THE UNIVERSITY OF CHICAGO

A Longitudinal Analysis of the Impact of Low Early Lingua-Franca Proficiency on Reading and Math Achievement Among Linguistically Minoritized K–5 Students in the US

By

Xinran Chen

August 2025

A paper submitted in partial fulfillment of the requirements for the
Master of Arts degree in the
Master of Arts Program in the Social Sciences

Faculty Advisor: Catherine C. Baumann, Ph.D.
Preceptor: Marshall Jean, Ph.D.

Abstract

While English is the United States' lingua franca, a growing proportion of US households regularly speaks a non-English language at home. Children from these households, however, often face academic disadvantages. Building on the existing literature, this thesis examines how the K–5 academic achievement of these linguistically minoritized (LM) children relates to their English proficiency when they enter kindergarten.

The dataset is sourced from the Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011). The analysis employs two-level hierarchical quadratic growth modeling on a selected subsample of LM children ($n = 5{,}021$). The outcomes are K–5 assessment scores in reading and math, and the main predictor is English proficiency at kindergarten entry. Drawing on cumulative disadvantage theory, the analysis adjusts for time-invariant covariates: gender, race/ethnicity, home language usage, baseline household socioeconomic status (SES), household structure, and school district impoverishment.

Regression results show an achievement gap in both reading and math between LM children from similar backgrounds who entered kindergarten with lower versus higher proficiency in English. Yet, due to the faster academic growth of those with initially lower proficiency, the gaps in the autumn of kindergarten (1.8 years of typical learning in reading and 0.93 years in math) narrowed to around 0.5 years for both subjects by the spring of Grade 5. These results remained qualitatively consistent after covariate adjustment. Notably, household SES exerted substantial influence in the opposite direction of early English proficiency.

These findings highlight the long-term impact of early English proficiency in the academic development of LM children and point to the crucial role of continual academic and linguistic supports that are suited to LM children's multilingual repertoires.

Introduction

In the United States, many different languages are spoken, and this linguistic diversity has been growing. The proportion of the US population aged five and above who speak a non-English language at home has steadily risen from 13.8% in 1990 (US Census Bureau, 1992) to 17.9% in 2000 (US Census Bureau, 2003), 20.6% in 2010 (US Census Bureau, n.d. a), and lastly 22.5% in 2023 (US Census Bureau, n.d. b). Surprisingly, this rise of non-English language usage has taken place despite the seemingly irresistible draw of English, which typically results in second-generation heritage speakers and third-generation English monolingualism.

However, students who speak a non-English language at home may be academically disadvantaged due to their language, alongside other demographic factors such as socioeconomic status (SES), household structure, and country of origin (Schmid, 2001). All these disadvantages may compound onto each other, and the impact could further deepen if the student commands a lower level of proficiency in the language of instruction, which is by default English (as the mainstream lingua franca, or LF). Adopting the framework of cumulative disadvantage (Nurius et al., 2015), students who speak a non-English language regularly at home are hereafter referred to as *linguistically minoritized* (LM; cf. Wang et al., 2021), a term intended as an agentive alternative to *language minority* (e.g., in Kieffer, 2011; Schmid, 2001). Note that LM in this thesis is defined in reference not to the English language per se, but rather to the status of English as the mainstream lingua franca (LF) in the US.

There are two related benefits to using "LM" over popular terms such as EL or ELL (i.e., English [Language] Learner).[1] First, as it is used in this thesis, LM does not presume any particular level of proficiency in the LF (here, English), which helps maintain conceptual clarity

[1] The EL or ELL terms will nonetheless be used sparingly below, when discussing previous research in which the authors did define such terms based on measures of English proficiency.

in discussing students' linguistic repertoires. Second, the LM grouping distinguishes between different pathways to a low proficiency in the LF and associated academic outcomes, particularly between homes in which LF exposure begins early and where it does not.[2] This distinction is important because the kinds of academic support that are appropriate for students similar in apparent achievement levels could differ between these linguistic home environments. In practice, LM students have been found to be disproportionately represented in lower academic tracks (Callahan, 2005) and in special education (Schmid, 2001), which could reflect an institutional conflation between linguistic proficiency and cognitive ability.

Broadly speaking, LM students of various ages and education levels can be set back by misalignment between the linguistic repertoire that the curriculum is designed for and what language skills the students currently possess. LM students may suffer extra cognitive burdens as they receive, process, and produce nonnative language while also processing the content knowledge contained therein (Hu and Lei, 2014), and such burdens can become more salient as students contend with increasingly complex materials (Ünsal et al., 2018). Not only could this make it more difficult for LM students to learn the content, but what they produce during exams administered in the nonnative language may underrepresent their true level of understanding. Whichever is the case, studies suggest a generally negative association between home-vs-school language misalignment and academic achievement for such varied student populations as "Standard English learners" in the US (Okoye-Johnson, 2011)[3], young NCLB test-takers

[2] Developmentally typical children can acquire any language as their native language if they take in reasonable amounts of linguistic input early on, but the acquisition process takes several years of time. Specialized language support can be seen as making up for language-acquisition experiences that LM children might not have had, so this would reasonably take up much time as well. Note that there is discussion further below, based on past studies, regarding the long process of acquiring a suitable level of English for both quotidian and academic usages.

[3] Dialectal variations in English in the US may readily evoke the familiar debate of what constitutes a language divide rather than a dialectal divide. Such debate is often based upon political identity and ideology rather than empirical evidence, e.g., measurement of mutual intelligibility. This thesis treats children speaking nonstandard English as LM, due to the similar practical challenges involved.

moderately proficient in English (Wright and Li, 2008), university students enrolled in a trilingual program in Italy (Bernhofer and Tonin, 2022), and ethnic-minority elementary school students in China (Yang et al., 2015).

In order to effectively assist with LM students' education, given the many aspects of their overall lived experience, it is imperative that researchers amass abundant information regarding what factors (linguistic and non-linguistic) may exacerbate or attenuate the known association between home-vs-school language misalignment and academic achievement. This is the key motivation for the study presented in the current thesis.

Literature Review

Globally, nondominant populations are often under pressure to assimilate. According to Bommes (2012), successful integration into modern society necessitates not only "structural assimilation"—i.e., some degree of conformity with the demands of structural positions such as job openings and association memberships—but also "cognitive assimilation", which includes the acquisition of "languages, skills, behavioural and situational patterns, normative knowledge, orientations towards mobility", etc. (p. 114). Of these aspects, it is noteworthy that a reasonable level of proficiency in a language locally used for academic purposes can theoretically facilitate the attainment of the rest. Given the fact that the modern quest for socioeconomic mobility in the US depends heavily on contextual factors (see Zhou's [1997] discussion of immigrant children's possible paths), a higher level of proficiency in English could potentially bring additional opportunities: social connections, economic wherewithal, or even political power.

Indeed, since the 1960s, one of the most consistent and least controversial policy goals in the K-12 education of LM students in the US has been to boost their proficiency in English, though not necessarily maintain proficiency in the non-English language(s) (Cohen, 1984; Orfield, 1986; Wright and Ricento, 2017). Despite strong evidence that pedagogy that includes LM students' home language is beneficial for their literacy and math development (Yu, 2024), the shortage of qualified bilingual teachers (see Gibney et al., 2021) and limited public support have meant that many LM students end up in an arrangement where they study content knowledge in LF-conducted (English) classrooms while also attending ESL or bilingual instruction. Whichever is the case, LM children who get educated in the US would typically eventually attain a level of English proficiency that allows them to complete many quotidian and academic tasks.

However, the timing of English (LF) acquisition could influence the developmental trajectory of LF proficiency, which may in turn impact the efficacy of LM students' learning. According to Hakuta and colleagues (2000), LM students in the US and Canada with limited English proficiency typically spent 3–5 years to attain an adequate level in oral English, and 4–7 years to attain the same in academic English. Due to the way in which developmentally typical children acquire languages (more by input than by direct instruction) as well as the ubiquity of English (as LF) in the US, proficiency attainment and academic achievement may come naturally if a reasonable level of English is acquired early on. For instance, a recent longitudinal study (Winsler et al., 2023) found that among lower-income LM students in Miami, students with early (before Grade 2) proficiency in English are associated with higher GPAs as well as higher reading and math achievement in Grade 5, and they are less likely to be retained in Grades 3–5. These effects were consistent whether or not covariates were adjusted for, and each year of delay in English acquisition was associated with incrementally less desirable academic outcomes for the lower-income LM students.

In addition to these static milestone results, studies have shown more nuanced differences in the trajectories of academic achievement between LM students who gain English proficiency earlier versus later, particularly in regard to the growth rate of the scores. For example, Kieffer's (2011) study suggests that after adjusting for SES and school concentration of poverty, LM students who were already fluent in English by kindergarten entry (i.e., passing the *pre*LAS[4] English language assessment) had higher English reading scores in kindergarten but lower overall growth rate through K–8 than LM students who gained fluency later.[5] The combined effect was an initial reading achievement gap between these two LM groups that shrank over time. These results are corroborated and expanded by three other studies that examined both English reading and math achievement, discussed below.

In one study, Halle and colleagues (2012) examined how a variety of cognitive and behavioral outcomes for a national sample of K–8 students related to LM status and the timeline of acquiring English proficiency. They divided the subsample of LM students into three groups based on how they performed on an English screener adapted from the same *pre*LAS assessment: those who passed the screener at kindergarten entry (n = 1,290), those who did not initially pass but were able to pass by the spring of Grade 1 (n = 980), and those who had not passed by the spring of Grade 1 (n = 410).[6] After adjusting for demographics and other contextually salient covariates, they found that LM students in the first group (i.e., those who acquired English proficiency the earliest, or the "early proficiency group") had the highest English reading and math scores in kindergarten among all three LM groups, but the rate of growth depended on both

---

[4] See Data Recognition Corporation (n.d.) for a description of the *pre*LAS assessment.
[5] Results from this study (Kieffer, 2011) further suggest that the LM students with lower early English proficiency eventually reached the achievement levels of comparable non-LM peers by Grade 8, while LM students with fluent early English proficiency even surpassed their non-LM peers in English reading achievement as early as Grade 3. In other words, the overall growth rates of both these LM groups were higher than the non-LM group.
[6] It seems likely that Halle et al. (2012) deliberately rounded these figures to the tens place, since the numbers do not sum to their reported total LM sample size of n = 2,670.

the timeline of proficiency acquisition and the academic subject assessed. Compared to the LM students in the early-proficiency group, those in the middle group (i.e., passed the screener sometime during Grades K–1) had a lower growth rate in English reading and math, while those in the late-proficiency group (i.e., not yet passed the screener by the spring of Grade 1) had a much higher growth rate in English reading and similar growth rate in math. Halle and colleagues (2012) thus inferred the following about the trajectories of the three groups of LM students' reading and math achievement: in reading, the initial achievement gap between the early- and middle-proficiency groups widened while that between the early- and late-proficiency groups shrank; whereas in math, the initial achievement gap between the early- and middle-proficiency groups also widened but the gap between the early- and late-proficiency groups persisted.

In another study that highlighted summer loss, Johnson (2023) traced the English reading and math achievement trajectories of a group of US public school students in Grades K–4, both in the academic year and over the summer. Within the group of LM students who were classified as ELs (i.e., low English proficiency) at any point during these grades (the "ever-ELs"; n = 8,206), the subgroup of students who stayed in EL status throughout these years (the "always-ELs"; n = 4,088) had lower kindergarten reading scores than the group average and lower reading growth rates in Grades K–1, but higher reading growth rates during the academic year in Grades 2–4. Meanwhile, in terms of math achievement, the always-ELs also had lower kindergarten math scores than the ever-EL group average, but the subgroup generally had slightly higher growth rates during the academic year throughout Grades K–4. Yet the gap between the always-EL subgroup and the whole group's average widened in English reading and persisted in math during K–4; this was attributed to significant loss during the summer, where the always-EL students experienced a more substantial score loss compared to the ever-EL group average, particularly over the summers after Grades 2 and 3 respectively (pp. 1059–1060).

A third study focused on a later timeframe for the attainment of English proficiency. In Pilger Suhr and colleagues' (2021) study on EL reclassification between Grades 3–5, they found that the earlier an LM student was reclassified as English-proficient, the higher their Grade 3 reading and math scores, the lower their reading growth rate, but the higher their growth rate in math achievement. The reading gaps narrowed while the math gaps widened between the LM groups with different timelines for English acquisition.

Collectively, the abovementioned four studies suggest that, when contextual factors are held constant, an LM student's earlier acquisition of English proficiency is generally associated with an academic advantage in both English reading and math that remains at least through the elementary grades, even though LM students who acquire English relatively late experience a higher overall growth rate in English reading.

Despite the discussion on overall growth rates, it is necessary to keep in mind that, as the abovementioned research shows, academic achievement growth curves can be nonlinear (e.g., Kieffer, 2011; Pilger Suhr et al., 2021). Both the piecewise modeling of IRT theta scores[7] in Kieffer's (2011) study (also cf. Johnson's [2023] piecewise approach) and the logarithmic modeling of Rasch unit scale scores in the Pilger Suhr et al. (2021) analysis highlight the nonlinearity. This may partially explain the surprising results that Halle et al. (2012) found from their multinomial linear modeling, and a nonlinear approach might have enabled a more nuanced understanding of their data.

[7] IRT = Item Response Theory. Theta scores are a measure of the student's latent ability in a certain academic subject, to be contrasted with other measures such as items-correct and quantiles.

Another point worthy of consideration is generalizability across US contexts. While local studies of specific classrooms (e.g., Haneda, 2009), specific schools (e.g., Sánchez et al., 2022), or specific school districts (e.g., Alvear, 2015) are helpful, it is uncertain to what extent the findings could be generalized to other US locales with different policies, practices, demographics, etc. A responsible study of the LM student population in the US as a whole would thus require nationally representative sampling.

Finally, while cross-sectional studies of LM students have been conducted (e.g., Abedi et al., 2006; Kieffer & Thompson, 2018), it is uncertain how to investigate the longer-term effects of interventions and contextual or structural factors, as part of the lived experience of LM students. As was the case for many of the studies reviewed above, a longitudinal data-collection design would more easily support such analysis.

Research questions

This thesis intends to examine how early English (as LF) proficiency relates to academic achievement in the longer term among LM students in the US. In accordance with the research reviewed above, academic achievement is operationalized as reading and math scores. Similarly, "early" here is defined as acquiring English proficiency by kindergarten entry, and the grades of interest are Grades K–5. The two research questions are as follows:

> (1) For linguistically minoritized (LM) students in the US, to what extent does early proficiency in the LF predict the growth trajectories of reading and math achievement from kindergarten to Grade 5?
>
> (2) To what degree does this predictive association remain after adjusting for child-level demographic and contextual covariates that might confound the association, such as gender, race/ethnicity, frequency of non-English language usage, baseline household SES, single-parent household status, and school-district poverty?

The second research question, in particular, attempts to separate the unique contribution of language proficiency, controlling for other potential predictors of the developmental trajectories of reading and math abilities. If low early English proficiency predicts a negative impact on LM students' reading and/or math learning even net of covariates, then this could justify programs that facilitate early LF (English) acquisition for LM students, and/or content-knowledge support that takes their linguistic needs into consideration. Theoretically speaking, this study will also contribute to a better understanding of how childhood reading and math development relates to personal and social factors that shape the lived experiences of LM children who are raised and educated in the US.

Data and methods

Data

This thesis analyzes data sourced from the Early Childhood Longitudinal Study – Kindergarten Class of 2010-11 (ECLS-K:2011; see Tourangeau et al., 2019). The dataset used for analysis is merged between original public-use data and preprocessed data. The preprocessing was conducted by the EDSO Archive Team at the University of Chicago.

The original longitudinal study by NCES (see Tourangeau et al., 2019) systematically selected a nationally representative sample of US kindergartners in the 2010-11 academic year (initial $n = 18{,}174$) and followed them through Spring 2016. Since a majority of students were supposed to be in 5th grade by then, the dataset provides K-5 data on most of the initially sampled students, despite attrition (resulting in $n = 11{,}445$ in Spring 2016). The data was collected in a total of 9 waves. The publicly accessible variables include (1) the children's academic, socioemotional, and other achievement measures, (2) family situations, (3) classroom and school contexts, and (4) demographic and other information about the children themselves and the individuals and institutions (including classrooms and schools) most relevant to a child's development. Many of these were measured across several waves, though some were measured only once.

Meanwhile, the preprocessed data comes from the EDSO Data Archive. Certain notable variables from the original dataset were selected, cleaned, and prepared by the EDSO Archive Team. The data was organized into several files, including time-level and child-level data files suitable for hierarchical analysis.

See Table 1 for an overview of the waves of data collection for ECLS-K:2011.

Table 1. The waves of ECLS-K:2011 at a glance[8]

Wave/TIME	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Total *n* =	15,756	17,207	5,230	15,132	4,735	13,850	12,896	12,102	11,445
Data collection	2010 Fall	2011 Spring	2011 Fall	2012 Spring	2012 Fall	2013 Spring	2014 Spring	2015 Spring	2016 Spring
Modal grade	**K**	**K**	**1**	**1**	**2**	**2**	**3**	**4**	**5**

Total unique children *n* = 18,174.

Language-related variables

Only a selection of variables collected from the respondents of the NCES study were available for analysis. Besides the EDSO Archive Team's own selection process for the preprocessed data, there is a deeper layer of redaction that NCES conducted in order to protect the privacy of the respondents. While the restricted version of the dataset does include many un-redacted variables, that version was not accessible for the current study. What *could* be accessed

[8] Part of the information in this table is sourced from the EDSO Archive Team's Analytical Codebook. Sample sizes are those that remain after the Archive Team's cleaning process.

was the public-use version of ECLS-K:2011, on which the EDSO Archive's preprocessed version was also based.

The LM subset was selected by examining the children's parents' responses to the following two questions. Data is available for Waves 1, 2, 4, 6, 7, and 9, although in Wave 7 the dataset only records responses to the first question.

(1) *Is any language other than English regularly spoken in your home?*

Recorded responses to Question 1 were Yes, No, Not applicable, Refused, Don't know, Not ascertained, and Missing (no value). "Not applicable" indicates that previous questions or the study methodology precludes an answer to this question, while missing data indicates nonresponse.

(2) *What is the primary language spoken in your home?*

Recorded responses to Question 2 were English, Spanish, Asian language, Other language, Not applicable, Refused, Don't know, Not ascertained, and Missing (no value). The interpretations for "Not applicable" and missing data remain the same.

This thesis operationalizes an LM student as someone whose respondent parent reported regularly speaking a non-English language at home (i.e., an answer of Yes to Question 1) in any wave. This forms the basis for selecting LM student data from the dataset as a whole.

Further, in order to account for the frequency of non-English language usage at home (i.e., Question 2), a separate binary variable PRIMARY was coded. The PRIMARY variable takes 1 if the respondent parent reported Spanish, Asian language, or Other language as the primary language spoken in the home, in any wave; and 0 otherwise. The set of children whose PRIMARY value is 1 was found to be a proper subset of the set of LM children derived above.

In order to determine early English proficiency at kindergarten entry, data-flag variables were examined that indicate how the children were routed through the academic assessments. A variable LOPROF was coded which takes value 1 if there is enough information to classify the child as having low English proficiency anytime in the first four waves (i.e., Grades K–1), and 0 if the child had high early proficiency, i.e., was able to pass an English proficiency screener by kindergarten entry. Note that the routing information was not collected by NCES beyond Wave 4 because by then, 99.9% of the students were reportedly so proficient in English that they were routed into assessments delivered entirely in English (Tourangeau et al., 2019, Ch. 2, p. 4).

More specifically, NCES employed an assessment routing process that depended upon both the child's home language and their performance in a *pre*LAS-based English language screener. Each child took the screener and a basic English reading test, and then they entered one of three possible assessment routes: Full in English, full in Spanish (except Science), or nothing. Below is the procedure that NCES used; note that when routing, they sourced the home language information from schools rather than directly from parents. Note that the English screener tests English at a relatively low level.

(1) Home language is English: >> Reading, Math, Executive function[9]; Science (Spring)
(2) Home language is Spanish:
    a. Passed screener: >> same assessments as English home-language speakers
    b. Failed screener: >> same assessments, but in Spanish and without Science
(3) Home language is another language:
    a. Passed screener: >> same assessments as English home-language speakers
    b. Failed screener: >> no further cognitive assessment

For each wave, the data flags were recorded as one of the following values.

[9] Executive function data is not analyzed in this thesis because the focus is on academic achievement as it is more traditionally defined.

The data flags:

- 0 = English speaker: (i) non-LM (case 1 above) or (ii) LM but passed the English screener (case 2a or 3a).
- 1 = Spanish speaker (case 2b).
- 2 = speaker of another language (case 3b).
- "-9" = "not ascertained", i.e., the assessment was coded as invalid.
- Missing data = nonresponse

Distribution of these values is given below in Table 2.

Table 2. The four waves of screener- and home-language-based flags in detail

Wave	**1 (Fall K)**	**2 (Spring K)**	**3 (Fall G1*)**	**4 (Spring G1)**
0 (English)	15,378	17,037	5,167	15,064
1 (Spanish)	344	157	39	21
2 ("other language")	63	21	0	0
Valid total	*15,785*	*17,215*	*5,206*	*15,085*
-9 (not ascertained)	6	0	24	47
Missing data	2,383	959	12,944	3,042
EDSO Archive sample total [10]	*15,756*	*17,207*	*5,230*	*15,132*
Total	*18,174*	*18,174*	*18,174*	*18,174*

*G1 = Grade 1.

A student was coded with a LOPROF value of 1 if in any of the first four waves their flag was 1 or 2; or a value of 0 if the flag was 0 wherever it was available (i.e., not missing or -9). If a student had either missing data or -9 for all of these flags, the student's data was discarded as part of the data cleaning process (discussed in more detail below), since there was not enough information to determine the student's early proficiency in English.

Data processing

As a first step, some of the entries on both the time-level and the child-level had to be removed due to data unavailability. On the time level, any entries were eliminated that had

[10] From the Analytical Codebook compiled by the EDSO Archive Team. Some students' entries were cleaned out from the EDSO Archive data at some point, but the discrepancy is small.

missing data for the age at assessment or had missing data for either the reading or math IRT theta score. On the child level, any entries were eliminated where any of the following data was not available: early English proficiency (see above), gender, race/ethnicity, baseline SES, baseline single-parent household status,[11] or the aggregate of school district poverty levels over the child's K–5 experience. As a general rule, the cleaned version of the time-variant (time-level) data corresponds with the time-invariant (child-level) data, and deleted Child IDs apply across both levels.

After selecting only the data associated with LM children, the sample for analysis consists of n = 33,532 child-wave entries on Level 1 and n = 5,021 children on Level 2. In order to facilitate calculation and interpretation, the continuous SES and school-district poverty measures were then both standardized to a mean of 0 and a standard deviation of 1.

Some descriptive statistics after the standardization are given below in Table 3.

Table 3. Child-level descriptive statistics for the selected LM sample (n = 5,021)

<table>
<tr><td colspan="2">Gender</td><td colspan="4" rowspan="3">Contextual variables</td></tr>
<tr><td>Female</td><td>2,506 (49.9%)</td></tr>
<tr><td>Male</td><td>2,515 (50.1%)</td></tr>
<tr><td colspan="2">Race/ethnicity[12]</td><td colspan="4">Continuous (standardized)</td></tr>
<tr><td>Hispanic</td><td>3,069 (61.1%)</td><td>Variable</td><td>Mean</td><td>SD</td><td>Range</td></tr>
<tr><td>White</td><td>483 (9.6%)</td><td>Baseline SES</td><td>0.00</td><td>1.00</td><td>-2.36, 3.61</td></tr>
<tr><td>Black</td><td>210 (4.2%)</td><td rowspan="3">District poverty (aggregated K–5)</td><td rowspan="3">0.00</td><td rowspan="3">1.00</td><td rowspan="3">-1.88, 2.94</td></tr>
<tr><td>Asian</td><td>1,030 (20.5%)</td></tr>
<tr><td>Other[13]</td><td>229 (4.6%)</td></tr>
<tr><td colspan="2">Language-related variables</td><td colspan="4">Categorical binary</td></tr>
<tr><td>PRIMARY = 1</td><td>2812 (56.0%)</td><td colspan="2" rowspan="2">Baseline single-parent household</td><td colspan="2" rowspan="2">1,017 (20.3%)</td></tr>
<tr><td>LOPROF = 1</td><td>394 (7.8%)</td></tr>
</table>

Note: PRIMARY = 1 means that a non-English language is used as the primary language at home. LOPROF = 1 means that the child had lower English proficiency at kindergarten entry, as indicated by the results from English screeners administered in Grades K–1.

[11] SES and single-parent household measures are drawn from the first available datapoint among the waves, and they are therefore time-invariant baseline measures on the child level.
[12] Non-Hispanic unless otherwise indicated.
[13] The Other category encompasses all the remaining racial/ethnic groups that were less represented.

All variables at a glance

Time level (Level 1)

1. The reading outcome used in this thesis is the IRT theta score for reading. This is calibrated through Grades K–5 and is a representation of the child's latent ability in reading comprehension, as opposed to items-correct on any particular test. This is a continuous variable, and in the analytical sample it ranges from -4.29 to 2.21 across all children and waves.

2. The math outcome used in this thesis is the IRT theta score for mathematics. This is also a continuous variable, and in the analytical sample it ranges from -3.78 to 2.68 across all children and waves.

3. The time variable used in this thesis is the age at which the children are assessed. Whereas the original study measured these ages in months, for the benefit of modeling the months have been converted into the equivalent number of years with decimals.

Child level (Level 2)

1. LOPROF: Binary, taking 1 for low early English proficiency, and 0 otherwise.

2. PRIMARY: Binary, taking 1 if primarily speaking non-English language at home.

3. FEMALE: Binary, taking 1 for female, and 0 for male.

4. Race/ethnicity dummy variables.

5. SES: Continuous measure of household socioeconomic status at baseline. This variable has been standardized for modeling purposes.

6. SINGPAR: Binary, taking 1 if single-parent household at baseline; 0 if not.

7. DISTPOV: Continuous measure of aggregated school district poverty across all the schools in which the child was enrolled in the K-5 period. This variable has also been standardized for modeling purposes.

Modeling

Hierarchical linear modeling (HLM) was employed for longitudinal analysis, due to three main benefits. First, and fundamentally, each child's prior academic achievement often predicts their later achievement (e.g., Jimerson et al., 2013), which can be considered a form of clustering. A hierarchical modeling approach would help account for individual differences by modeling each student separately. Second, HLM handles certain kinds of missing time-level data better, by modeling trajectories with all available measurement occasions. This is important because not every child is measured on all occasions. By using HLM methods, the missing measurements do not impact the modeling process in any major way. Finally, not only the intercept but also the linear slope and the quadratic coefficient can be explicitly a function of the predictor and covariates, which is key to answering the second research question.

As was discussed earlier, growth curves are often nonlinear, and so this thesis uses quadratic modeling as a relatively straightforward approach to nonlinearity. The full quadratic model is specified as follows; note that the AGE variable is centered on 8 years of age, which is close to the grand mean age across all child-wave entries.

Level 1.

$$\text{THETA}_{ti} = \pi_{0i} + \pi_{1i}(\text{AGE}'_{ti}) + \pi_{2i}(\text{AGE}'_{ti})^2 + e_{ti}$$

t = Wave of data collection; i = child.

THETA = The IRT theta score in question: reading or math.

π_{0i} = Child i's score at age 8, in theta units.

π_{1i} = Child i's instantaneous growth rate at age 8. This is close to the average of the instantaneous growth rates and therefore is a good approximation of the child's overall growth rate over K-5.

$\pi_{2i} = \frac{1}{2}$ of Child i's score growth acceleration, assumed to be constant over K-5.

e_{ti} = Level 1 residual (Normal), i.e., the remaining unexplained variation in the scores. We assume that $e_{ti} \sim N(0, \sigma^2)$.

Level 2.

$$\pi_{0i} = \beta_{00} + \sum_{k=1}^{L} \beta_{0k} v_{ki} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \sum_{k=1}^{M} \beta_{1k} v_{ki} + r_{1i}$$

$$\pi_{2i} = \beta_{20} + \sum_{k=1}^{N} \beta_{2k} v_{ki} + r_{2i}$$

i = child.

v_{ki} are the child-level variables of interest and L, M, N are the number of such variables included in each of the equations. These numbers vary across the models used.

β_{00} = expected value of π_{0i}, i.e., score at age 8 assuming zero value for all child-level variables.[14]

β_{10} = expected value of π_{1i}, i.e., instantaneous growth rate at 8 or approximately the overall K–5 growth rate, if all the child-level variables take zero.

β_{20} = expected value of π_{2i}, i.e., one-half of growth acceleration, if all the child-level variables take zero.

For brevity, definitions for the effect βs have been omitted. These consist of the explanatory variables that are incorporated in each model.

r_{0i} = unique (random) effect of Child i on π_{0i} (i.e., score achievement at age 8) conditioning on all other variables.

r_{1i} = unique (random) effect of Child i on π_{1i} (i.e., instantaneous growth rate at age 8, or approximately the overall K–5 growth rate) conditioning on all other variables.

[14] Zeros for all child-level variables refers to a child who is male, of a race/ethnicity excluded by all the racial dummy variables (depending on the model, either a non-Hispanic child or a white child), whose primary home language is not a non-English language, whose early English proficiency is reasonably high, whose SES measure is average, whose school-district poverty measure is average, and who does not live in a single-parent household.

r_{2i} = unique (random) effect of Child i on π_{2i} (i.e., $\frac{1}{2}$ of growth acceleration over K-5) conditioning on all other variables.

These random effects are assumed to be in joint multivariate Normal distribution, i.e.,

$$\begin{pmatrix} r_{0i} \\ r_{1i} \\ r_{2i} \end{pmatrix} \sim N\left(\vec{\mathbf{0}}, \begin{pmatrix} \tau_{00} & \tau_{01} & \tau_{02} \\ \tau_{10} & \tau_{11} & \tau_{12} \\ \tau_{20} & \tau_{21} & \tau_{22} \end{pmatrix}\right).$$

Results and Discussion

Main findings

The main results of this thesis were obtained through a total of six quadratic models, consisting of three for reading (Models 1–3) and three for math (Models 4–6). In each case, the three quadratic models were formed, respectively, (1) with low early English proficiency as the only predictor, (2) adjusting for an additional Hispanic vs. non-Hispanic binary as well as other non-racial covariates, and (3) with the full set of variables, including every racial dummy contrasting with non-Hispanic white students. Hispanic ethnicity was specially considered in Models 2 and 5 both because of how the ECLS-K:2011 study treated racial/ethnic identity categories, and because it is often associated with proficiency in Spanish, which is a dominant minority language in the US. Indeed, as the descriptive statistics in Table 3 previously showed, students identified as Hispanic constitutes a majority of the LM subsample.

The results from the regression are presented below; due to formatting reasons, the first regression table (Table 4) begins on the next page.[15]

[15] *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, and (ns) = statistically nonsignificant.

Table 4. Regression results for Reading scores

	Model 1	SE	**Model 2**	SE	**Model 3**	SE
Intercept						
Intercept	0.86***	0.01	0.86***	0.01	0.82***	0.02
Low early English proficiency	-0.39***	0.02	-0.26***	0.02	-0.26***	0.02
Primary home language is non-English			0.01 (ns)	0.01	-0.001 (ns)	0.012
Female			0.09***	0.01	0.09***	0.01
Hispanic			-0.08***	0.01	-0.03 (ns)	0.02
Black					0.00 (ns)	0.03
Asian					0.10***	0.02
Other					0.02 (ns)	0.03
SES measure, standardized			0.13***	0.01	0.12***	0.01
Single parent household			-0.04*	0.02	-0.03*	0.02
District poverty aggregate, standardized			-0.01 (ns)	0.01	-0.01 (ns)	0.01
Linear						
Intercept	0.491***	0.002	0.497***	0.005	0.525***	0.007
Low early English proficiency	0.137***	0.009	0.102***	0.009	0.102***	0.009
Primary home language is non-English			-0.006 (ns)	0.004	0.001 (ns)	0.004
Female			-0.018***	0.004	-0.017***	0.004
Hispanic			0.013*	0.005	-0.019**	0.007
Black					-0.021 (ns)	0.012
Asian					-0.056***	0.008
Other					-0.011 (ns)	0.011
SES measure, standardized			-0.045***	0.002	-0.043***	0.002
Single parent household			0.011*	0.005	0.009 (ns)	0.005
District poverty aggregate, standardized			-0.004 (ns)	0.002	-0.005*	0.002
Quadratic						
Intercept	-0.114***	0.001	-0.113***	0.002	-0.120***	0.003
Low early English proficiency	-0.030***	0.003	-0.020***	0.003	-0.020***	0.003
Primary home language is non-English			0.003 (ns)	0.002	0.000 (ns)	0.002
Female			-0.000 (ns)	0.002	-0.000 (ns)	0.001
Hispanic			-0.005**	0.002	0.003 (ns)	0.003
Black					0.000 (ns)	0.004
Asian					0.015***	0.003
Other					0.004 (ns)	0.004
SES measure, standardized			0.015***	0.001	0.014***	0.001
Single parent household			-0.005*	0.002	-0.004*	0.002
District poverty aggregate, standardized			0.003**	0.001	0.003***	0.001

Table 5. Random effects for Reading scores

Variance components	Model 1	Model 2	Model 3
Intercept τ_{00}	0.142***	0.118***	0.117***
Linear τ_{11}	0.017***	0.015***	0.014***
Quadratic τ_{22}	0.001***	0.001***	0.001***
Time-level e	0.101	0.101	0.101
Deviance	35352.9	34229.8	34219.1

Table 6. Regression results for Math scores

	Model 4	SE	**Model 5**	SE	**Model 6**	SE
Intercept						
Intercept	0.95***	0.01	1.03***	0.02	0.95***	0.03
Low early English proficiency	-0.37***	0.03	-0.20***	0.03	-0.21***	0.03
Primary home language is non-English			0.07***	0.01	0.03*	0.02
Female			-0.05**	0.01	-0.05***	0.01
Hispanic			-0.16***	0.02	-0.06*	0.03
Black					-0.09*	0.04
Asian					0.22***	0.03
Other					0.02 (ns)	0.04
SES measure, standardized			0.17***	0.01	0.16***	0.01
Single parent household			-0.05**	0.02	-0.04*	0.02
District poverty aggregate, standardized			-0.02**	0.01	-0.02*	0.01
Linear						
Intercept	0.578***	0.002	0.586***	0.004	0.604***	0.005
Low early English proficiency	0.052***	0.006	0.035***	0.006	0.035***	0.006
Primary home language is non-English			0.003 (ns)	0.003	0.006 (ns)	0.003
Female			-0.020***	0.003	-0.020***	0.003
Hispanic			0.003 (ns)	0.004	-0.017**	0.006
Black					-0.029**	0.010
Asian					-0.029***	0.006
Other					-0.012 (ns)	0.009
SES measure, standardized			-0.021***	0.002	-0.021***	0.002
Single parent household			-0.002 (ns)	0.004	-0.003 (ns)	0.004
District poverty aggregate, standardized			-0.005**	0.002	-0.005**	0.002
Quadratic						
Intercept	-0.106***	0.001	-0.105***	0.001	-0.105***	0.002
Low early English proficiency	-0.008**	0.003	-0.003 (ns)	0.003	-0.003 (ns)	0.003
Primary home language is non-English			-0.002 (ns)	0.001	-0.002 (ns)	0.001
Female			0.004***	0.001	0.004***	0.001
Hispanic			-0.003*	0.001	-0.002 (ns)	0.002
Black					0.002 (ns)	0.003
Asian					0.001 (ns)	0.002
Other					-0.003 (ns)	0.003
SES measure, standardized			0.007***	0.001	0.007***	0.001
Single parent household			-0.003*	0.002	-0.003*	0.002
District poverty aggregate, standardized			0.004***	0.001	0.004***	0.001

Table 7. Random effects for Math scores

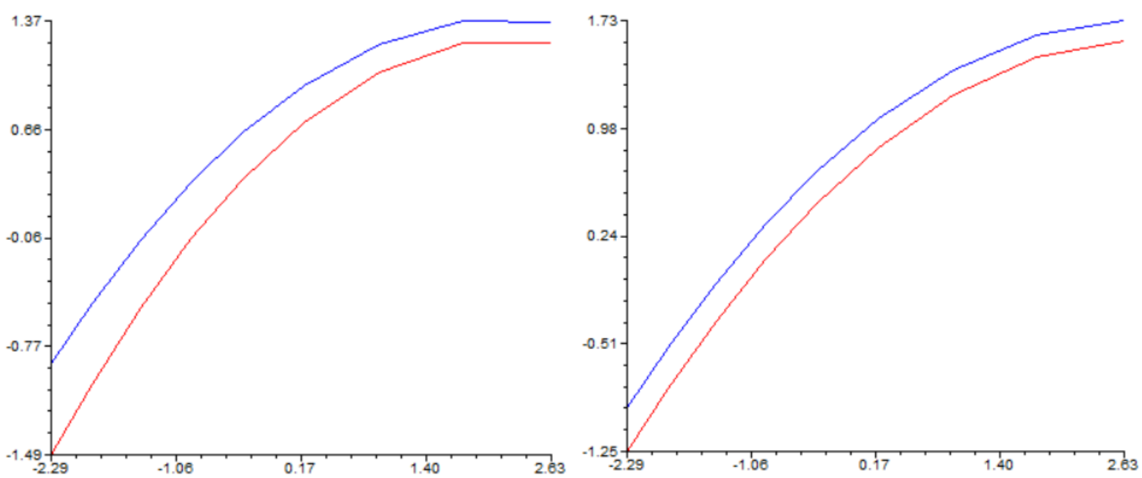
Variance components	Model 4	Model 5	Model 6
Intercept τ_{00}	0.259***	0.211***	0.206***
Linear τ_{11}	0.008***	0.008***	0.008***
Quadratic τ_{22}	0.001***	0.001***	0.001***
Time-level e	0.059	0.059	0.059
Deviance	23595.0	22405.0	22333.7

For reading achievement, low early English proficiency is associated with a lower reading score at age 8 across all the reading models (every $p < 0.001$) but higher instantaneous growth rate at that point (every $p < 0.001$) and more negative acceleration in the learning of reading (every $p < 0.001$). Since the instantaneous growth rate at age 8 is also approximately the overall growth rate in the entire K–5 period, these findings combined suggest a narrowing of the achievement gap between students whose early English proficiency was low and those whose proficiency was higher. Based on results from the unconditional Model 1, preliminary calculation shows that the average reading achievement gap between LM students with low vs. high early English proficiency is around 1.8 years of typical learning growth in reading in the autumn of kindergarten, which narrows to around 0.50 years of typical learning growth by the spring of Grade 5.

In regard to math achievement, low early English proficiency is similarly associated with a lower math score at age 8 across all the math models (every $p < 0.001$) and higher instantaneous growth rate at that age (every $p < 0.001$). However, unlike the reading growth trajectories, the moderately negative association between low early English proficiency and learning acceleration in math disappears after adjusting for covariates, suggesting that this parameter of math growth trajectories is similar between those with low vs. high early English proficiency (Model 4: $p < 0.01$; Models 5 and 6: $p > 0.05$, statistically nonsignificant). These findings, much like the ones above in reading achievement, also imply a narrowing gap between the low- and high-early-proficiency groups. Based on results from the unconditional Model 4, preliminary calculation shows that the average math achievement gap between LM students with low vs. high early English proficiency is around 0.93 years of typical learning growth in math in the autumn of kindergarten, which narrows to around 0.49 years of typical learning growth by the spring of Grade 5.

From both the numeric regression results above and Figure 1 below, it can be seen that the inclusion of covariates does not qualitatively change the associations between low early English proficiency and the reading or math growth trajectories (except for the acceleration term for math, which becomes statistically insignificant as noted). This suggests that the development of these achievement gaps deserve attention in their own right, besides other predictors of lower reading and math achievement and growth described further below.

Figure 1. Comparison between the average trajectories for LM students with low vs. high early English proficiency, in reading and in math.



On the left are reading trajectories, and on the right are math trajectories. The *x*-axes express age centered on 8, and the *y*-axes are the theta scores. In each graph, the blue (top) curve illustrates the average trajectory of those with higher early English proficiency, and the red (bottom) curve illustrates that of those with lower early English proficiency. To generate these graphs, Models 3 and 6 were used, which had the most covariates as well as the best model fit for reading and math respectively.

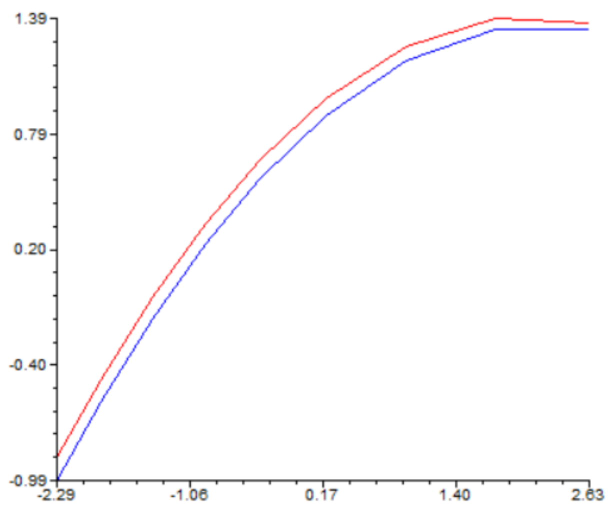
Covariates: Reading achievement

Speaking primarily a non-English language at home is not statistically significantly associated with any differences in the growth trajectory of reading. Notably, the Pearson correlation coefficient between low early English proficiency and primary usage of a non-

English language at home was around 0.14, indicating a very weak relationship between these characteristics. The regression results suggest that the early English proficiency level has more predictive power for reading growth than the frequency of non-English usage at home per se.

Female gender is associated with higher reading achievement at age 8 (every $p < 0.001$) and lower instantaneous growth rates (every $p < 0.001$). Meanwhile, there is no statistically significant relationship between female gender and the acceleration of growth in reading. From Figure 2 below we can observe that there is a gender achievement gap in reading that slightly narrows through K–5.

Figure 2. Gender differences in reading growth.

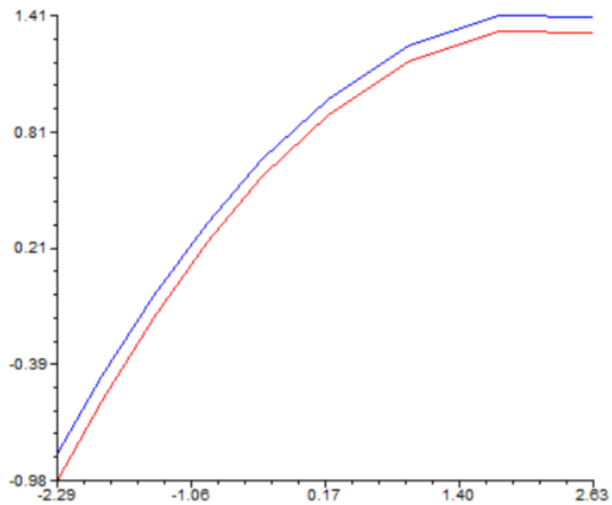


The *x* axis is age centered on 8, and the *y*-axis is the reading theta score. The red (top) curve is for female students, and the blue (bottom) curve is for male students.

Hispanic students are associated with lower reading achievement at age 8 when compared with non-Hispanic students (Model 2, $p < 0.001$), but not statistically significantly differentiated from white students (Model 3). Hispanic students are associated with higher instantaneous growth in reading compared with non-Hispanic students as a whole (Model 2, $p < 0.05$), but lower instantaneous growth compared with white students (Model 3, $p < 0.01$). There is also an association with somewhat more negative acceleration in reading growth compared with non-

Hispanic students (Model 2, $p < 0.01$), but not when compared with white students (Model 3). Figure 3 below (based on Model 2) demonstrates that a gap in reading achievement persists between Hispanic and non-Hispanic students, even when other relevant factors are controlled for.

Figure 3. Hispanic vs. non-Hispanic students' reading growth.



The x axis is age centered on 8, and the y-axis is the reading theta score. The blue (top) curve is for non-Hispanic students, and the red (bottom) curve is for Hispanic students.

In regard to other races/ethnicities, compared with white students of similar backgrounds, Asian students are associated with higher reading achievement at age 8 ($p < 0.001$), lower reading growth ($p < 0.001$), and less negative acceleration in reading growth ($p < 0.001$). Meanwhile, Black students and students in the Other racial/ethnic category are not associated with statistically significantly different trajectories compared with white students of similar backgrounds. This may be either a matter of statistical power due to the smaller sample sizes of these groups, or a matter of intra-group heterogeneity.

Higher household SES is associated with higher reading achievement at age 8 (every $p < 0.001$), lower reading growth (every $p < 0.001$), and less negative acceleration of reading growth (every $p < 0.001$). The contribution of household SES is substantial: the entire range of SES data spans more than 5 SD units, and a 2 SD difference can approximately cancel out the intercept and growth-rate (slope) effects of low early English proficiency for reading achievement.

Living in a single-parent household at baseline is associated with somewhat lower reading achievement at age 8 (every $p < 0.05$) and possibly marginally higher reading growth (Model 2: $p < 0.05$; Model 3: $p > 0.05$, nonsignificant). It is also associated with a somewhat more negative acceleration in reading growth (every $p < 0.05$). Although the magnitudes of all these coefficients are all rather small, the point holds that living in a single-parent household may be associated with long-term impact on reading. Note that household structure and SES may interact, which is a point that calls for more research.

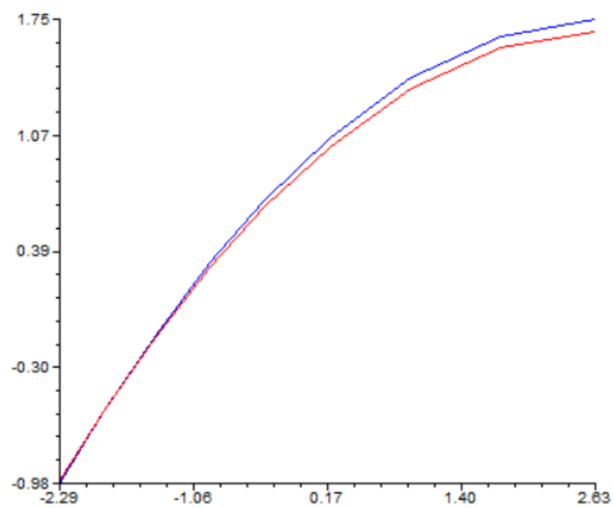
Finally, enrollment in relatively impoverished school districts is associated with a possibly lower reading growth rate (Model 2: nonsignificant; Model 3: $p < 0.05$), but most definitely a less negative acceleration in reading growth (Model 2: $p < 0.01$; Model 3: $p < 0.001$). There is no statistically significant association between aggregate school district poverty and the reading achievement at age 8. These results suggest that with all other factors controlled for, the effect of school district poverty on reading achievement may be cumulative and may only surface later in the LM students' academic career.

Covariates: Math achievement

Speaking primarily a non-English language at home is associated with higher math achievement at age 8 (Model 5: $p < 0.001$; Model 6: $p < 0.05$), though otherwise the math growth trajectories of such students are not statistically significantly differentiated from those of others, in terms of growth rate or acceleration in growth. Importantly, these results are obtained while other covariates are controlled for, which merits further research.

Female gender is associated with slightly lower math achievement at age 8 (Model 2: $p < 0.01$; Model 3: $p < 0.001$), lower growth rates in math (every $p < 0.001$), and less negative acceleration in math learning (every $p < 0.001$). From the regression results above and Figure 4 below, we can observe that in contrast to the abovementioned results on gender differences in reading, here we find an almost opposite pattern: female students begin kindergarten with math achievement similar to male students, but an achievement gap forms and widens through Grades K–5, where female students are slightly behind.

Figure 4. Gender differences in math growth

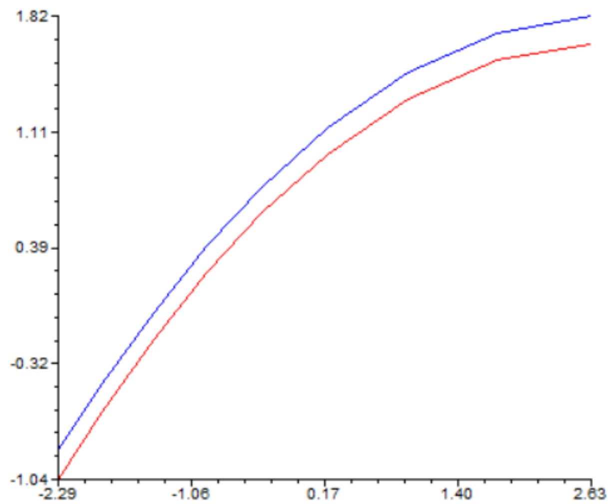


The *x* axis is age centered on 8, and the *y*-axis is the math theta score. The red (bottom) curve is for female students, and the blue (top) curve is for male students.

Hispanic students are associated with lower math achievement at age 8 when compared both with non-Hispanic students (Model 5, $p < 0.001$) and with white students (Model 6, $p < 0.05$). Hispanic students are not statistically significantly differentiated with non-Hispanic students in growth rates in math (Model 5), but they are associated with a lower growth rate when compared with white students of similar backgrounds (Model 6, $p < 0.01$). Hispanic students are associated with somewhat lower acceleration in math learning when compared with non-Hispanic students (Model 5, $p < 0.05$), but not when compared with white students (Model

6, statistically nonsignificant). Figure 5 below (based on Model 5) illustrates a persistent math achievement gap between Hispanic and non-Hispanic students over Grades K–5.

Figure 5. Hispanic vs. non-Hispanic students' math growth.



The *x* axis is age centered on 8, and the *y*-axis is the math theta score. The blue (top) curve is for non-Hispanic students, and the red (bottom) curve is for Hispanic students.

In regard to other races/ethnicities, compared with white students of similar backgrounds, Asian students are associated with higher math achievement at age 8 ($p < 0.001$) and lower rate of growth in math ($p < 0.001$). Black students are associated with lower math achievement at age 8 ($p < 0.05$) and lower growth rate in math ($p < 0.01$). Neither of these groups, however, are associated with a statistically significant difference in the acceleration of math growth compared to white students of similar backgrounds (each $p > 0.05$). Further, students in the Other racial/ethnic category are again not associated with statistically significantly different trajectories of math growth compared with white students of similar backgrounds ($p > 0.05$). Once more, statistical power and/or intragroup heterogeneity may be at play here.

Higher household SES is associated with higher math achievement at age 8 (every $p < 0.001$), lower growth rate in math (every $p < 0.001$), and less negative acceleration in math growth (every $p < 0.001$). The contribution of SES is again substantial: this time, around 1.2 SD of SES difference numerically cancels out the net effect of low early English proficiency on the intercept, and around 1.7 SD of SES difference cancels out the net effect of the same on the growth-rate term for math.

Living in a single-parent household at baseline is associated with somewhat lower math achievement at age 8 (Model 5: $p < 0.01$; Model 6: $p < 0.05$), a statistically nonsignificant difference in the rate of math growth, and somewhat more negative acceleration in math growth (every $p < 0.05$). Once more, the small but significant relationship between single parenthood and math growth trajectories should not be overlooked, and it should be noted that single parenthood may interact with household SES to complicate the picture.

Last but not least, enrollment in relatively impoverished school districts is associated with lower math achievement at age 8 (Model 5: $p < 0.01$; Model 6: $p < 0.05$), lower growth rate in math (every $p < 0.01$), and less negative acceleration in math growth (every $p < 0.001$). In contrast with the results for reading achievement, here it seems that the potentially negative effect of school district impoverishment on LM students' math development is perceptible even within the K–5 period.

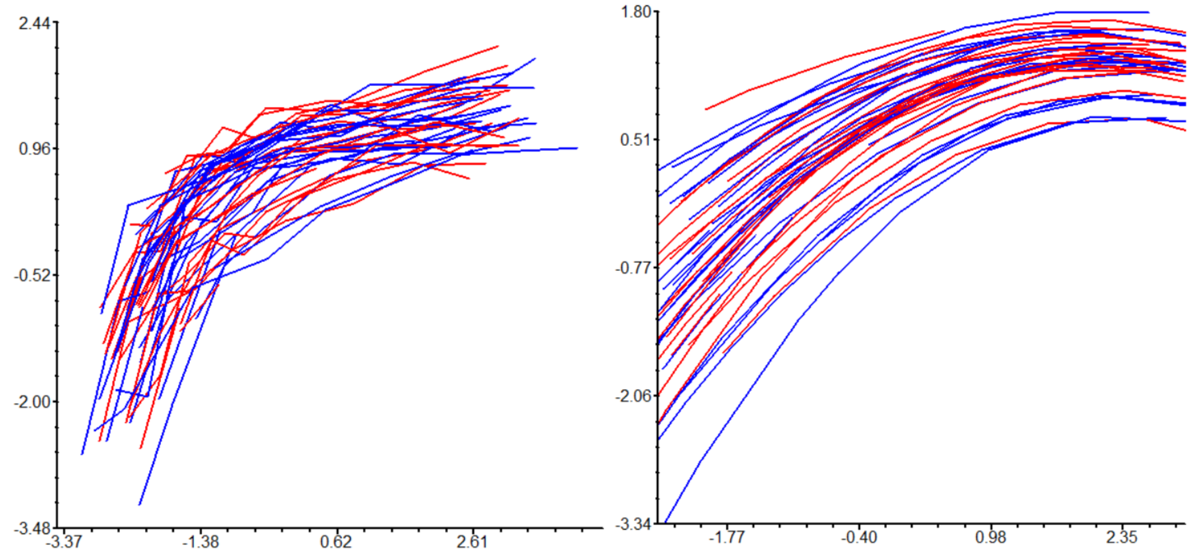
Robustness of findings

For all effects that are statistically significant, the directionality remains unchanged between the simpler and more complex models, except when it comes to the Hispanic binary variable that signaled two different comparisons: one with non-Hispanic students (Models 2 and 5) and the other with white students (Models 3 and 6).

Collinearity was checked between all child-level variables included in the models. The first step was to check for pairwise correlation, with the notable exception of the meaningless pairings between race/ethnicity dummies, which were by definition mutually exclusive. Among the Pearson correlation coefficients that resulted from this procedure, the highest coefficients by absolute value were between Hispanic ethnicity and household SES (r = -0.47), between school district poverty and SES (r = -0.36), and between Asian race and SES (r = 0.35). Further, in order to check for multicollinearity, Variance Inflation Factor (VIF) values were also computed for each of the predictor variables on the child level. The highest VIF values obtained were 3.3 for Hispanic ethnicity and 2.6 for Asian race. These values stay within acceptable thresholds such as 5 or 10 (e.g., O'brien, 2007).

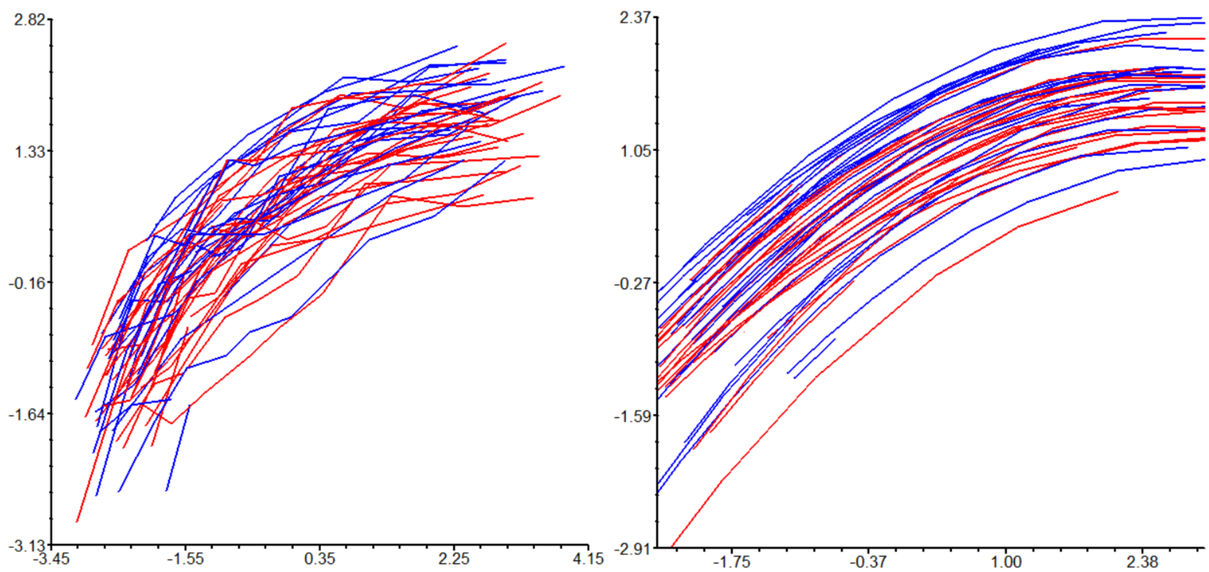
In regard to the modeling approach, the main reason for the use of quadratic models on the time-level in this thesis is that the quadratic model is the simplest polynomial model that has a curvature and can be readily estimated using HLM software. Polynomial coefficients are generally more readily interpretable, but especially so if we remain on the quadratic level and refrain from adding higher-power terms. Notably, while linear versions of the quadratic models were also attempted, i.e., without the quadratic term on the time-level and without the corresponding coefficients on the child-level, the linear versions were excluded from further analysis and reporting. Not only was the linear models' goodness-of-fit always worse than the analogous quadratic model, but the linear models also do not account for the curvature of the learning trajectories. Below are two pairs of graphs utilizing a 1% random sample of the data, featuring either the reading scores or the math scores. In each pair, the graph on the left presents raw, unadjusted student learning trajectories, while the graph on the right presents model-based estimated trajectories. Compared with linear models, quadratic models would seem to provide a more reasonable fit.

Figure 6. Raw trajectories and model-based estimated trajectories (Reading)



The *x* axis is age centered on 8, and the *y*-axis is the reading theta score. The red curves represent girls' trajectories, and the blue curves represent boys' trajectories.

Figure 7. Raw trajectories and model-based estimated trajectories (Math)



The *x* axis is age centered on 8, and the *y*-axis is the math theta score. The red curves represent girls' trajectories, and the blue curves represent boys' trajectories.

Conclusion

The multiple disadvantages that LM students often encounter need to be addressed, and within the current US educational framework a safe option is to focus on building viable and ideally self-improving linguistic repertoires in LM students. Due to the centrality of language in human communication in general and within formal education in particular, the development of language proficiency in the lingua franca may promise LM students not only better grades, but also potentially smoother navigation within a society that mostly speaks the lingua franca.

The main findings of this thesis are concerned with the association between LM students' early English proficiency and their learning trajectories in reading and math. Whether with or without adjustment for covariates, this thesis has found that there exists an achievement gap in reading and math between LM students with lower and higher early English proficiency. However, in the course of K–5, the students with lower proficiency gradually proceed toward closing these gaps. By the spring of Grade 5, it is estimated that the initial gap in reading narrowed by about 70% (around 1.3 years of typical learning) and the initial gap in math narrowed by about 47% (around 0.44 years of typical learning). Applied to real-world situations, these findings may be an indication of more sensible and more broadly accessible bilingual and ESL programs in the US, post-NCLB. However, one should be reminded that since language acquisition is destined to be a long-term endeavor, such linguistic (and sometimes content-knowledge) programmatic support must be continued or expanded in order that these results are maintained or further improved.

It is important to note that this thesis does come with a few limitations. First, it could be enhanced by adding theoretically motivated interaction terms. Second, other types of growth curves may also be explored, such as logistic or cubic. This may either corroborate or disprove some of the findings presented here. Finally, it would be ideal to have a better variable indicating the child's English (LF) proficiency, perhaps within the restricted-use dataset.

Data and code availability

The original data from ECLS-K:2011 is downloadable from the NCES website (https://nces.ed.gov/ecls/dataproducts.asp).

The code used for analysis can be provided upon request.

HLM analysis was conducted on software created by Raudenbush et al. (n.d.).

References

Abedi, J., Courtney, M., Leon, S., Kao, J., & Azzam, T. (2006). *English language learners and math achievement: A study of opportunity to learn and language accommodation* [Research Report]. National Center for Research on Evaluation, Standards, and Student Testing (CRESST). https://eric.ed.gov/?id=ED495848

Alvear, S. A. (2015). *Learning more than language: An examination of student achievement in English immersion and bilingual programs* [Master's thesis, Rice University]. Rice University Theses and Dissertations. https://hdl.handle.net/1911/87694

Bernhofer, J., & Tonin, M. (2022). The effect of the language of instruction on academic performance. *Labour Economics*, *78*, 102218. https://doi.org/10.1016/j.labeco.2022.102218

Bommes, M. (2012). Transnationalism or assimilation? In C. Boswell & G. D'Amato (Eds.), *Immigration and social systems: Collected essays of Michael Bommes* (pp. 107–124). Amsterdam University Press. https://www.jstor.org/stable/j.ctt46mtgj.10

Callahan, R. M. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal*, *42*(2), 305–328. https://doi.org/10.3102/00028312042002305

Cohen, G. (1984). The politics of bilingual education. *Oxford Review of Education*, *10*(2), 225–241. https://www.jstor.org/stable/1050049

Data Recognition Corporation. (n.d.). *Assess language proficiency in early childhood*. LAS Links. Retrieved July 22, 2025, from https://laslinks.com/prelas/

Gibney, D. T., Kelly, H., Rutherford-Quach, S., Riccards, J. B., & Parker, C. (2021). *Addressing the bilingual teacher shortage*. Comprehensive Center Network.

Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* (Policy Report 2000–1). University of California Linguistic Minority Research Institute.

Halle, T., Hair, E., Wandner, L., McNamara, M., & Chien, N. (2012). Predictors and outcomes of early versus later English language proficiency among English language learners. *Early Childhood Research Quarterly*, *27*(1), 1–20. https://doi.org/10.1016/j.ecresq.2011.07.004

Haneda, M. (2009). Learning about the past and preparing for the future: A longitudinal investigation of a Grade 7 'sheltered' social studies class. *Language and Education*, *23*(4), 335–352. https://doi.org/10.1080/09500780902954265

Hu, G., & Lei, J. (2014). English-medium instruction in Chinese higher education: A case study. *Higher Education*, *67*(5), 551–567. https://doi.org/10.1007/s10734-013-9661-5

Jimerson, S. R., Hong, S., Stage, S., & Gerber, M. (2013). Examining oral reading fluency trajectories among English language learners and English speaking students. *Journal of New Approaches in Educational Research*, *2*(1), 3–11. https://doi.org/10.7821/naer.2.1.3-11

Johnson, A. (2023). Achievement and growth for English learners. *Educational Policy*, *37*(4), 1044–1069. https://doi.org/10.1177/08959048211049419

Kieffer, M. J. (2011). Converging trajectories: Reading growth in language minority learners and their classmates, kindergarten to Grade 8. *American Educational Research Journal*, *48*(5), 1187–1225. https://doi.org/10.3102/0002831211419490

Kieffer, M. J., & Thompson, K. D. (2018). Hidden progress of multilingual students on NAEP. *Educational Researcher*, *47*(6), 391–398. https://doi.org/10.3102/0013189X18777740

Nurius, P. S., Prince, D. M., & Rocha, A. (2015). Cumulative disadvantage and youth well-being: A multi-domain examination with life course implications. *Child & Adolescent Social Work Journal 32*(6), 567–576. https://doi.org/10.1007/s10560-015-0396-2

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Okoye-Johnson, O. (2011). Intangible heritage of Standard English learners: the "invisible" subgroup in the United States of America? Implications for closing the achievement gap. *Sage Open*, *1*(1), 2158244011408441. https://doi.org/10.1177/2158244011408441

Orfield, G. (1986). Hispanic education: Challenges, research, and policies. *American Journal of Education*, *95*(1), 1–25. https://doi.org/10.1086/444290

Pilger Suhr, M., Nese, J. F. T., & Alonzo, J. (2021). Parallel reading and mathematics growth for English learners: Does timing of reclassification matter? *Journal of School Psychology*, *85*, 94–112. https://doi.org/10.1016/j.jsp.2021.02.003

Raudenbush, S., Bryk, A., & Congdon, R. (n.d.). *HLM for Windows* (Version 7.92d) [Windows].

Sánchez, M. T., Menken, K., & Pappas, L. N. (2022). "What are you doing to us?!": Mediating English-only policies to sustain a bilingual education program. *International Multilingual Research Journal*, *16*(4), 291–307. https://doi.org/10.1080/19313152.2021.2015936

Schmid, C. L. (2001). Educational achievement, language-minority students, and the new second generation. *Sociology of Education*, *74*, 71. https://doi.org/10.2307/2673254

Tourangeau, K., Nord, C., Lê, T., Wallner-Allen, K., Vaden-Kiernan, N., Blaker, L., and Najarian, M. (2019). *Early Childhood Longitudinal Study, Kindergarten Class of 2010–11 (ECLS-K:2011) User's Manual for the ECLS-K:2011 Kindergarten–Fifth Grade Data*

File and Electronic Codebook, Public Version (NCES 2019-051). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

U.S. Census Bureau (1992). *Summary tables on language use and English ability: 1990*. https://www.census.gov/data/tables/time-series/dec/cph-series/cph-l/cph-l-96.html

U.S. Census Bureau (2003). *Summary tables on language use and English ability: 2000*. https://www.census.gov/data/tables/2000/dec/phc-t-20.html

U.S. Census Bureau. (n.d. a). Language spoken at home. *American Community Survey, ACS 1-year estimates subject tables, Table S1601*. Retrieved June 24, 2025, from https://data.census.gov/table/ACSST1Y2010.S1601?q=Language+Spoken+at+Home

U.S. Census Bureau. (n.d. b). Language spoken at home. *American Community Survey, ACS 1-year estimates subject tables, Table S1601*. Retrieved June 24, 2025, from https://data.census.gov/table/ACSST1Y2023.S1601?q=Language+Spoken+at+Home

Ünsal, Z., Jakobson, B., Molander, B.-O., & Wickman, P.-O. (2018). Science education in a bilingual class: Problematising a translational practice. *Cultural Studies of Science Education*, *13*(2), 317–340. https://doi.org/10.1007/s11422-016-9747-3

Wang, S., Lang, N., Bunch, G. C., Basch, S., McHugh, S. R., Huitzilopochtli, S., & Callanan, M. (2021). Dismantling persistent deficit narratives about the language and literacy of culturally and linguistically minoritized children and youth: counter-possibilities. *Frontiers in Education*, *6*, 641796. https://doi.org/10.3389/feduc.2021.641796

Winsler, A., Rozell, N., Tucker, T. L., & Norvell, G. (2023). Earlier mastery of English predicts 5th Grade academic outcomes for low-income dual language learners in Miami, USA. *Bilingualism: Language and Cognition*, *26*(5), 910–923. https://doi.org/10.1017/S136672892300007X

Wright, W. E., & Li, X. (2008). High-stakes math tests: How No Child Left Behind leaves newcomer English language learners behind. *Language Policy*, *7*(3), 237–266. https://doi.org/10.1007/s10993-008-9099-2

Wright, W. E., & Ricento, T. (2017). Language policy and education in the USA. In T. L. McCarty & S. May (Eds.), *Language policy and political issues in education* (pp. 383–400). Springer International Publishing. https://doi.org/10.1007/978-3-319-02344-1_29

Yang, Y., Wang, H., Zhang, L., Sylvia, S., Luo, R., Shi, Y., Wang, W., & Rozelle, S. (2015). The Han-minority achievement gap, language, and returns to schools in rural China. *Economic Development and Cultural Change*, *63*(2), 319–359. https://doi.org/10.1086/679070

Yu, K. (2024). Bilingual program effectiveness: An evaluation of meta-analytic methods and findings. *International Journal of Bilingual Education and Bilingualism*, *27*(6), 744–771. https://doi.org/10.1080/13670050.2023.2232522

Zhou, M. (1997). Segmented assimilation: Issues, controversies, and recent research on the new second generation. *The International Migration Review*, *31*(4), 975–1008. https://www.jstor.org/stable/2547421