I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11524

Language Assimilation and Performance in Achievement Tests among Immigrant Children: Evidence from a Field Experiment

Sankar Mukhopadhyay

MAY 2018

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11524

Language Assimilation and Performance in Achievement Tests among Immigrant Children: Evidence from a Field Experiment

Sankar Mukhopadhyay
University of Nevada Reno and IZA

MAY 2018

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

Language Assimilation and Performance in Achievement Tests among Immigrant Children: Evidence from a Field Experiment

We provide new evidence about language assimilation and its effect on test scores using data from two rounds (conducted approximately six years apart) of the New Immigrants Survey (NIS). As part of the NIS interviews, U.S. born and foreign-born children of immigrants were asked to take Woodcock-Johnson achievement test. In both rounds, prior to the administration of tests, children of Hispanic origin were randomly assigned to take the tests either in Spanish or in English. Therefore, we can attribute the difference in scores to language proficiency and directly estimate the rate of assimilation. Our results suggest that in reading tests, U.S. born children of Hispanic immigrants perform better, when they are assigned to take the tests in English, and the advantage remains stable across two rounds of interviews. However, there is substantial heterogeneity. For example, U.S. born children at the top of score distribution perform better when they take tests in Spanish. Foreign-born children of Hispanic immigrants exhibit Spanish dominance during the first round, but it declines and in some cases completely disappears by the second round. We find that foreign-born children who immigrated to the U.S. after age six, exhibit Spanish dominance in reading tests in the first round. However, during the six years between interviews, Spanish dominance disappears among foreign-born children who immigrated between the ages of six and eight (in reading) and in all children (in math). Moreover, for children who still have Spanish dominance in reading, the score differences have narrowed.

JEL Classification: J15, I20, Z13

Keywords: immigrants, language assimilation, test score, education, children

Corresponding author:
Sankar Mukhopadhyay
University of Nevada
Reno, NV 89557
USA

E-mail: sankarm@unr.edu

1. Introduction

As of 2015, about 26% of all children in the U.S. are immigrant children[1] (2.8 million first generation and 15.9 million second generation)[2]. About 58% of all immigrant children in the U.S. are of Hispanic origin[3]. In other words, approximately one out of seven children living in the U.S. today is of Hispanic origin. Kindler (2002) reported that about 11% of public school students are English Language Learners (ELL)[4] and about 79% of all ELL students are of Hispanic origin (Calderón, Slavin, & Sánchez 2011). Furthermore, about 50% of all ELL students are foreign-born, even though they are only about 4% of all students (García, Kleifgen, and Falchi 2008). The situation in the U.S. is not unique. Numbers in many European countries are comparable. Harte, Herrera, and Stepanek (2016) reported that that in 2012, about 46% of all children in Luxembourg are immigrants (first or second generation). In addition, 10% or more of all children in Austria, Belgium, Croatia, France, Germany, Greece, Sweden, Switzerland, and the UK are immigrants (first or second generation). Among them, about 4% of all children in the EU-15 countries are first generation (Janta and Harte, 2016).

Immigrant children tend to perform worse in tests than natives do (OECD 2010, Brind et al. 2008, Dustmann et al. 2012, De Paola and Brunello 2016 among others). While all of these studies mention deficiency in host country language as a factor, it is not clear how much of this performance gap can be attributed to language ability. For example, Jakubowski (2011) report that socio-economic variables can explain only a part of this gap, Dustmann et al. (2012) report that, in some countries (though not in all countries), the performance gap disappears once we control for socio-economic variables. Heubert and Hauser (1999) observed "*If a student is not proficient in the language of the test, her performance is likely to be affected by construct-irrelevant variance—that is, her test score is likely to underestimate her knowledge of the subject being tested*" (pp. 225). Despite these concerns, these test scores are widely used in the U.S. to evaluate students, teachers, and schools. Students, who do not perform adequately in tests, may be placed on special

[1] Sometimes also referred to as children with migrant background.
[2] First generation refers to children who are foreign-born and second generation refers to children who have at least one parent who is foreign born. Children who immigrated before the age of 12 are sometimes referred to as 1.5 generation.
[3] available at https://www.childtrends.org/wpcontent/uploads/2013/07/110_Immigrant_Children.pdf.

tracks, or they may have to repeat grades, and in some cases, they may not be able to graduate (Heubert and Hauser 1999; Shepard & Smith, 1989; Sullivan et. al. 2005, Nouwen et al. 2015). These may have a long-term effect on their educational outcomes and therefore on their labor market outcomes (Kossoudji 1988; Tainer 1988; Bleakley and Chin 2004; Berman, Lang, and Siniver 2003; Schnepf, 2007; Dustmann and van Soest 2002). Janta and Harte (2016) report that in all EU countries young immigrants (15-24 years old) are more likely to be in the NEET (not in education, employment, or training) category compared to natives. Furthermore, Language assimilation is important not only for labor market success, but also for other socio-demographic outcomes such as marriage, divorce, and fertility (Bleakley and Chin 2010).

With the passage of the *English Language Acquisition, Language Enhancement, and Academic Achievement Act of 2001*, the focus in the U.S. has shifted to English only (as opposed to bilingual) instruction. In addition to the changes at the Federal level, several states (Arizona, California, and Massachusetts) have also imposed independent restrictions on bilingual educational programs (Rolstad, Mahoney, & Glass 2005). Therefore, academic success of millions of immigrant children may critically depend on their English ability. Given the size of the affected population and potentially serious consequences of low test-scores in achievement tests, it is important for us to understand how English skills affect test scores of immigrant children in the U.S. In this paper, we provide new evidence about language assimilation using data from two rounds of the New Immigrants Survey (NIS). The NIS interviewed a nationally representative sample of newly admitted legal permanent residents (LPRs) to the U.S. and their family members. In both rounds, eligible children (both first and second-generation) of principal immigrants were asked to take Woodcock-Johnson achievement test – version III (WJ-ACH III). As part of the NIS interviews, a *language experiment* was conducted to assess the effect of testing language on test scores. For this purpose, children of Hispanic origin[5] were randomly assigned to take the tests either in Spanish or in English. Randomization was implemented,

[5] Children who had a parent born in a Spanish speaking country and parent's first language was Spanish then were eligible for the language experiment (Akresh and Akresh 2011).

independently, prior to each round. We show that the randomization was effective in both rounds[6].

In this paper, we address two questions. *First*, with two rounds of survey we observe same children about six years apart. Therefore, we directly estimate the rate of language assimilation for both U.S. born and foreign-born children. We also explore heterogeneity in language assimilation. *Second*, we explore how age at the time of immigration and time spent in the U.S. affects language assimilation among foreign-born children. We estimate how the score differentials changed between two rounds of interviews for a given age at immigration. For foreign-born children, age at immigration and time spent in the U.S. may have independent effects on English acquisition. There may an interaction effect as well. Thus, we disentangle "age-at-immigration" effect from "time-spent-in-the-U.S." effect.

This paper is most closely related to Akresh and Akresh (2011), who used data from the first round of the NIS interviews to show that U.S. born children of Hispanic immigrants perform better, when they take tests in English (English dominance) and foreign-born children of Hispanic immigrants perform better, when they take tests in Spanish (Spanish dominance). Despite of the similarities, there are important differences between this paper and Akresh and Akresh (2011). In Akresh and Akresh (2011), about 50% of foreign-born children in their sample spent less than two years in the U.S. at the time of their study. In other words, the outcomes analyzed in Akresh and Akresh (2011) are intermediate outcomes and not final language learning outcomes. We show that this has important implications for overall conclusions. . Moreover, we explore and report heterogeneity in the language effect not addressed in Akresh and Akresh (2011). In addition, Akresh and Akresh (2011) use cross-sectional data (only the first round of the NIS) and therefore cannot disentangle "age-at-immigration" effect from "time-spent-in-the-U.S." effect. However, a small number of papers have previously used longitudinal data to disentangle these two effects (Conger 2008, Slama 2014). However, these papers study amount to time needed to achieve "minimal English proficiency". Minimal

[6] For comparability between WJ ACH III tests conducted in English and Spanish please see Schrank, McGrew, and Woodcock (2001) and Schrank et al. (2005).

efficiency is not equivalent to native-like efficiency. Test score of a student who has minimal English proficiency may still not reflect their true ability. Furthermore, these studies are based on observational data and the identification exclusively relies on having large number of controls and fixed effects.

Age at immigration is one of the most important determinants of language acquisition process. Lenneberg (1967) found that children who start to learn a second language before age nine, are substantially more likely to master that language compared to children who are first exposed to a new language after age nine (also known as Critical Period Hypothesis or CPH). However, two influential studies in this literature arrive at contradictory conclusions regarding when that critical period begins to end. Johnson and Newport (1989) found that language acquisition window may start to close as soon as seven, on the other hand Birdsong and Molis (2001) found that language acquisition window might extend to as late as 15. For a comprehensive review of this literature, please see Newport (2002) and Birdsong (2006). Akresh and Akresh (2011) found that foreign-born children of Hispanic descent who immigrated to the U.S. after age six or so have lower achievement scores when tests are administered in English compared to Spanish. This result is inconsistent with Birdsong and Molis (2001) who found that language acquisition window might extend to as late as 15 for children of Hispanic descent[7]. The differences between these two studies may stem from the fact that Birdsong and Molis (2001) study adults who have resided in the U.S. for about 10 years on an average. On the other hand, Akresh and Akresh (2011) subjects are children (between the ages of 3 and 12) who have been in the U.S. for only about three years on an average, and about 50% of children spent less than two years in the U.S. at the time of their interview. Therefore, they have not reached their final (steady) state in language acquisition process. In this paper, we attempt to reconcile these inconsistencies.

Our results suggest that the U.S. born children of Hispanic immigrants exhibit English dominance in Reading tests in both rounds. The score differentials[8] in reading tests remain stable across two rounds of interviews. However, our results also suggest that

[7] Subjects in Johnson and Newport (1989) study were of Chinese and Korean descent.
[8] The difference in scores between children who were assigned to take the tests in English, and those who were assigned to take the tests in Spanish

average score differentials mask substantial heterogeneity. While U.S. born children of Hispanic immigrants at the bottom end (below 50th percentile) of the reading score distribution show English dominance, those in the middle (between 50th and 80th percentile) of reading score distribution are equally proficient in English and Spanish, and those at the top (90th percentile or above) exhibit Spanish dominance. In Mathematics tests, there is some evidence of English dominance in the first round, but it disappears by the second round.

Our results show that foreign-born children of Hispanic immigrants were Spanish dominant during the first round of interviews. However, between rounds, the Spanish dominance among this group declines by about one standard deviation (SD), and in some cases completely disappears, by the second round. This suggests rapid, and in some cases, complete language assimilation among foreign-born children. We find that foreign-born children, who immigrated to the U.S. after age six, show Spanish dominance in reading tests in the first round (which is consistent with Akresh and Akresh 2011). However, by the second round, only children who immigrated after age eight, show Spanish dominance in reading tests. In other words, during the six years between the two rounds of interviews, Spanish dominance has disappeared among foreign-born children who immigrated between the ages of six and eight (in reading). Moreover, for children who still have Spanish dominance in reading, the score differences have narrowed. In addition, Spanish dominance has disappeared in all children (in mathematics).

Our results also suggest that time taken to achieve final (steady state) language learning outcome depends on age at immigration. We find that almost all foreign-born children achieve the final (steady) state within four years of immigrating to the U.S. Akresh and Akresh (2011) found that foreign-born children, as young as age six at the time of immigration, perform worse when asked to take the tests in English. That result is an artifact of children not having enough time to learn English at the time of their interview (50% of children in their sample spent less than two years in the U.S. at the time of their study). In other words, the outcomes analyzed in Akresh and Akresh (2011) are intermediate outcomes and not final language learning outcomes. We discuss this issue in details in Section 3.2.2.C.

Rest of the paper is organized in the following way: Section 2 describes the data, Section 3 presents the empirical results and Section 4 concludes.

2. Data

We use data from the New Immigrant Survey (NIS)[9], which interviewed a nationally representative sample of newly admitted legal permanent residents (LPRs) to the U.S. and their family members. The first round of the NIS (conducted in 2002 and 2003) interviewed a sample of 8,573 adult immigrants who received their LPR status in the previous year. In the second round, 4,363 adult immigrants (a subset of the first round respondents) were interviewed. Jasso and Rosenzweig (2010) note that despite the high attrition rate, the second round sample remained representative of the first round. We discuss and present evidence to the same effect below.

In both rounds of the NIS, eligible children took up to four tests from the WJ-ACH III test. Children with age six or above took four tests: two reading and two mathematics tests. The two reading tests are Letter Word Identification (LWI) and Passage Comprehension (PC). Two mathematics tests are Applied Problems (AP) and Calculation (C) tests. Children between the ages of three and six took only two (LWI and AP) tests. We focus on children who took all four tests since for these children a composite score for their reading (called *Brief Reading* or BR) and mathematical (called *Brief Math* or BM) ability can be computed using "*Compuscore*" (a software that is part of WJ-ACH III evaluation[10]). Brief Reading (BR) score is a weighted average of reading (LWI and PC) tests and Brief Math (BM) score is a weighted average of mathematics (AP and Calculation) tests. We use BR and BM scores as primary outcomes of interest in our analysis. In the rest of the paper, we refer to them and reading and mathematics scores

[9] See nis.princeton.edu for additional details.
[10] Available from Houghton Mifflin Harcourt publishers at http://www.hmhco.com/hmh-assessments/cognitive-intelligence/compuscore

respectively. While we use BR and BM scores primarily for ease of exposition[11], these scores are accepted as formal evidence of achievement[12].

For our analysis, we create two samples. Our first sample is a balanced panel, which consists of children for whom complete information from both rounds are available. Our second sample is an unbalanced panel, which includes all available observations with complete information in at least one of the rounds. Table 1 shows how we arrive at our final sample sizes both for the balanced panel and for the unbalanced panel.

As part of the NIS interview, 3856 children of immigrants took the WJ-ACH III tests in the first round (2003) and 1503 children of immigrants took the WJ-ACH III test in the second round (2009). In each round, Children of Spanish origin (parent born in a Spanish speaking country or first language Spanish) were randomly assigned to take the tests either in English or in Spanish (language experiment). In the first (second) round, 1445 (557) children were part of language experiment. Out of them, 1036 (519) children met the age criteria (age six or older at the time of interview) to take all four tests in the first (second) round. Among them, 875 (518) finished the WJ-ACH tests and therefore have valid scores in the first (second) round. Finally, we exclude 59 (35) children from the first (second) because their country of birth or date of birth is not available, leaving us with a total of 1299 observations. Among them, 816 (483) observations are from the first (second) round. If we impose the criteria that the children would have to present in both interviews then we end up with 604 observations from 302 children (balanced panel).

Our unbalanced (balanced) panel consists of 1299 (604) observations. Out of them, 938 (438) observations are from U.S. born children and 361 (166) are from foreign-born children. First, we check whether the randomization was effective in both rounds. Akresh and Akresh (2011) have shown that the randomization was effective in the first round. They check the balance for the combined sample (U.S. born and foreign-born children taken together). However, since the analysis is carried out separately for U.S. born and foreign-born children, we test whether the randomization was preserved in our analysis samples.

[11] We conducted analysis using LWI, PC, AP, and C scores as separate variables (available on request). They provide broadly similar result.

[12] Please see http://www.hmhco.com/hmh-assessments/cognitive-intelligence/wjiii-nu-brief-battery for more details.

We also test whether the randomization was effective in the second round, which was not used in Akresh and Akresh (2011). Table 2 presents the results for balancing test for U.S. born children (Panel A for balanced panel and Panel B unbalanced panel). Table 3 presents the corresponding results for foreign-born children. Columns 1 and 2 (4 and 5) shows the summary statistics for children who took the tests in English and Spanish respectively in the first (second) round. Column 3 (6) shows the difference between Columns 1 and 2 (4 and 5). We check whether there is any difference between groups in the characteristics of the children (gender, age, years spent in the U.S., and age at the time of immigrating to the U.S.) or their parents[13] (years of education, whether they have any U.S. education, years spent in the U.S., English ability, and whether they adjusted their visa status to LPR). If randomization were effective, then we would expect that there would no significant difference in these variables. That is indeed what we find in both Panel A and Panel B. Two exceptions are gender distribution (in the second round) and years in the U.S. by parent (in the first round) in the U.S. born children unbalanced panel sample. Therefore, we conclude that randomization was effective in both rounds, and randomization was preserved in our analysis samples.

3. Results

We use age adjusted standardized scores computed by "*Compuscore*". "*Compuscore*" is a software (that is part of WJ-ACH III evaluation) that produces standardized test scores by comparing raw test scores against a national sample. The standardized scores have a population mean of 100 and population standard deviation is 15 at any given age. We present the analysis for U.S. born children in Section 3.1 and foreign-born children in Section 3.2.

3.1. U.S. Born Children

3.1.1. Language effect

The sample averages of standardized scores for U.S. born children are presented in Table 4. Panel A (B) presents the results for the balanced (unbalanced) panel. Columns 1-

[13] For parental characteristics, we use characteristic from the parent who was the principal immigrant in the family.

3 present the results for the first round and columns 4-6 present the results for second round. Given that, in each round, the testing language was randomly assigned, difference in means (column 3 for the first round, and column 6 for the second round) represents the causal effect of language on test scores. Column 7 represents the change in language effect between rounds.

Panel A (Column 3) shows that in the first round, U.S. born children who were assigned to take the tests in English scored 7.90 points more in reading BR than children who were assigned to take the test in Spanish (101.13 vs. 93.24). In Mathematics, the difference was smaller (3.22 points) and the difference was not statistically significant. Columns 4-6 present the results from the second wave when the same tests taken by the same individuals about six years after the first wave. The average BR score for U.S. born children who were assigned to take the tests in English (Spanish) in the second round is 94.38 (86.62). The difference 7.77 points remained similar to the difference from the first round. In math (BM), the difference is insignificant in the second wave as well. Column 7 presents the change in difference across waves. Tests suggest that the difference in achievement gap (column (3) – column (6)) due to language did not change between the two waves for U.S. born children of Hispanic immigrants. In both rounds, U.S. born children of Hispanic immigrants show English dominance in reading tests (BR) and the magnitude of English dominance does not change over time. In the mathematics tests (BM), the score difference are small and statistically insignificant in both rounds. The change in difference in not significant either.

Panel (B) presents the same descriptive statistics for the unbalanced sample. As discussed above, the unbalanced panel include all available observations and therefore is substantially bigger. Sample sizes are about twice that of the balanced sample. We first note that the average scores for all the tests for all groups are similar in Panel A and Panel B. This is reassuring because given the smaller sample size (especially in the second wave) one might be concerned about non-random attrition. However, comparing Panels A and B suggests that test scores for balanced and unbalanced samples are similar despite high attrition rate. One exception is that the score differential in mathematics (BM) score is significant in the first round for the unbalanced sample.

In each round, participants were randomly assigned to take the tests either in English or in Spanish. Therefore, the difference in means is a consistent estimate of language effect. However, controlling for additional covariates may reduce sampling variance. The regression results are presented in Table 5. Columns 1 and 2 of Panel A present the OLS estimates for balanced panel sample. As one would expect with randomization, the OLS results are similar to difference in means reported in Table 4. For the U.S. born children there is significant English dominance in reading (BR) score in both rounds and the magnitude does not change across rounds (Column 1). In mathematics (BM), there is no effect of language of test on test scores in either round. Columns 3 and 4 of Panel A present results for unbalanced panel sample. The results are similar to columns 1 and 2, with one exception. In the first round, those who were assigned to take the mathematics tests in English scored significantly more that those were assigned to take the mathematics tests in Spanish (column 4). The difference is not significant in the second round anymore. This is consistent with unbalanced panel results (Panel B) in Table 4.

Since the language of test was randomly assigned prior to each round of interview, same child may have taken tests in different language in different round. For example, a child who was randomly assigned to take the tests in English in first round may be assigned to take the tests in Spanish in the second round. Thus, the assignment of language for a given child across interviews is also random. Therefore, we can implement a first-difference (FD) regression using the balanced panel. Panel B of Table 5 shows the results from first difference (FD) regression. Results for reading (BR) scores are similar to OLS. For mathematics (BM) score, the language effect is significant in the first round, which is different from OLS results for balanced sample, but consistent with the OLS results for unbalanced sample (column 4 of Panel A).

3.1.2 Heterogeneity in the language effect

Our discussion so far shows that for U.S. born immigrant children, there is no change in the nature of language effect on reading scores between two rounds. However, it seems that the English dominance in mathematics scores has disappeared between rounds. In this section, we explore how the language effect varies across score distribution and across age distribution.

3.1.2.1. Language effect and score distribution

We estimate an Unconditional Quantile Regression (UQR) model, developed by Firpo *et al*. (2009), to check how the effect of language on test scores varies across the unconditional test score distributions. Figure 1 presents the coefficient estimates (and 95% confidence intervals) from UQR regressions[14] using the balanced panel sample. The corresponding figure for the unbalanced panel sample is presented in Appendix Figure A1. The left panel shows the results for reading (BR) score and the right panel shows the results for mathematics (BM) score. In the left panel, there are two figures: one for the first round and one for the second round. The first figure shows that in the first round, English is most dominant at the bottom end of BR score distribution. The score differential (between test takers in English as opposed to Spanish) declines as BR score increases. At the 90th percentile of the BR score distribution, the point estimates suggest Spanish dominance, although it is not statistically significant. The second round results are similar. English dominance is at its strongest at the 10th percentile of BR score distribution. However, at the 80th and 90th percentiles, there is Spanish dominance, and they are statistically significant at 95 percent level of significance. These results suggest that English dominance reported in Tables 4 and 5 using mean based estimators mask substantial heterogeneity.

The right panel also has two figures: one for the first round and one for the second round mathematics (BM) scores. The first round results seems to suggest English dominance at the bottom end of the mathematics (BM) score distribution. This is not very clear in Figure 1 (balanced panel), but it is in Figure A1 (unbalanced) panel. In the second round, we do not see any language effect on mathematics (BM) score in any part of the distribution.

3.1.2.2. Language effect and age

Next, we explore how the language effect varies with age. Figure 2 show the results from kernel weighted local polynomial regressions of reading (BR) and mathematics (BM) scores (with 95% confidence intervals) for the balanced panel. Results

[14] Confidence Intervals are based on bootstrapped standard errors (500 replications).

for the unbalanced panel are in Appendix Figure A2. Age of the children varies between 6 and 12 at the time of the first round and between 6 and 17 at the time of second round of surveys[15]. We pool the data from both rounds to check how the language effect varies over the age-range 6 to 17. Results in both Panel A (reading) and Panel B (mathematics) show that the average standardized score of U.S. born children of Hispanic immigrants declines with age. It holds for both languages (English and Spanish). Note that these are standardized scores, and therefore the population average should be 100 at every age. This may suggest that the reading and mathematical skills of U.S. born children of Hispanic origin do not progress at the national average rate.

Panel A suggests that at each age, the average BR score of those who were assigned to the tests in English is higher than those who were assigned to the tests in Spanish. This is consistent with English dominance among U.S. born children reported in Tables 4 and 5. However, the difference starts to narrow after age 14, and for children above age 16 English dominance is not statistically significant. This pattern more distinctly visible in the unbalanced panel regression results (Figure A2, Panel A). In Figure A2, English dominance is present at all ages below 15, but not in children above 15. These results are consistent with pattern of racial homophily reported in the literature. For example, Aboud, Mendelson, and Purdy (2003) found that cross-racial friendships decline with age. Shrum et al. (1988) found that cross-racial friendship are less common than what would be expected given the demographics of a school. However, within that context, cross-racial friendship are relatively common in elementary school and declines steadily after that. For example, number of observed cross-racial friendship is about 65% of what we would expect given the demographics in the third grade. It declines to about 10% by the end of middle school and remains there during the high school years (Shrum et al, 1988). It is also consistent with Angerer at al. (2016) who conducted a field experiment and found that inclination among children to cooperate with other children who speak the same language, and discriminate against children who speak a different language increases with age. Panel B shows the corresponding results for mathematics

[15] Technically age varies between 3 and 12 at the time of first round and between 6 and 18 at the time of second round of surveys. However, we exclude children below six as they are not eligible to take all four WJ ACH III tests.

(BM) score. BM scores are similar for those who are assigned to take the tests in English and those who are assigned to take the tests in Spanish. This is consistent with Tables 4 and 5.

Therefore, our analysis suggests that there is heterogeneity in reading score across the unconditional BR score distribution and across age of the children.

3.2. Foreign-born children

3.2.1. Language effect

Tables 6 and 7 present the differences in means and regression estimates for foreign-born children. Tables 6 and 7 (foreign-born children) have a structure that is similar to Tables 4 and 5 (U.S. born children). The sample averages of standardized scores for foreign-born children are presented in Table 6. Panel A (B) presents the results for the balanced (unbalanced) panel for foreign born children. Columns 1-3 present the results for the first round and columns 4-6 present the results for second round. Given that in each round the testing language was randomly assigned, difference in means (column 3 for the first round, and column 6 for the second round) represents the causal effect of language on test scores. Column 7 represents the change in language effect between rounds.

Children, who took the tests in English in the first round, scored lower that those who took the tests in Spanish. In reading (BR), the disadvantage is 25.57 points, and in mathematics (BM), the disadvantage is 13.45 points. The second round results (columns 4-6) suggest that the English disadvantage in reading has decreased to 8.58 points (from 25.57 points in the first round). Therefore, the change in English disadvantage (shown in column 7) is 16.99 points and it is statistically significant. The English disadvantage in the mathematics (BM) score also declined by a statistically significant 11.73 points. In fact, the English disadvantage in BM score has completely disappeared by second round (column 6).

Panel B presents the descriptive statistics for the unbalanced panel sample. As discussed above, the unbalanced panel include all available observations, and therefore has a bigger sample size. The sample averages as well as language effect patterns are similar to the results reported above. The only difference is that the change in language effect in mathematics (BM) is not statistically significant.

Table 7 presents regression results. Columns 1 (reading) and 2 (mathematics) show the results from OLS regression based on balanced panel sample. A Note below Table 7 shows the complete list of covariates in the regressions. Estimates suggest that English disadvantage faced by foreign-born children declined dramatically and in some cases (mathematics) completely disappeared by the second round. As expected with experimental data, inclusion of covariates have little effect on estimates of change in language effect across rounds, and the standard errors are slightly lower. The regression results for unbalanced panel sample (columns 3 and 4) suggest that English disadvantage have disappeared completely in both reading and mathematics. The regression results are also consistent with mean differences reported in Table 6. One exception is that the decline in English disadvantage in unbalanced sample (column 4) is now significant. Panel B presents the FD regressions. The results are similar to OLS.

3.2.2. Heterogeneity in the language effect

Our discussion so far shows that for young foreign-born immigrants, the years in the U.S. in between two rounds of interviews significantly improved their English skills. In this section, we explore heterogeneity in the language effect among foreign-born children. In particular, we are interested in how language effect varies by i) quantiles of unconditional score distribution, ii) age at immigration, and iii) number of years in the U.S. Note that while the first exercise is same as the U.S. born children (Section 3.1.2.1), the last two are different. For foreign-born children, age of a child is summation of age at immigration and number of years in the U.S. Both age at immigration and years in the U.S. may have an independent effect on English acquisition. There may an interaction effect as well. We explore these in this Section.

3.2.2.1. Language effect and score distribution

We estimate an Unconditional Quantile Regression (UQR) model for foreign-born children as well. Figure 3 presents the coefficient estimates (and 95% confidence intervals) from UQR regressions[16]. The corresponding figure for the unbalanced panel sample is presented in Appendix Figure A3. The left panel shows the results for reading (BR) score and the right panel shows the results for mathematics (BM) score. In the left panel, there are two figures: one for the first round and one for the second round. The first figure shows that in the first round, the language effect was evident across the whole spectrum of reading (BR) score distribution. However, the second figure in the left panel shows children in the middle of the distribution are most English deficient. Children in either end of score distribution exhibit little to no English deficiency.

The right panel also has two figures: one for the first round and one for the second round mathematics scores. The first figure (of the right panel) shows that children below the 20th percentile of BM score distribution do significantly worse when the tests are given in English. However students above the 20th percentile do equally well in Spanish and English. Therefore, unlike reading (where the language effect is present in all students), the language ability only affects students with relatively low math scores. By the second round, all foreign-born students do equally well whether they take the math tests in Spanish or English. These results suggest that there is substantial variability in the language effect across score distributions and across test subjects.

3.2.2.2. Age of immigration and language acquisition

Since we can observe same foreign-born immigrant children about six years apart, we can directly estimate how the rate of language assimilation differs by age at the time of immigrating to the U.S. Figures 4 and 5 presents the results from kernel weighted local polynomial regressions of reading and math scores respectively for the balanced panel. The corresponding figures for the unbalanced panel are in appendix figures A4 and A5.

Figure 4 shows the results from kernel weighted local polynomial regressions for reading (BR) scores with 95% confidence intervals. The left panel is for the first round, and right panel is for the second round. The bold line shows the scores for children who took the reading tests in English and the shaded area represents 95% confidence interval around the estimates. The dashed lines show the scores (along with 95% confidence

[16] Confidence Intervals are based on bootstrapped standard errors (500 replications).

interval) for children who took the tests in Spanish. First round results (left panel) show that for children who immigrated before age six, the test scores are not different based on test language. However, for children who immigrated to the U.S. after age six, test scores begin to diverge as children who were assigned to take the tests in English do significantly worse than children who were assigned to take the tests in Spanish. This is consistent with Akresh and Akresh (2011). The right panel shows the test scores for the *same* children about six years after the first test (second round). This time the test scores are significantly different only for children who immigrated after age eight or so. Furthermore, for children who still have Spanish dominance (above age eight) the score differences have declined. Comparing the left panel and right panel also shows that the solid line (tests scores for children who took the tests in English) has become flatter in the second round, suggesting that children who arrived in the U.S. after age six, are accumulating English skills. The dashed line (test scores for children who the tests in Spanish) also becomes flatter suggesting their Spanish skills are depreciating as they spend more time in the U.S. Taken together, these lead to a decline in Spanish dominance among foreign-born children above age six.

Figures 5 presents the corresponding figure for the mathematics (BM) scores. In general, the effect of language on scores is less pronounced for mathematics (BM) scores compared to reading scores. This is consistent with results in Table 6. Figure 5 suggests that in the first round, the mathematics scores of children who immigrated after age seven, and took the tests in Spanish were significantly higher than their counterparts who took the tests in English. The right panel of Figure 5 suggests that there the language effect on BM scores have disappeared by the second round. Figures (A4 and A5) based on unbalanced sample suggest similar results.

3.2.2.3. Years in the U.S. and language acquisition

Our analysis of Figure 4 (balanced panel) and Figure A4 (unbalanced panel) suggests for foreign-born children who immigrated to the U.S. after age eight, the language effect persists (on an average) in reading scores even after spending at least six years in the U.S. Please note that the average number of years spent in the U.S. at the time of second round of interviews is about nine years (see Columns 4 and 5 of Table 3), which is comparable to Johnson and Newport (1989) and Birdsong and Molis (2001). Therefore, our results seems to be consistent with the supposition that language-learning window begins to close after age eight, which is consistent with Lenneberg (1967) and Johnson and Newport (1989). On the other hand, our results may not be consistent with Birdsong and Molis (2001) who argue that language-learning window may extend to age 15 or so. To provide a definite answer, we explore how long does it take for the language acquisition process to complete in child immigrants. Is there a possibility that even more time (after at least six years) in the U.S. may result in children who immigrated after the age of eight becoming equally proficient in English and Spanish. For this purpose, we exploit the fact different children in our sample have spent different number of years in the U.S. We focus on reading scores, since at the time of the second round of interviews we observe a language effect only in reading scores of children who immigrated after the age of eight. We pool the observations from two rounds together and we use the unbalanced panel to maximize our sample size.

Figure 6 shows that the first four years spent in the U.S. are most important for English language acquisition among foreign-born children. However, after the fourth year the marginal benefit of spending more time in the U.S. in negligible. Figure 6 also suggest that among those who have spent more than four years in the U.S., there is no significant difference between the BR scores of those who were assigned to take the tests in English and those who were assigned to take the tests in Spanish.

The marginal improvement, on an average, in English skills after spending more than four years in the U.S. seems negligible for foreign-born children. However, it may be that children who immigrated to the U.S. at an older age (and therefore have a bigger English deficiency to start with) may continue to improve their English skills even after their fourth year in the U.S. To explore this we plot how reading (BR) scores vary with years in the U.S. conditional on age at immigration. Ideally, we would like to plot one figure for each age at immigration. However, given our sample size restrictions, we plot for two or three year bands (cohorts) of age-at-immigration groups. Figure 7 shows how the BR scores vary with years in the U.S. for four groups: age at immigration four or five, six or seven, eight or nine, and ten or eleven. Since all children were at least six years old at the time of the first round of interviews all children in panel A (age at immigration four or five) have already spent at least two years in the U.S. at the time of first round of interviews. For children who immigrated before the age of seven (Panels A and B), we do not find any significant language effect on BR score. For children who immigrated at age eight or nine (Panel C) or 10 or 11 (Panel D), the newly arrived immigrant children perform significantly worse when they are assigned to take the tests in English. However, the difference narrows with years in the U.S., especially for the first four years. After that, the score gap becomes insignificant and remains stable. In Figure 8, where we group children based on three-year bands, show similar results. For children who immigrated after the age of nine, score gap remains not only significant, but also stable after four years.

Next, we explore how the six years in between interviews altered the language effect for a given child with a specific age at immigration, and years in the U.S. at the time of the first interview. We use FD regressions. Based on our results so far we split the foreign-born children sample into two groups: those who immigrated at or before age eight, and those who immigrated after age eight. Each of these two groups are then further sub-divided based on how many years they have spent at the time of the first round of interviews: zero to three or four and above. Since all children were at or below 12 at the time of the first round, we do not have any children who immigrated after age eight (i.e. nine and above) and spent four or more years in the U.S. at the time of first round of interviews. The regression results for the three remaining groups are presented in Table 8. Since these are FD regressions, they are based on the balanced panel sample. The sample sizes for each of these regressions are small but the results show that standard errors are not large, especially relative to the point estimates. The results suggest that children who were eight or below at the time of immigration and had spent less than four years at time of the first round of interviews (column 1), were English deficient at the time of first round of interviews. However, that deficiency disappeared by the time of the second round. Children who were 8 or below at the time of immigration and had spent four or more years at time of the first round of interviews (column 2), were equally proficient in English and Spanish at the time of first round of interviews and that result did change in the second round. Finally, that children who were above 8 at the time of immigration and had spent less than four years at time of the first round of interviews (column 3), were English deficient at the time of first round of interviews. For this group English deficiency narrowed but they remain English deficient even at the time of the second round. Taken with results of Figures 6-8, this suggests that these children may never become completely proficient in English.

Using the first round data from the NIS, Akresh and Akresh (2011) found that children as young as age six at the time of immigration, may not have native-like English skills. Foreign-born children have spent about 3.3 years in the U.S. (on an average) at the time of first round of survey. Furthermore, about 50% of foreign–born children have spent two years or less at the time of first round of survey. Our results suggest that children in their sample did not have enough time to learn English to reach final (steady) state. This is especially true for children who immigrated after the age of eight. Our results suggest that children (especially those immigrate after the age of eight) need to spend at least four years before the score differentials becomes stable. In other words, the outcomes analyzed in Akresh and Akresh (2011) are intermediate language learning outcomes, and not final language learning outcomes. By the time of the second round of interviews, all children have spent at least six years in the U.S. Therefore, the results from second wave can be considered final language learning outcome. In other words, the score differentials that we observe at the time of second round of interviews are likely to persist.

Before we conclude, it is worth noting that from a policy perspective, both intermediate and final language learning outcomes are important as schools decide how to best integrate foreign-born immigrant children in different stages of their lives. Our results also suggest that language effect depends on the subject of the test; language-effect is more of an issue in reading tests compared to mathematics tests. This suggests that schools should be more careful in interpreting scores of reading tests of immigrant children, especially those who have spent less than four years in the host country.

4. Conclusion

In this paper, we provide new evidence about rate of language assimilation using data from two rounds of the New Immigrants Survey. In both rounds, children were asked to take the Woodcock-Johnson achievement test (WJ-ACH III). In each round, some children of Hispanic origin were randomly assigned to take the tests in Spanish, while others were assigned to take the tests in English. Therefore, the difference in scores can be *causally* attributed to language proficiency. The longitudinal structure allows us to estimate the rate of language assimilation among U.S. born and foreign-born children of Hispanic origin.

Our results suggest that the English dominance in reading tests remain stable across two rounds of interviews among U.S. born children of Hispanic immigrants. However, our results also suggest that average score differentials mask substantial heterogeneity. While U.S. born children of Hispanic immigrants at the bottom end (below 50^{th} percentile) of the reading score distribution show English dominance, those in the middle (between 50^{th} and 80^{th} percentile) of reading score distribution are equally proficient in English and Spanish, and those at the top (90^{th} percentile or above) show Spanish dominance. In Mathematics tests, we find some evidence of English dominance in the first round but none in the second round.

Our results on foreign-born children of Hispanic immigrants show that the Spanish dominance among this group declines significantly, and in some cases completely disappears by the second round. This suggests rapid, and in some cases, complete language assimilation among foreign-born children. We find that while foreign-born children of Hispanic immigrants who immigrated to the U.S. after age six show Spanish dominance in reading tests in the first round, by the second round, only children who immigrated after age eight show Spanish dominance. In other words, during the six years between interviews, Spanish dominance has disappeared among foreign-born children who immigrated between the ages of six and eight (in reading) and in all children (in math). Moreover, for children who still have Spanish dominance in reading, the score differences have narrowed. Our results suggest that children (especially those immigrate after the age of eight) need to spend at least four years before they achieve final (steady state) language learning outcome. Our results also suggest that the differentials observed after spending four years in the host country (U.S.) are likely to persist in the future as well. In other words, foreign-born children who immigrated to the U.S. after age eight may always be at a disadvantage when taking tests in English.

References

Aboud, F., Mendelson, M., & Purdy, K. (2003). Cross-race peer relations and friendship quality. *International Journal of Behavioral Development*, *27*(2), 165-173.

Akresh, R., & Akresh, I. R. (2011). Using achievement tests to measure language assimilation and language bias among the children of immigrants. *Journal of Human Resources*, *46*(3), 647-667

Angerer, S., Glätzle-Rützler, D., Lergetporer, P., & Sutter, M. (2016). Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city. *European Economic Review*, *90*, 254-264.

Berman, Eli & Lang, Kevin & Siniver, Erez, 2003. Language-skill complementarity: returns to immigrant language acquisition, *Labour Economics*, Elsevier, vol. 10(3), pages 265-290, June.

Birdsong, David. Age and second language acquisition and processing: A selective overview. *Language Learning* 2006; 56:9–49.

Birdsong, D., & Molis, M. (2001). On the Evidence for Maturational Constraints in Second-Language Acquisition. *Journal of Memory and Language*, *44*(2), 235-249. DOI: 10.1006/jmla.2000.2750

Bleakley, Hoyt, and Aimee Chin. 2004. "Language Skills and Earnings: Evidence from Childhood Immigrants." *Review of Economics and Statistics* 86(2):481–96

———. 2010. "Age at Arrival, English Proficiency, and Social Assimilation among U. S. Immigrants." *American Economic Journal: Applied Economics* 2(1):165–92.

Brind, Tom, Caroline Harper & Karen Moore. 2008. 'Education for migrant, minority and marginalised children in Europe. A report commissioned by the Open Society Institute's Education Support Programme.' Open Society Foundations, 31 January. Available at https://www.opensocietyfoundations.org/sites/default/files/review_20080131.pdf

Calderón, M., Slavin, R., & Sánchez, M. (2011). Effective instruction for English learners. *The Future of Children*, *21*(1), 103-127.

Child Trends. 2014. "Immigrant Children." Retrieved April 11, 2018, available at (https://www.childtrends.org/wp-content/uploads/2013/07/110_Immigrant_Children.pdf).

Conger, D. (2009). Testing, time limits, and English learners: Does age of school entry affect how quickly students can learn English?. *Social Science Research*, *38*(2), 383-396.

De Paola, M., & Brunello, G. (2016). Education as a tool for the economic integration of migrants.

Dustmann, Christian, Tommaso Frattini & Gianandrea Lanzara. 2012. 'Educational Achievement of Second-Generation Immigrants: An International Comparison. *Economic Policy* 27: 143–185.

Dustmann, Christian and van Soest, Arthur, (2002), Language and the Earnings of Immigrants, *ILR Review*, 55, issue 3, p. 473-492.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953-973.

García, O., Kleifgen, J. A., & Falchi, L. (2008). From English Language Learners to Emergent Bilinguals. Equity Matters. Research Review No. 1. *Campaign for Educational Equity, Teachers College, Columbia University*.

Harte, E., Herrera, F., & Stepanek, M. (2016). Education of EU migrant children in EU Member States. RAND. Available at https://www.rand.org/content/dam/rand/pubs/research_reports/RR1700/RR1715/RAND_RR1715.pdf

Heubert, J., & Hauser, R. (Eds.). (1999). High stakes testing for tracking, promotion, and graduation. Washington, DC: National Academy Press.

Jakubowski, Maciej. 2011. 'How are school systems adapting to increasing numbers of immigrant students?' *PISA in Focus* 2011/11. Available at http://www.oecd.org/pisa/pisaproducts/pisainfocus/49264831.pdf

Janta, B., & Harte, E. (2016). Education of migrant children: education policy responses for the inclusion of migrant children in Europe. RAND. Available at https://www.rand.org/content/dam/rand/pubs/research_reports/RR1600/RR1655/RAND_RR1655.pdf

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60-99.

Kindler, A. L. (2002). Survey of the states' limited English proficient students and available educational programs and services: 2000–2001 summary report. Washington, DC: National Clearinghouse for English Language Acquisition

Kossoudji, Sherrie A, (1988), English Language Ability and the Labor Market Opportunities of Hispanic and East Asian Immigrant Men, *Journal of Labor Economics*, 6, issue 2, p. 205-28,

Lenneberg, Eric H. Biological Foundation of Language. Wiley & Sons; New York: 1967.

Newport, Elissa L. Critical Periods in Language Development. In: Nadel, L., editor. Encyclopedia of Cognitive Science. Macmillan Publishers Ltd./ Nature Publishing Group; London: 2002.

Nouwen, Ward, Noel Clycq & Daniela Ulična. 2015. 'Reducing the Risk that Youth with a Migrant Background in Europe will Leave School Early.' *SIRIUS Network Policy Brief Series* 6. Available at http://www.sirius-migrationeducation.org/wp content/uploads/2015/02/SIRIUSEarlySchoolLeaving-FINAL.pdf

OECD. 2010. 'Equal opportunities? The labour market integration of the children of immigrants.' OECD, 12 May. Available at http://www.oecd-ilibrary.org/social-issues-migration-health/equal-opportunities_9789264086395-en

Rolstad, K., Mahoney, K., & Glass, G. V. (2005). The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational policy*, *19*(4), 572-594.

Schnepf, S. (2007). Immigrants' educational disadvantage: An examination across ten countries and three surveys. *Journal of Population Economics, 20*(3), 527-545.

Shepard, L. A., & Smith M. L. (Eds.). (1989). Flunking grades: Research and policies on retention. Philadelphia: The Falmer Press.

Shrum, W., Cheek Jr, N. H., & MacD, S. (1988). Friendship in school: Gender and racial homophily. *Sociology of Education*, 227-239.

Slama, R. B. (2014). Investigating whether and when English learners are reclassified into mainstream classrooms in the United States: A discrete-time survival analysis. *American Educational Research Journal*, *51*(2), 220-252.

Sullivan, P., Yeager, M., Chudowsky, N., Kober, N., O'Brien, E., & Gayler, K. (2005). State high school exit exams: States try harder, but gaps persist. Washington, DC: Center on Education Policy.

Tainer, E. (1988). English language proficiency and the determination of earnings among foreign-born men. *Journal of Human Resources*, 108-122.

Table 1: Sample criteria and sample sizes

	Round 1	Round2	Total	Name of sample
# of children who took the WJ ACH III tests	3856	1503		
# of children part of Language experiment (randomization)	1445	557		
# Eligible for all four tests(Age>=6 at the time of interview)	1036	519		
# Completed WJ tests	875	518		
& Other Baseline covariates available	816	483	1299	Unbalanced Panel
& were present in both interviews	302	302	604	Balanced Panel

Table 2: Balancing test for covariates: U.S. born balanced and unbalanced panel samples

Panel A: Balanced Panel Sample						
	First Round			Second Round		
	English	Spanish	Difference	English	Spanish	Difference
Age at immigration	-	-	-	-	-	-
	-	-	-	-	-	-
Years in the U.S.	-	-	-	-	-	-
	-	-	-	-	-	-
Age	9.29	9.01	0.275	14.28	14.10	0.183
	[1.88]	[1.96]	(0.261)	[1.78]	[1.77]	(0.240)
Female	0.50	0.59	-0.0914	0.59	0.50	0.0909
	[0.50]	[0.49]	(0.0682)	[0.49]	[0.50]	(0.0675)
Parent Characteristics						
Educ. (yrs.)	8.43	8.05	0.375	8.39	8.16	0.238
	[3.76]	[3.71]	(0.511)	[3.49]	[3.95]	(0.506)
Adjustee	0.97	0.96	0.0113	0.96	0.97	-0.00368
	[0.18]	[0.20]	(0.0258)	[0.19]	[0.18]	(0.0255)
years in the U.S.	13.33	13.82	-0.492	18.99	19.40	-0.410
	[5.24]	[5.06]	(0.706)	[4.37]	[5.64]	(0.687)
Any U.S. educ.	0.18	0.13	0.0535	0.16	0.16	0.00694
	[0.39]	[0.34]	(0.0502)	[0.37]	[0.36]	(0.0498)
English below avg.	0.84	0.77	0.0671	0.79	0.83	-0.0463
	[0.37]	[0.42]	(0.0534)	[0.41]	[0.37]	(0.0529)
N	126	93	219	104	115	219
Panel B: Unbalanced Panel Sample						
Age at immigration	-	-	-	-	-	-
	-	-	-	-	-	-
Years in the U.S.	-	-	-	-	-	-
	-	-	-	-	-	-
Age	9.35	9.52	-0.173	12.91	12.89	0.0159
	[2.08]	[2.17]	(0.178)	[2.57]	[2.53]	(0.267)
Female	0.48	0.55	-0.0640	0.57	0.48	0.0912*
	[0.50]	[0.50]	(0.0420)	[0.50]	[0.50]	(0.0521)
Parent Characteristics						
Educ. (yrs.)	8.71	8.41	0.296	8.81	8.53	0.280
	[4.06]	[3.95]	(0.337)	[3.85]	[4.10]	(0.415)
Adjustee	0.95	0.95	-0.00180	0.96	0.95	0.00268
	[0.22]	[0.22]	(0.0185)	[0.21]	[0.21]	(0.0220)
years in the U.S.	13.43	14.28	-0.844*	18.21	18.41	-0.200
	[5.47]	[4.91]	(0.438)	[5.22]	[5.73]	(0.573)
Any U.S. educ.	0.22	0.21	0.0125	0.21	0.19	0.0174
	[0.41]	[0.41]	(0.0345)	[0.41]	[0.39]	(0.0418)
English below avg.	0.81	0.80	0.0127	0.75	0.78	-0.0250
	[0.39]	[0.40]	(0.0334)	[0.43]	[0.42]	(0.0443)
N	310	261	571	178	189	367

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard deviations are in brackets and standard errors are in parenthesis.

Table 3: Balancing test for covariates: foreign-born balanced and unbalanced panel samples

Panel A: Balanced Panel Sample						
	First Round			Second Round		
	English	Spanish	Difference	English	Spanish	Difference
Age at immigration	6.05	6.74	-0.694	6.46	6.35	0.105
	[3.62]	[3.57]	(0.789)	[3.75]	[3.43]	(0.797)
Years in the U.S.	3.58	3.51	0.0634	9.39	9.14	0.256
	[3.61]	[3.95]	(0.833)	[3.87]	[3.43]	(0.813)
Age	9.63	10.26	-0.631	15.13	14.78	0.347
	[1.78]	[1.73]	(0.386)	[1.63]	[1.65]	(0.362)
Female	0.43	0.30	0.123	0.35	0.38	-0.031
	[0.50]	[0.46]	(0.106)	[0.48]	[0.49]	(0.107)
Parent Characteristics						
Educ. (yrs.)	10.73	10.16	0.562	10.07	10.89	-0.827
	[4.41]	[4.51]	(0.981)	[4.19]	[4.77]	(0.983)
Adjustee	0.47	0.53	-0.0599	0.50	0.51	-0.013
	[0.51]	[0.50]	(0.111)	[0.51]	[0.51]	(0.112)
years in the U.S.	4.95	5.21	-0.259	10.76	10.89	-0.131
	[5.26]	[7.91]	(1.487)	[7.16]	[6.06]	(1.478)
Any U.S. educ.	0.13	0.16	-0.0378	0.17	0.11	0.0658
	[0.33]	[0.37]	(0.0781)	[0.38]	[0.31]	(0.078)
English below avg.	0.82	0.77	0.0576	0.80	0.78	0.0206
	[0.38]	[0.43]	(0.0895)	[0.40]	[0.42]	(0.090)
N	40	43	83	46	37	83
Panel B: Unbalanced Panel Sample						
Age at immigration	7.26	6.85	0.409	5.98	5.76	0.220
	[3.37]	[3.50]	(0.439)	[3.49]	[3.14]	(0.625)
Years in the U.S.	3.10	3.58	-0.481	8.52	8.51	0.0133
	[3.15]	[3.61]	(0.433)	[3.55]	[3.20]	(0.637)
Age	10.36	10.43	-0.0720	13.89	13.73	0.167
	[1.94]	[2.00]	(0.252)	[2.70]	[2.45]	(0.486)
Female	1.46	1.43	0.0354	1.43	1.31	0.117
	[0.50]	[0.50]	(0.0637)	[0.50]	[0.47]	(0.0909)
Parent Characteristics						
Educ. (yrs.)	11.44	10.68	0.761	10.22	10.88	-0.667
	[4.79]	[4.46]	(0.591)	[4.26]	[4.84]	(0.845)
Adjustee	0.49	0.52	-0.0366	0.42	0.45	-0.0356
	[0.50]	[0.50]	(0.0641)	[0.50]	[0.50]	(0.0934)
years in the U.S.	4.60	4.67	-0.0743	9.52	10.33	-0.810
	[6.19]	[6.44]	(0.807)	[6.35]	[5.94]	(1.155)
Any U.S. educ.	0.20	0.14	0.0613	0.14	0.10	0.0404
	[0.40]	[0.35]	(0.0477)	[0.35]	[0.30]	(0.0614)
English below avg.	0.74	0.80	-0.0628	0.82	0.76	0.0507
	[0.44]	[0.40]	(0.0541)	[0.39]	[0.43]	(0.0763)
N	121	124	245	65	51	116

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard deviations are in brackets and standard errors are in parenthesis.

Table 4: Change in Difference in Means by language of test: U.S. born children

Panel A: Balanced Panel Sample							
	First Round			Second Round			
	English	Spanish	Difference	English	Spanish	Difference	Change in Difference
	(c1)	(c2)	(c3)	(c4)	(c5)	(c6)	(c7)
			(c1)-(c2)			(c4)-(c5)	(c6)-(c3)
Reading	101.13	93.24	7.898**	94.38	86.62	7.767**	-0.13
	[17.35]	[30.12]	(3.229)	[14.60]	[29.67]	(3.213)	(4.55)
Math	106.24	103.02	3.217	96.49	95.73	0.760	-2.45
	[18.49]	[27.56]	(3.114)	[16.86]	[19.40]	(2.469)	(3.97)
N	126	93	219	104	115	219	438
Panel B: Unbalanced Panel Sample							
Reading	101.4	93.62	7.780***	95.75	84.43	11.32***	3.54
	[17.08]	[28.46]	(1.932)	[14.83]	[30.95]	(2.558)	(3.16)
Math	106.28	101.55	4.729**	98.56	96.87	1.688	-3.04
	[19.38]	[26.91]	(1.943)	[17.62]	[20.29]	(1.989)	(2.90)
N	310	261	571	178	189	367	938

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard deviations are in brackets and standard errors are in parenthesis.

Table 5: Regression estimates: U.S. born children

	Balanced panel		Unbalanced panel	
	Reading	Mathematics	Reading	Mathematics
Panel A: OLS				
English* I(year=1)	7.894**	2.577	7.631***	4.056**
	(3.378)	(3.065)	(1.991)	(1.928)
English* I(year=2)	7.342**	1.378	10.89***	1.847
	(3.056)	(2.371)	(2.489)	(1.922)
Observations	438	438	938	938
Difference in English Coefficient across rounds	-0.553	-1.199	3.263	-2.209
	(4.587)	(3.817)	(3.195)	(2.716)
Panel B: First difference regression				
English* I(year=1)	10.01***	5.716*		
	(3.467)	(3.118)		
English* I(year=2)	8.918***	1.963		
	(3.409)	(3.071)		
Observations	219	219		
Difference in English Coefficient across rounds	-1.090	-3.752		
	(3.944)	(4.766)		

Note 1: additional regressors included in the regressions: children characteristics (gender, years in U.S., age at immigration), parent characteristics (total years of education, any U.S. education, below average English), and dummy for second period.

Note2: Robust standard error in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Change in Difference in Means by language of test: foreign-born children

Panel A: Balanced Panel Sample							
	First Round			Second Round			
	English	Spanish	Difference	English	Spanish	Difference	Change in Difference
	(c1)	(c2)	(c3)	(c4)	(c5)	(c6)	(c7)
			(c1)-(c2)			(c4)-(c5)	(c6)-(c3)
Reading	85.83	111.40	-25.57***	91.50	100.08	-8.581**	16.99**
	[27.13]	[23.47]	(5.586)	[12.51]	[20.14]	(3.786)	(6.749)
Math	89.20	102.65	-13.45*	92.76	94.49	-1.726	11.73*
	[28.98]	[18.50]	(5.299)	[14.49]	[21.19]	(3.927)	(6.755)
N	40	43	83	46	37	83	166
Panel B: Unbalanced Panel Sample							
Reading	86.53	107.08	-20.55***	94.48	98.06	-3.582	16.99***
	[28.22]	[24.58]	(3.379)	[14.91]	[24.62]	(3.697)	(5.154)
Math	93.03	100.97	-7.935*	97.28	97.86	-0.586	7.325
	[28.79]	[18.73]	(3.096)	[18.12]	[20.52]	(3.593)	(4.776)
N	121	124	245	65	51	116	361

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard deviations are in brackets and standard errors are in parenthesis.

Table 7: Regression estimates: foreign-born children

	OLS		FD	
	Reading	Mathematics	Reading	Mathematics
Panel A: OLS				
English* I(year=1)	-26.64***	-15.98***	-21.86***	-8.677***
	(5.395)	(5.003)	(3.166)	(2.825)
English* I(year=2)	-7.084*	0.308	-2.008	1.572
	(3.862)	(4.372)	(4.123)	(3.508)
Observations	166	166	361	361
Difference in English Coefficient across rounds	19.56***	16.29**	19.86***	10.25**
	(6.577)	(6.771)	(5.209)	(4.550)
Panel B: First difference regression				
English* I(year=1)	-33.23***	-18.59***		
	(6.025)	(6.459)		
English* I(year=2)	-13.67**	-0.500		
	(5.888)	(6.643)		
Observations	83	83		
Difference in English Coefficient across rounds	19.56***	18.09*		
	(6.957)	(10.25)		

Note 1: additional regressors included in the regressions: children characteristics (gender, years in U.S., age at immigration), parent characteristics (total years of education, any U.S. education, below average English), and dummy for second period.

Note2: Robust standard error in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 8: FD regression: balanced panel of foreign-born children

	Age at immigration <=8		Age at immigration>8	
	# years in U.S. at the time of first round<4	# years in U.S. at the time of first round>=4	# years in U.S. at the time of first round<4	# years in U.S. at the time of first round>=4
English* I(year=1)	-35.40***	-3.230	-60.44***	-
	(9.581)	(8.939)	(8.323)	-
English* I(year=2)	-8.536	2.709	-21.23**	-
	(9.826)	(9.331)	(8.813)	-
Observations	18	36	29	0
Difference in English Coefficient across rounds	26.87*	5.939	39.20***	-
	(13.25)	(7.701)	(13.19)	-

Figure1: Unconditional Quantile Regression estimates of language effect on test scores (U.S. born balanced panel sample)

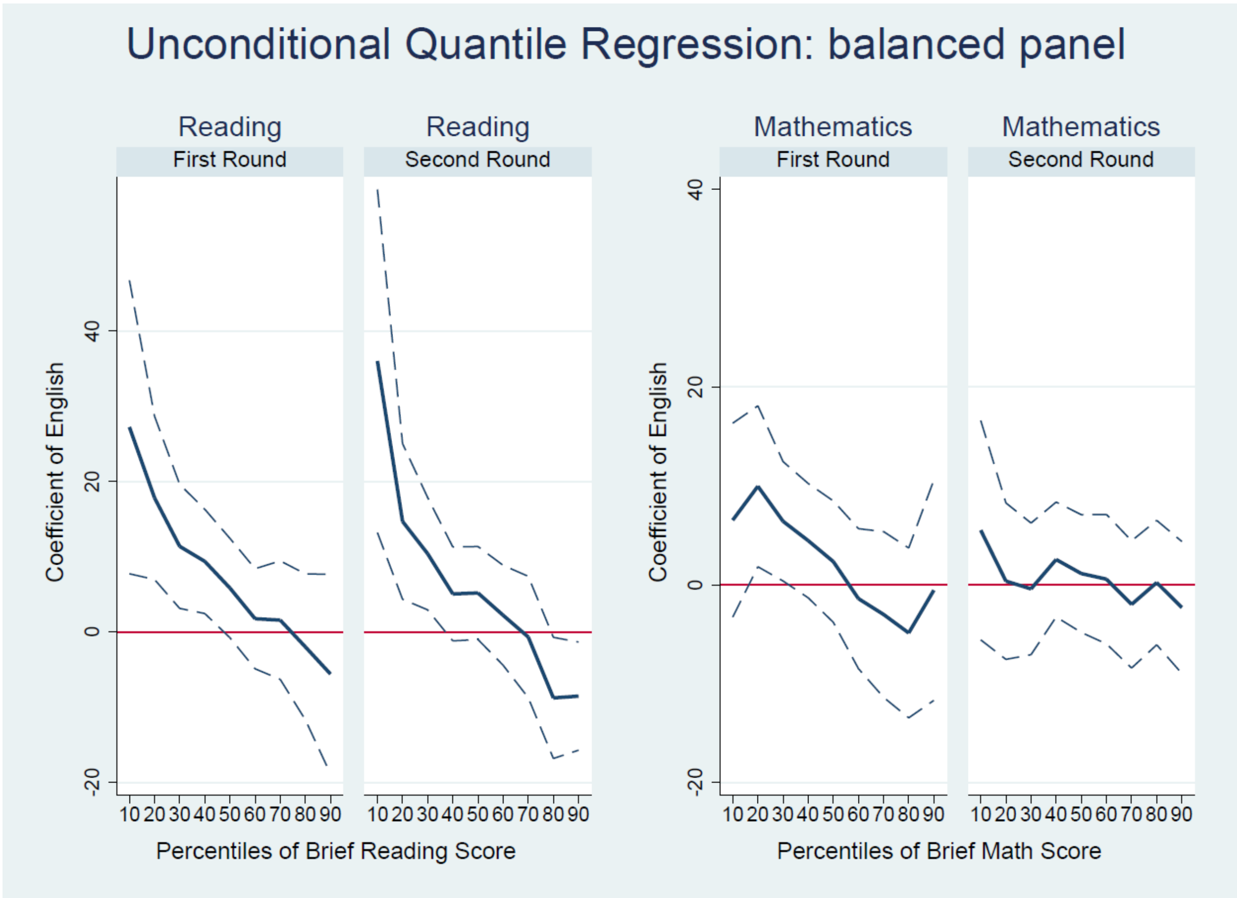


Figure 2: Age and language effect: Balanced Panel

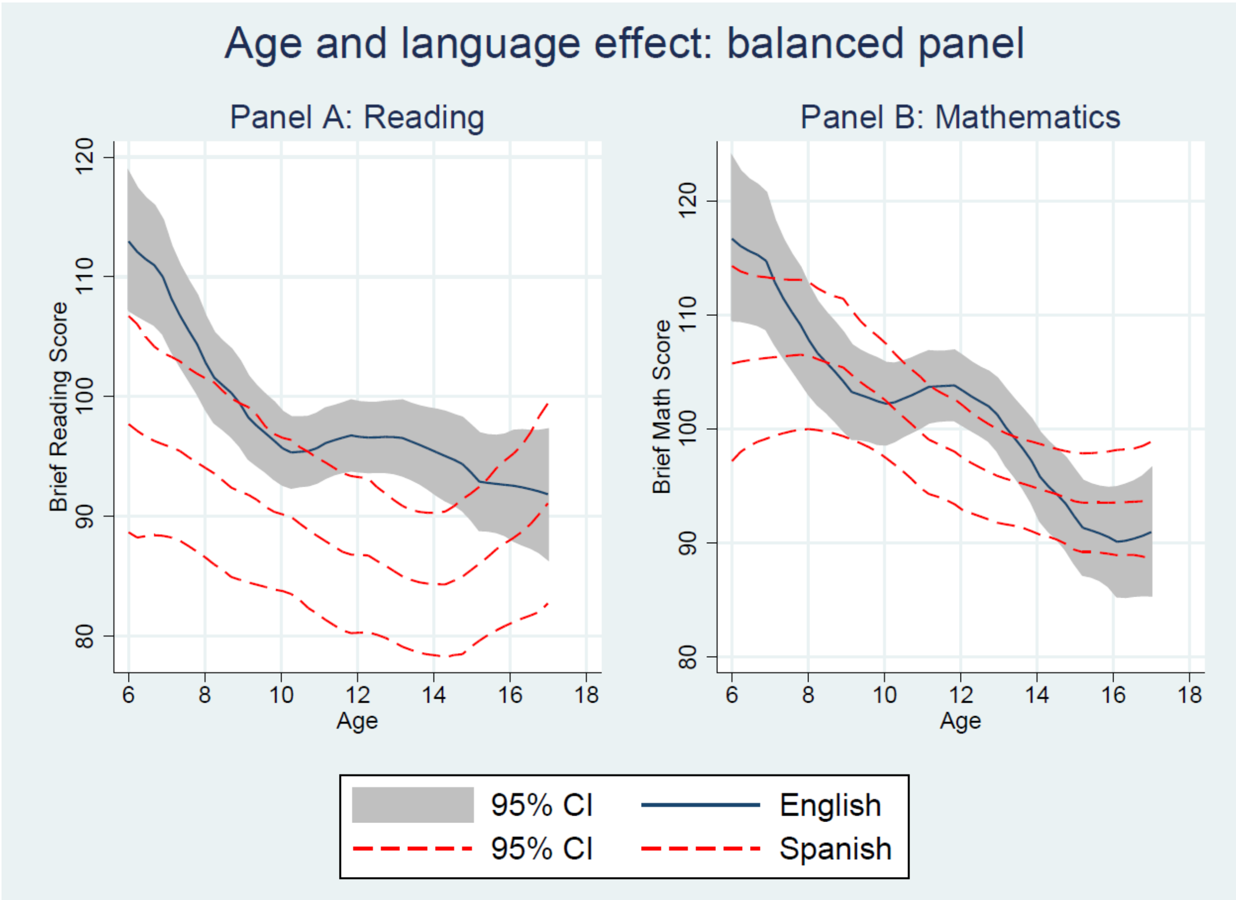


Figure 3: Unconditional Quantile Regression estimates of language effect on test scores (foreign-born balanced panel sample)

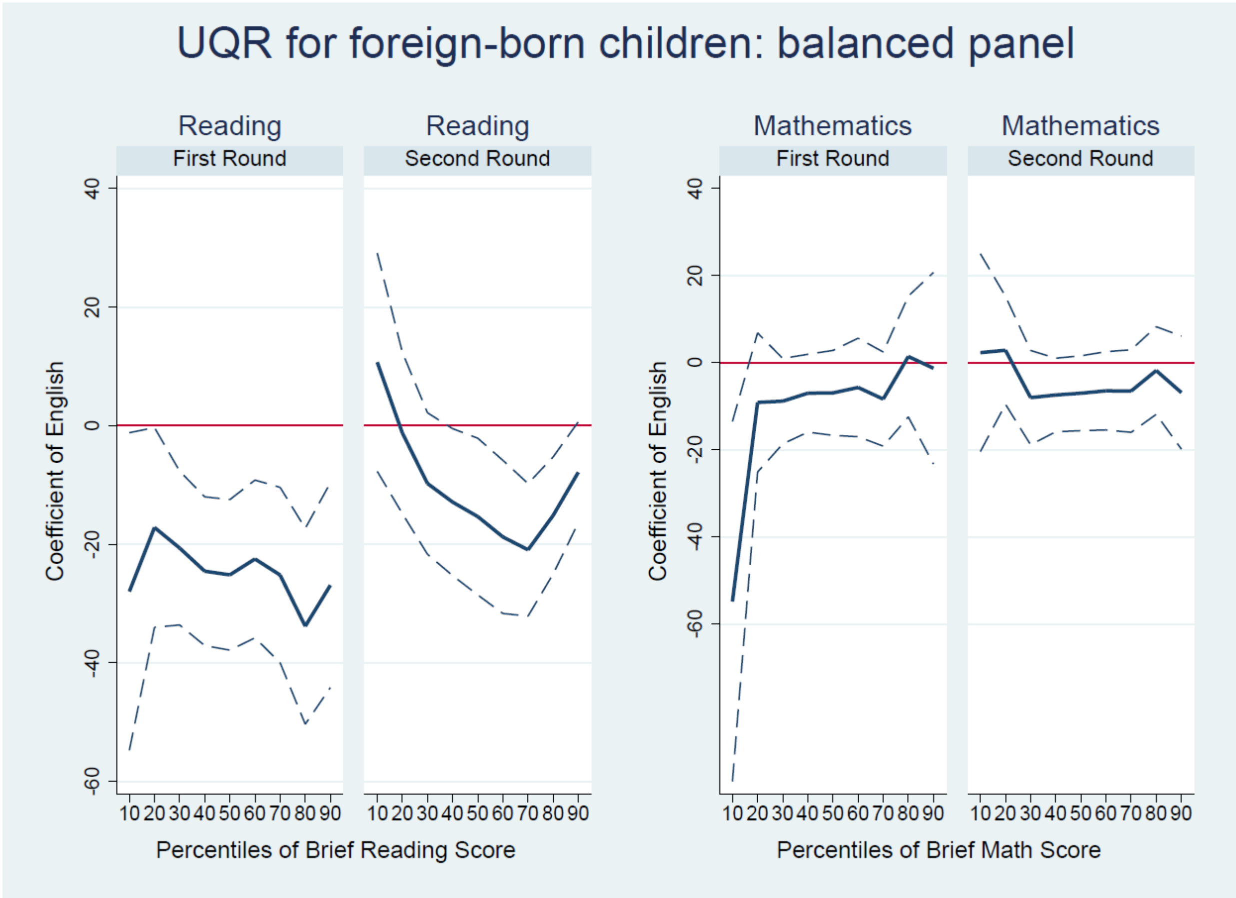


Figure 4: Age at immigration and Brief Reading (BR) score of foreign-born children: Balanced Panel

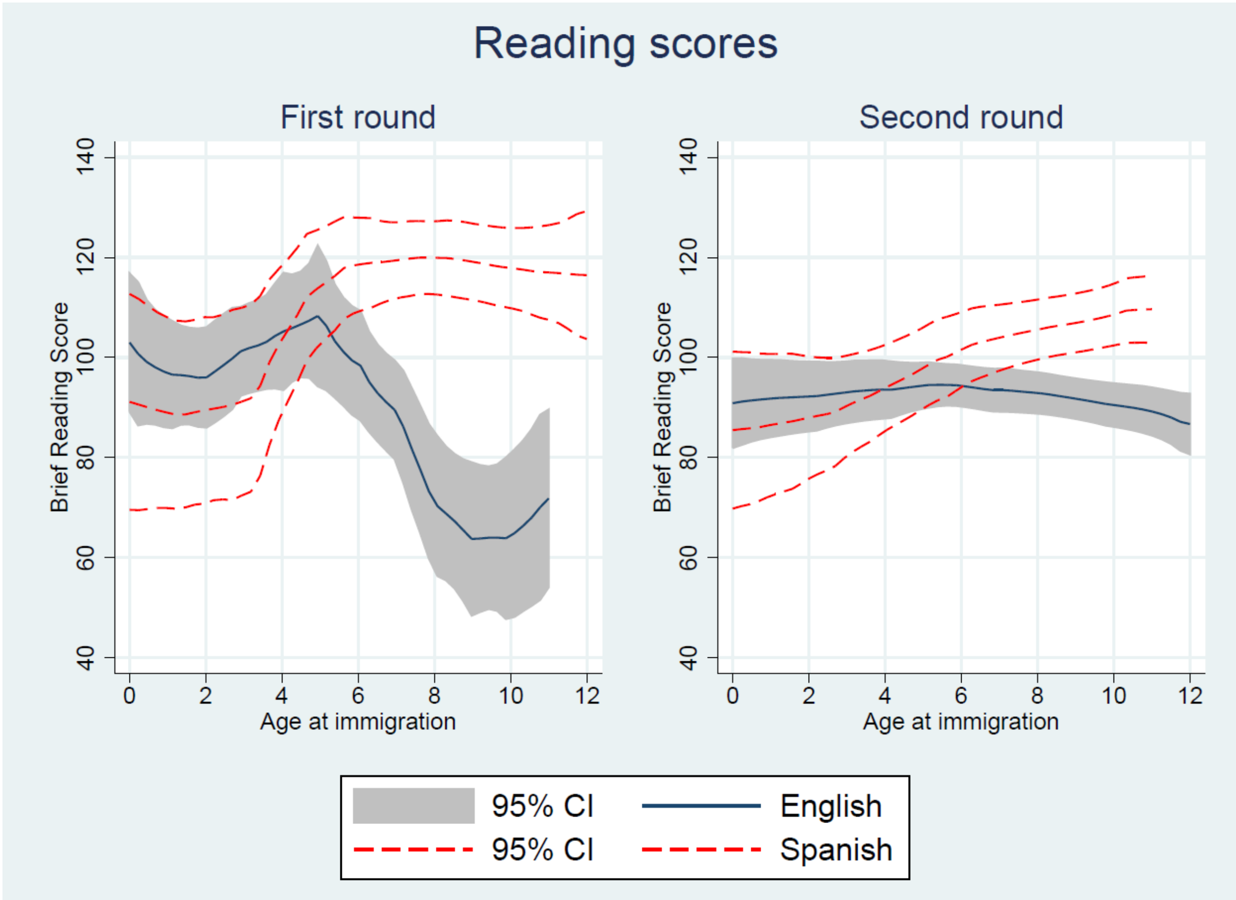


Figure 5: Age at immigration and Brief Math (BM) score of foreign-born children: Balanced Panel

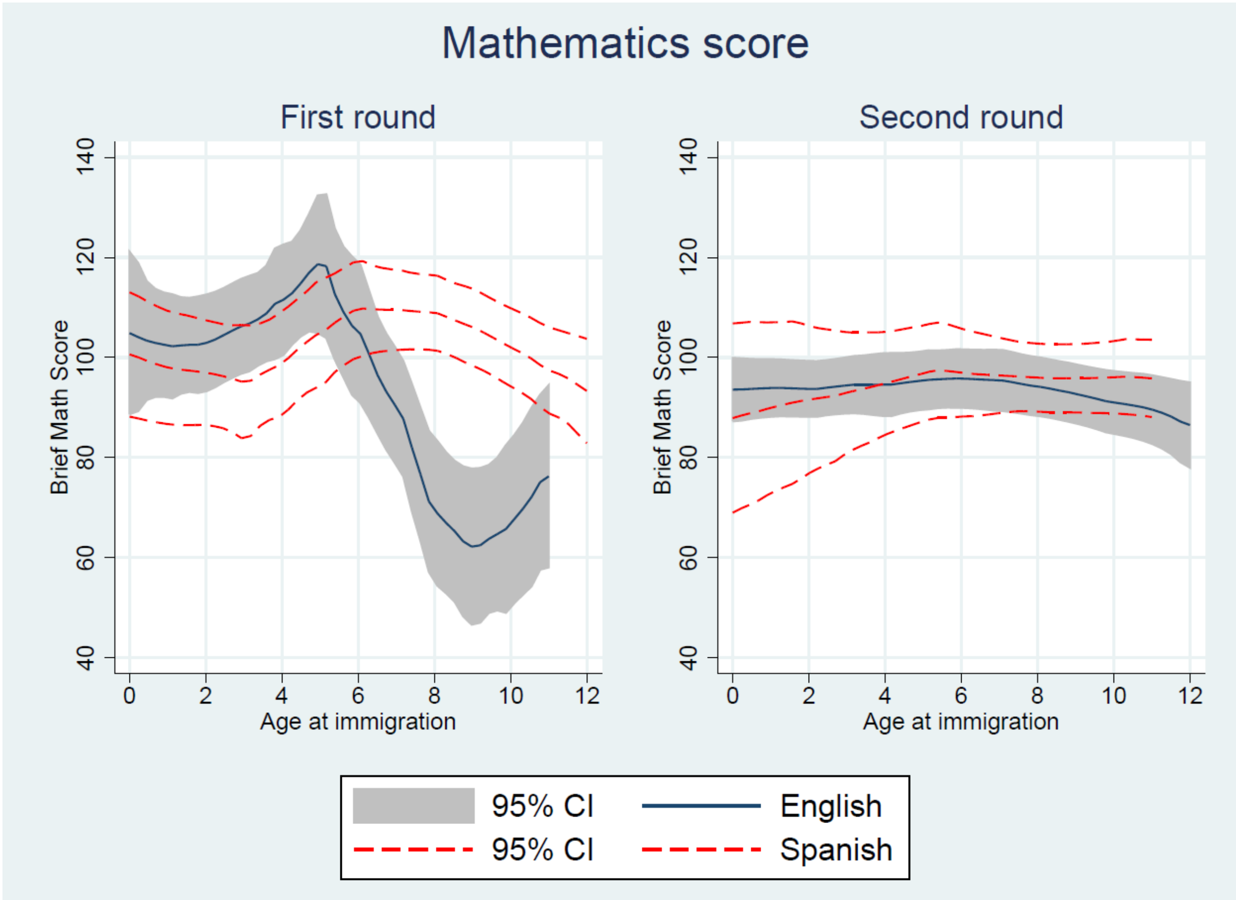


Figure 6: Reading score and years spent in the U.S.: foreign-born unbalanced panel

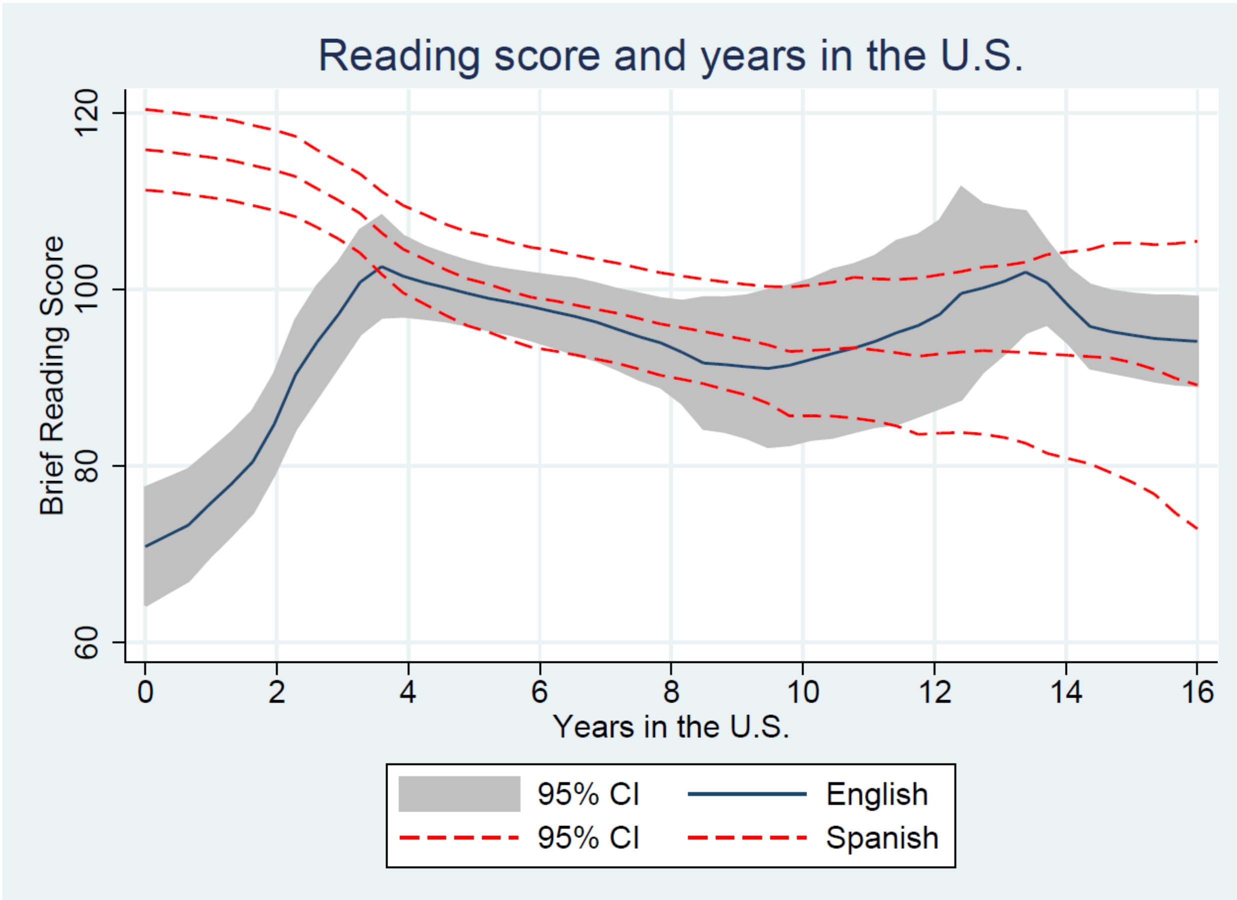


Figure 7: Reading score and years spent in the U.S. among foreign-born unbalanced panel: by age at immigration

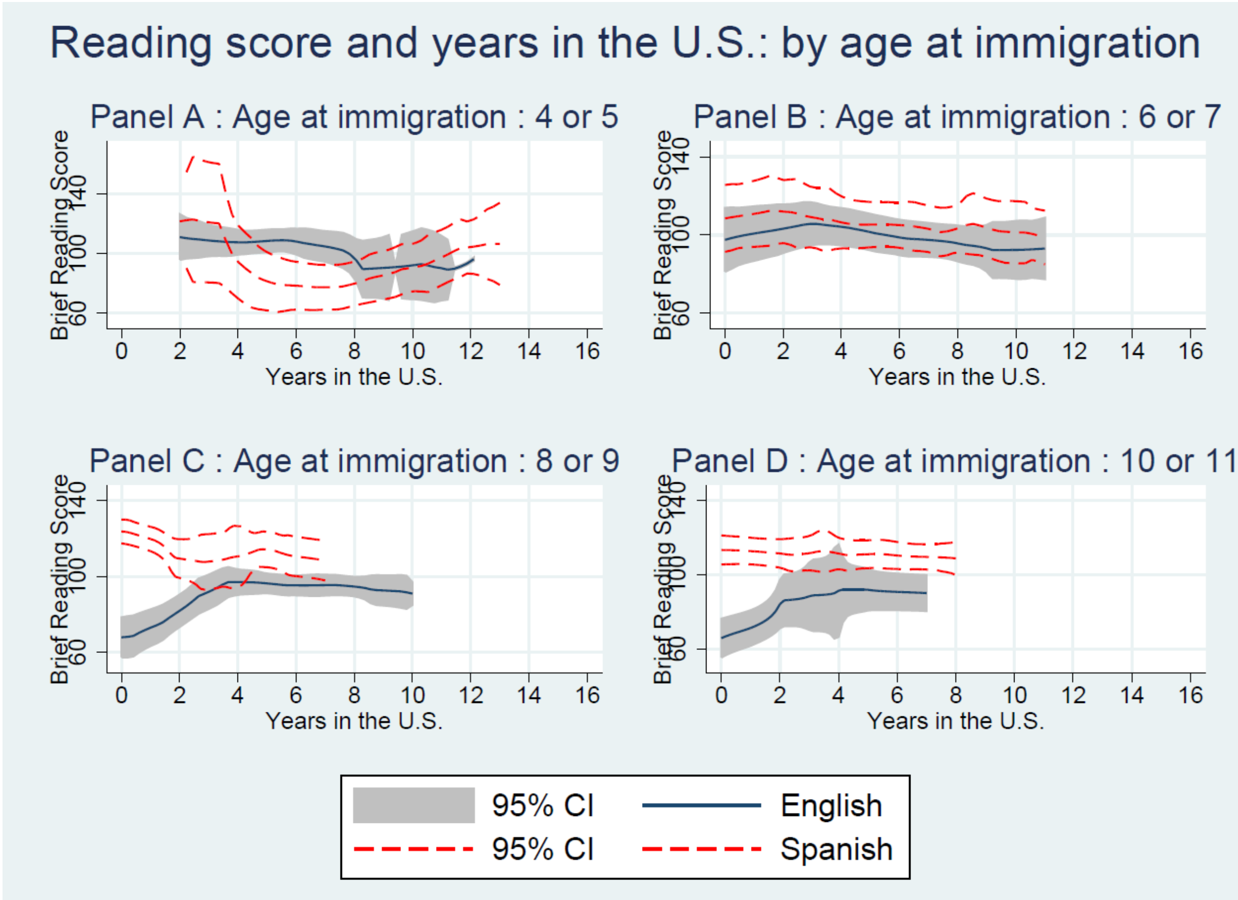
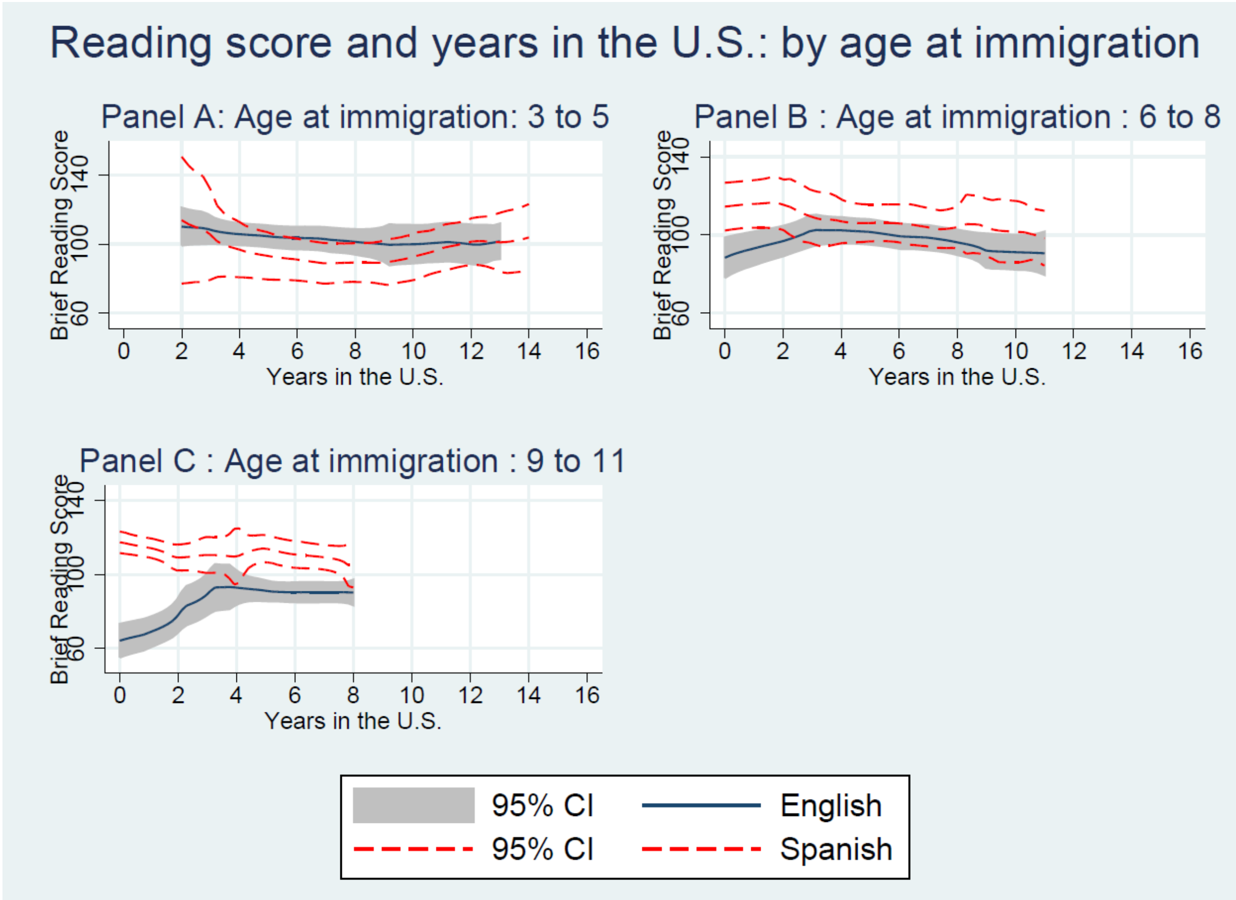


Figure 8: Reading score and years spent in the U.S. among foreign-born unbalanced panel: by age at immigration



Appendix

Figure A1: Unconditional Quantile Regression estimates of language effect on test scores of U.S. born children (unbalanced panel sample)

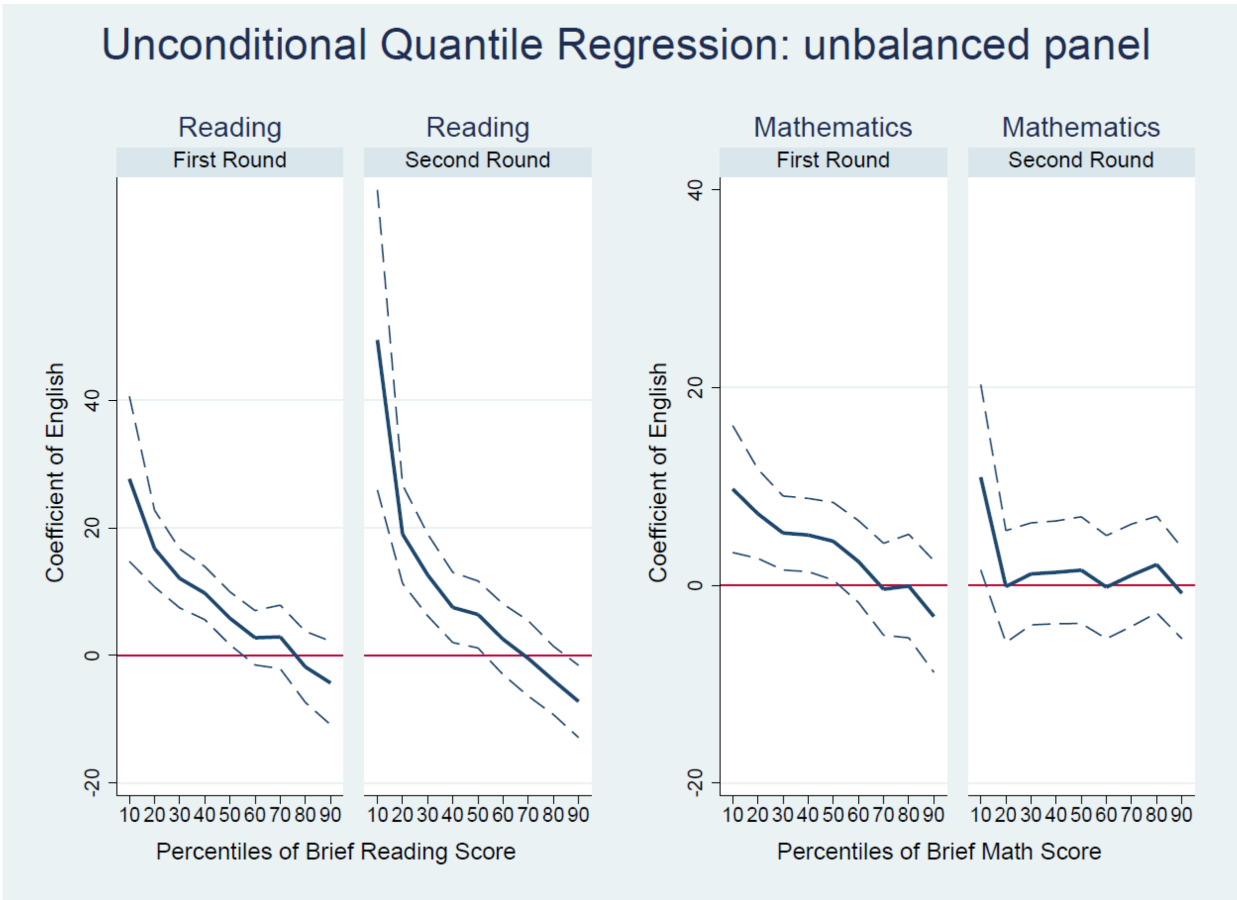


Figure A2: Age and language effect: unbalanced Panel

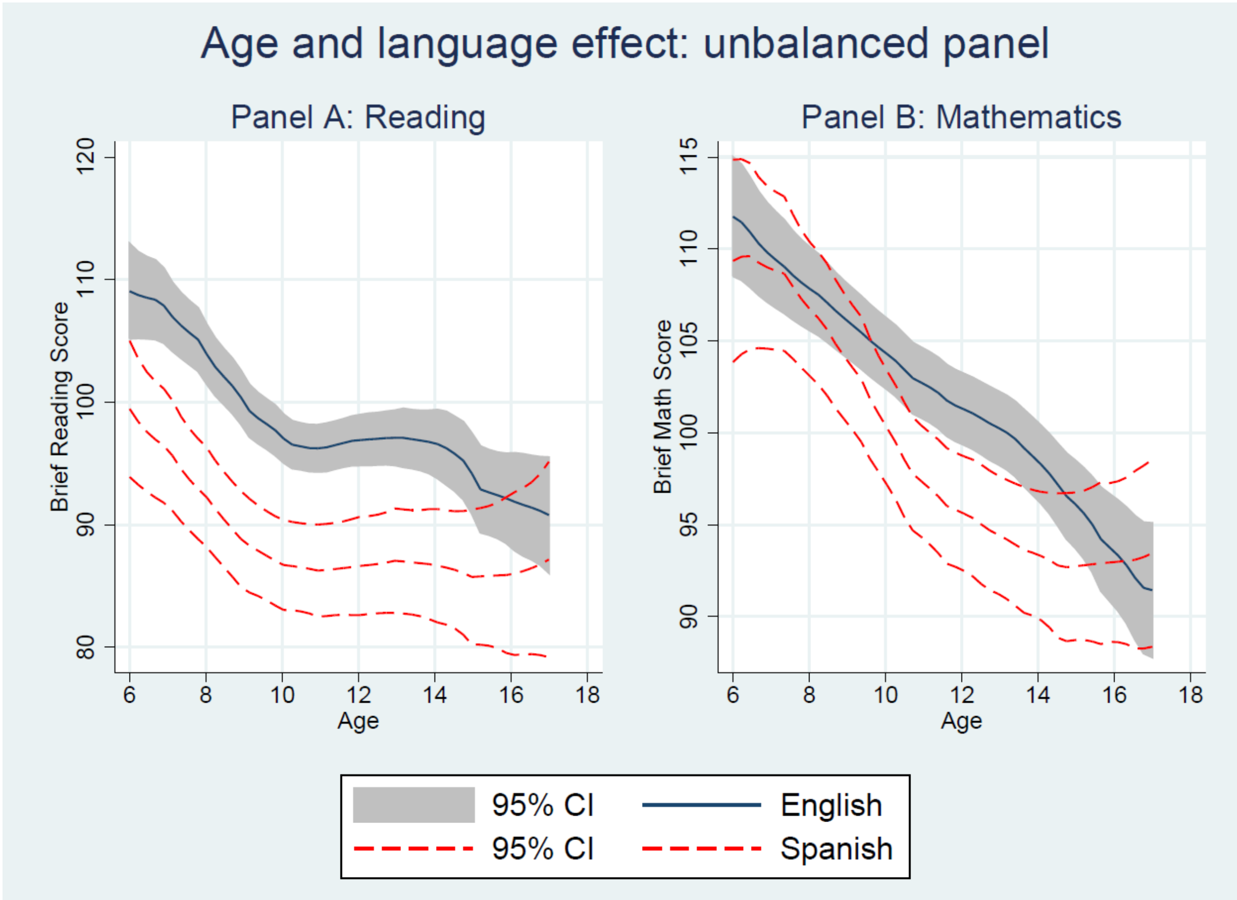


Figure A3: Unconditional Quantile Regression estimates of language effect on test scores of foreign-born children (unbalanced panel sample)

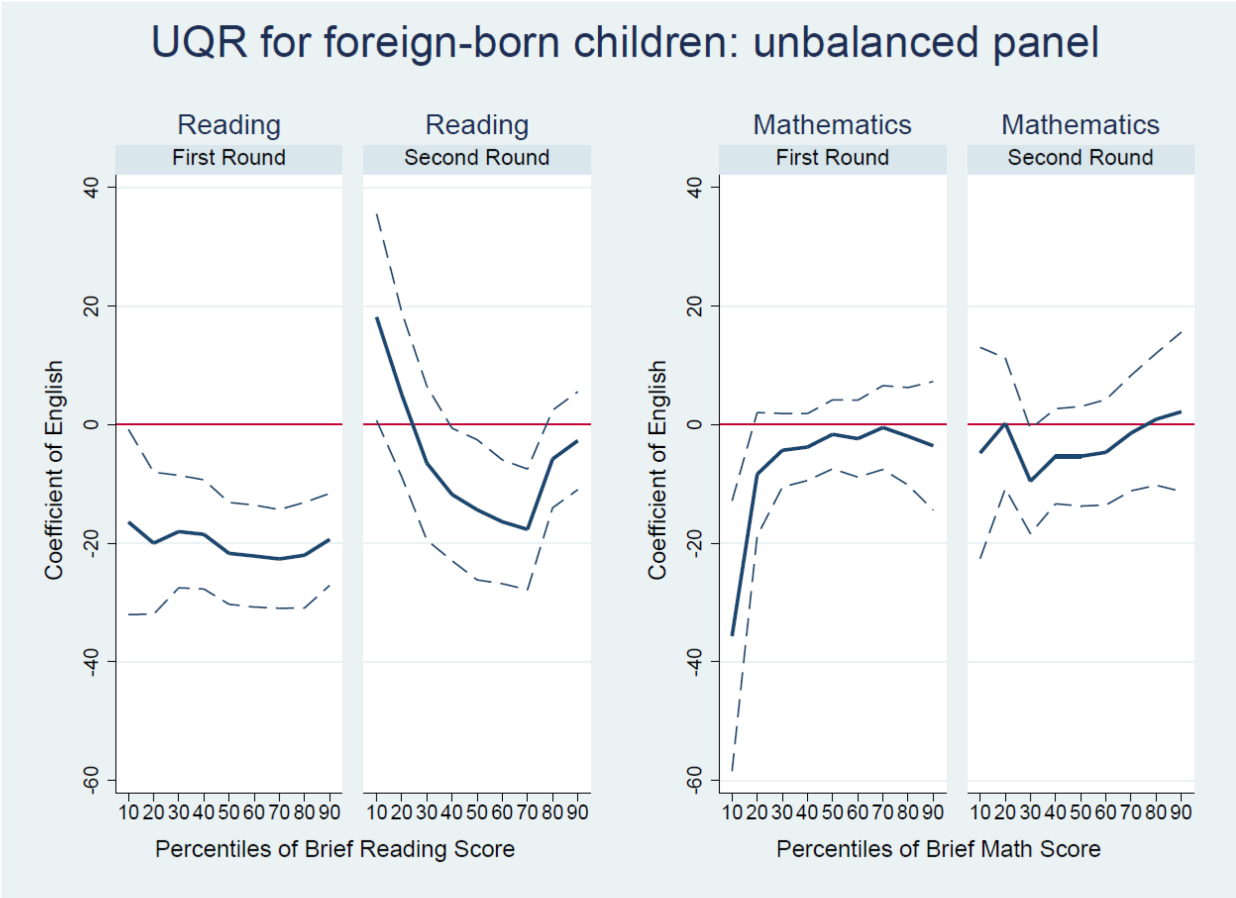


Figure A4: Age at immigration and Brief Reading (BR) score of foreign-born children: unbalanced Panel

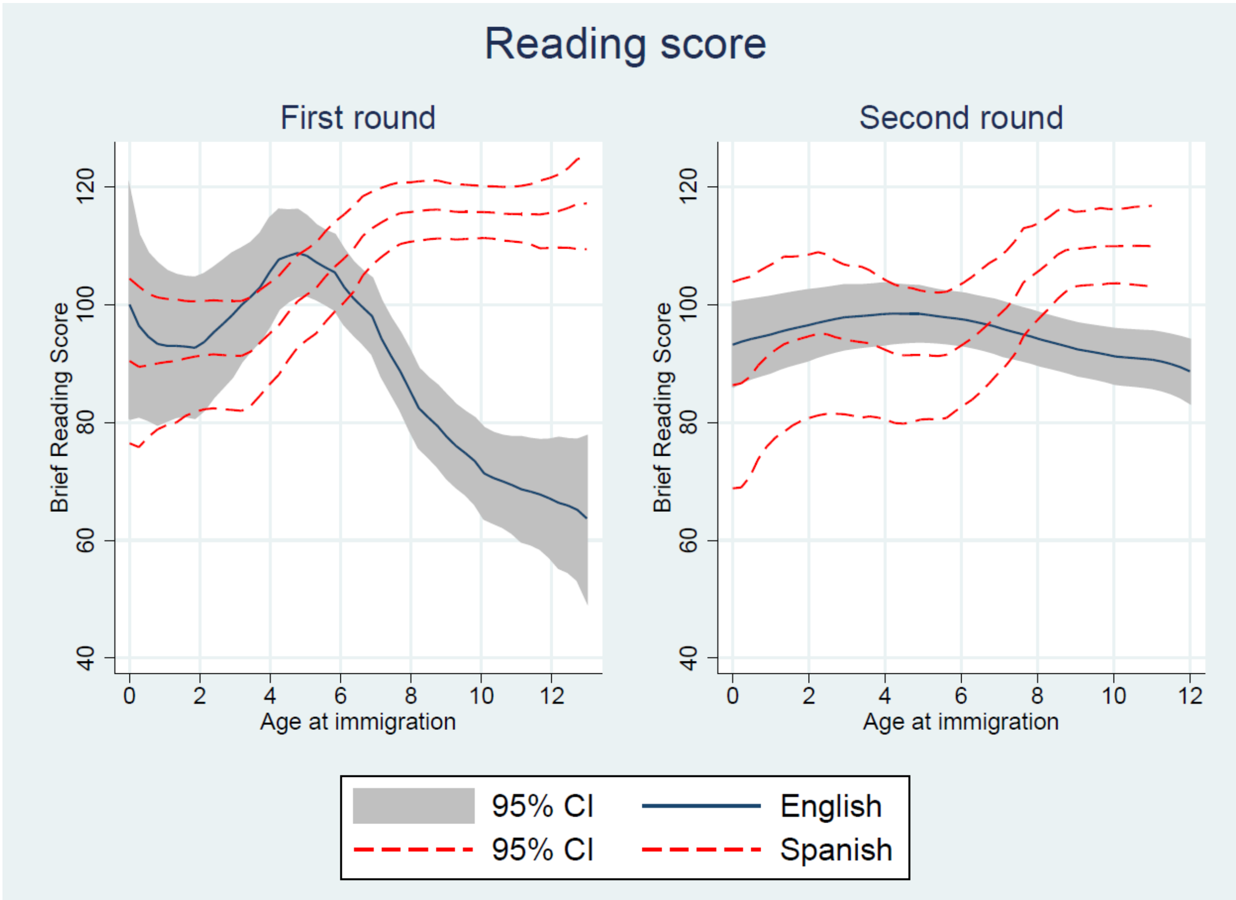


Figure A5: Age at immigration and Brief Mathematics (BM) score of foreign-born children: unbalanced Panel