanish	0.02	-0.104	-0.052	-0.024
	(0.23)	(0.235)	(0.246)	(0.242)
Test Spanish × neither language	-0.955	-1.011	-1.257	-0.970
	(0.761)	(0.781)	(0.798)	(0.786)
Boy		0.019	0.026	0.013
		(0.067)	(0.071)	(0.069)
Number of siblings		-0.036	-0.017	-0.017
		(0.033)	(0.034)	(0.034)
Mother's years of education		0.033	0.036	0.032
		(0.012)***	(0.012)***	(0.012)***
Father's years of education		0.016	0.018	0.012
		(0.015)	(0.016)	(0.016)
Mother's years of US education		0.01	0.012	0.016
		(0.032)	(0.033)	(0.033)
Father's years of US education		-0.054	-0.063	-0.051
		(0.033)	(0.035)*	(0.035)
Parents cohabiting		0.434	0.391	0.35
		(0.239)*	(0.252)	(0.247)
Parents separated		0.041	0.01	0.015
		(0.271)	(0.283)	(0.278)
Parents divorced		0.563	0.638	0.613
		(0.297)*	(0.311)**	(0.305)**
Single parent		0.092	0.086	0.05
		(0.25)	(0.269)	(0.264)
Widowed parent		-0.519	-1.846	-1.944
		(0.861)	(1.192)	(1.170)*
Father at home		0.168	0.302	0.241
		(0.503)	(0.536)	(0.527)
Mother's age		0.003	0.002	0.002
		(0.009)	(0.009)	(0.009)
Father's age		-0.007	-0.010	-0.011
		(0.008)	(0.008)	(0.008)
Age mother missing		0.3	0.274	0.173
		(0.305)	(0.31)	(0.305)
Age father missing		0.068	0.084	0.045
		(0.287)	(0.295)	(0.29)
Std. sum correctly repeated digits				0.186
				(0.041)***
US state of residence		x	x	x
Age		x	x	x
Years in US		x	x	x
Age × years in US			x	x
Number of observations	652	652	652	652
R-squared	0.019	0.15	0.204	0.234
F statistic	1.796	1.381	1.112	1.312

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. The dependent variable is the standardized calculation test score. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A8: Robustness: Bilingual advantage; controlling for geographical variation in median household income and per pupil spending on public education

	(1)	(2)	(3)	(4)
Bilingual reader	0.243	0.269	0.26	0.263
	(0.103)**	(0.11)**	(0.115)**	(0.113)**
Individual characteristics	No	Yes	Yes	Yes
State/region variables	No	Yes	Yes	Yes
Age by Years in US	No	No	Yes	Yes
Innate ability	No	No	No	Yes
Number of observations	652	652	652	652
R-squared	0.044	0.135	0.2	0.224
F statistic	4.232	1.481	1.229	1.402

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. All columns control for language proficiency, test language, and the last two variables interacted. The dependent variable is the standardized calculation test score. Individual characteristics refer to age, years in US, gender, number of siblings, parental age, education, marital status, and country of origin. State/region variables refer to median household income and per pupil expenditure on public education by state/region of parental residence in 2003. Innate ability refers to the age-standardized sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A9: Robustness: Using raw test scores instead of standardized test scores – Different measures of bilingualism and cognitive advantage

	Reads	Understands	Speaks	Writes
Bilingual	2.131	-0.820	-0.448	2.179
	(0.804)***	(1.049)	(1.002)	(0.821)***
Individual characteristics	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes
Number of observations	652	652	652	652
R-squared	0.542	0.533	0.531	0.544
F statistic	5.078	4.906	4.858	5.119

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. All columns control for years in US, age of the child, language proficiency, test language, and the last two interacted. The dependent variable is the raw calculation test score. Innate ability refers to the raw correctly repeated sum of digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A10: Robustness: Logit model for scoring above mean on calculation test

	Reads	Understands	Speaks	Writes
Bilingual	0.357	-0.207	-0.307	0.818
	(0.369)	(0.448)	(0.422)	(0.383)**
Individual characteristics	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes
Number of observations	597	597	597	597

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by a logistic regression with the dependent variable redefined as a dummy variable for having ir not scored above the age-dependent mean on the calculation test score. All columns control for years in US, age of the child, language proficiency, test language, and the last two interacted. Innate ability refers to the age-standardized correctly repeated sum of digits. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Falsification: Difference in scores on language-dependent tests between bilinguals and monolinguals

	Reads	Understands	Speaks	Writes
A: Passage Comprehension				
Bilingual	0.107	-.028	-.009	0.091
	(0.11)	(0.146)	(0.14)	(0.112)
R-squared	0.325	0.296	0.286	0.338
F statistic	2.07	1.803	1.717	2.191
B: Letter-Word identification				
Bilingual	0.139	-.199	-.066	0.13
	(0.112)	(0.148)	(0.143)	(0.112)
R-squared	0.403	0.371	0.358	0.425
F statistic	2.9	2.531	2.397	3.178
C: Applied Problems				
Bilingual	0.097	-.074	-.049	0.171
	(0.113)	(0.147)	(0.14)	(0.116)
R-squared	0.28	0.273	0.267	0.279
F statistic	1.672	1.614	1.565	1.661
Individual characteristics	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes
Number of observations	652	652	652	652

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS. All columns control for years in US, age of the child, language proficiency, test language, and the last two interacted. The dependent variable in Panel A is the standardized passage comprehension test score, in Panel B the standardized score on the Letter-Word identification test, and in Panel C the standardized score on Applied Problems. Innate ability refers to the standardized sum of correctly repeated digits. *** p<0.01, ** p<0.05, * p<0.1.

Table A12: Differences between Children of Mexican and Non-Mexican parents

Variable	Mexican Parents	Non-Mexican parents	Difference
Standardized calculation score	0.083	0.074	0.009
Calculation score	14.395	15.536	-1.141**
Standardized passage comprehension score	0.074	0.055	0.019
Passage comprehension score	18.505	19.403	-0.898
Standardized Applied Problems score	0.123	0.145	-0.022
Applied Problems score	25.714	27.322	-1.608*
Standardized Letter Word identification score	0.060	0.141	- 0.081
Letter Word identification score	40.292	44.105	- 3.813**
WJ tests in Spanish	0.480	0.523	-0.043
Standardized correct digits	-0.312	-0.247	-0.065
Correct digits	8.775	9.341	-0.566**
Bilingual Speaker	0.687	0.622	0.065*
Spanish Monolingual Speaker	0.158	0.235	-0.077**
English Monolingual Speaker	0.143	0.139	0.004
Bilingual Understander	0.693	0.653	0.040
Spanish Monolingual Understander	0.155	0.220	-0.065**
English Monolingual Understander	0.146	0.118	0.028
Bilingual Reader	0.353	0.359	0.007
Spanish Monolingual Reader	0.143	0.173	0.031
English Monolingual Reader	0.377	0.368	0.009
Bilingual Writer	0.322	0.328	-0.006
Spanish Monolingual Writer	0.128	0.177	-0.049*
English Monolingual Writer	0.416	0.378	0.039
Boy	0.483	0.548	-0.065*
Age	8.699	9.155	-0.456**
Number of siblings	2.0729	1.4923	0.5807***
Years in US	7.690	6.613	1.077**
Mother's age	35.035	36.014	-0.980**
Father's age	37.537	38.882	-1.346***
Mother's years of schooling	8.795	10.024	-1.229***
Father's years of schooling	9.668	10.091	-0.424**
Mother's years of schooling in US	0.514	0.383	0.131
Father's years of schooling in US	0.448	0.579	-0.131
Married Parents	0.854	0.759	0.096***
Cohabiting Parents	0.040	0.090	-0.050***
Separated Parents	0.052	0.022	0.030 **
Divorced Parents	0.015	0.034	-0.019*
Single Parents	0.037	0.096	-0.060***
Widowed Parents	0.003	0	0.003
Mother at home	0.979	0.916	0.062***
Father at home	0.866	0.783	0.083***
Mother is Mexican	0.906	0	0.906***
Father is Mexican	0.708	0	0.708***
Median Households income by state/region	46142.15	46287.09	-144.941
Per pupil expenditure on public education	8748.222	9685.274	- 937.053***

Notes: This table presents the average observable characteristics of children with and without Mexican parents who are part of the sample used in the estimation. The last column includes the differences between the two values. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A13: Heterogeneity: Bilingualism and cognitive advantage by parental education

	Bilingual writer		Bilingual reader	
	$\geq$ 10 years	$<$ 10 years	$\geq$ 10 years	$<$ 10 years
Bilingual	0.302	0.391	0.399	0.355
	(0.186)	(0.164)**	(0.177)**	(0.163)**
Individual characteristics	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes
Number of observations	307	345	307	345
R-squared	0.41	0.342	0.421	0.333
F statistic	1.256	1.563	1.314	1.497

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS separately for children of parents with less (or more than) 10 years of education. All columns control for years in US, age of the child, language proficiency, test language, and the last two interacted. The dependent variable is the standardized calculation test score. Innate ability refers to the age-standardized sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A14: Bilingualism and cognitive advantage by family type, alternative definition

	Bilingual writer		Bilingual reader	
	Traditional family	Non-traditional	Traditional	Non-traditional
Bilingual	0.486	0.245	0.214	0.331
	(0.172)***	(0.167)	(0.163)	(0.166)**
Individual characteristics	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes
Number of observations	288	364	288	364
R-squared	0.451	0.371	0.437	0.357
F statistic	1.772	1.55	1.67	1.462

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by OLS separately for children in traditional and non-traditional families, where traditional families are defined as households with married parents, living together and at least three children. All columns control for years in US, age of the child, language proficiency, test language, and the last two interacted. The dependent variable is the standardized calculation test score. Innate ability refers to the age-standardized sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A15: The bilingual advantage: Heterogeneity by ability

	15%	25%	50%	75%	85%
Bilingual writer	0.218	0.263	0.26	0.206	0.2
	(0.246)	(0.122)**	(0.111)**	(0.085)**	(0.094)**
Individual characteristics	Yes	Yes	Yes	Yes	Yes
Age by Years in US	Yes	Yes	Yes	Yes	Yes
Innate ability	Yes	Yes	Yes	Yes	Yes
Number of observations	652	652	652	652	652
Pseudo r-squared	0.2921	0.2519	0.1924	0.2200	0.2547

Notes: Data is from the NIS 2003 and refers to children of Spanish-speaking immigrants ages 6 to 12. Coefficients are the result of estimating Equation 1 by quantile regressions. The dependent variable is the standardized reading test score. All columns control for years in US, language proficiency, test language, and age of the child. Innate ability refers to the age-standardized sum of correctly repeated digits. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.