Open Access Articles

English Learners' Time to Reclassification: An Analysis

The Faculty of Oregon State University has made this article openly available.
Please share how this access benefits you. Your story matters.

Citation	Thompson, K. D. (2015). English Learners' Time to Reclassification: An Analysis. [Article in Press]. Educational Policy. doi:10.1177/0895904815598394
DOI	10.1177/0895904815598394
Publisher	SAGE Publications
Version	Accepted Manuscript
Terms of Use	http://cdss.library.oregonstate.edu/sa-termsofuse

Title:

English learners' time to reclassification: An analysis

Author:

Name:	Karen D. Thompson
Affiliation:	Oregon State University
Street Address:	301B Furman Hall Corvallis, OR 97337
Email:	karen.thompson@oregonstate.edu
Phone:	(541) 737-2988

Note: This research was made possible in part by a Stanford Interdisciplinary Graduate Fellowship and by a Stanford Graduate School of Education Dissertation Support Grant. I thank LAUSD staff, particularly Cynthia Lim and Kathy Hayes, for sharing district data and for providing important insights about data features, program characteristics, and interpretations of findings. In addition, I thank Kenji Hakuta, Sean Reardon, Guadalupe Valdés, Ed Haertel, John Rickford, Patricia Gándara, Robert Linquanti, Joseph Robinson-Cimpian, and Ilana Umansky for useful comments and discussions during the writing of the manuscript. The research presented here does not necessarily reflect the views of any funder or LAUSD.

Abstract

This study uses 9 years of longitudinal, student-level data from the Los Angeles Unified School District to provide updated, empirically–based estimates of the time necessary for English learners (ELs) to become reclassified as proficient in English, as well as factors associated with variation in time to reclassification. To illustrate how different aspects of proficiency develop, estimates of the time necessary for ELs to attain six separate reclassification criteria are provided. Findings corroborate prior cross-sectional research suggesting that the development of full proficiency in a second language typically takes 4 to 7 years. However, after 9 years in the district, approximately one-fourth of students had not been reclassified. There appears to be a *reclassification window* during the upper elementary grades, and students not reclassified by this point in time become less likely ever to do so. Findings illustrate the crucial role that students' initial academic language proficiencies, both in English and their primary language, play in their likelihood of reclassification. This work has implications for the design of next-generation assessment and accountability systems, as well as for instructional practices.

A large, growing proportion of students enter U.S. schools as English learners (ELs). Current accountability systems require that states establish targets for ELs' English proficiency development. However, these targets are not always empirically grounded. Using 9 years of student-level, longitudinal data from the Los Angeles Unified School District (LAUSD), this study examines students' English proficiency development, estimating how long it takes students who enter the district as ELs in kindergarten to attain each of six separate criteria necessary to be reclassified and exit EL services.

Results corroborate prior research suggesting that English proficiency, when defined to encompass text-based literacy practices, does not develop quickly. Although a majority of students attain speaking and listening proficiency in English after only 2 years in the district, attaining proficiency on measures of reading and writing in English takes considerably longer. Specifically, the time necessary for at least 60% of students who enter LAUSD as ELs in kindergarten to score proficient on literacy-based measures ranges from 4 to 7 years (4 years for the content-area English Language Arts assessment, 4 years for the writing portion of the English Language Proficiency (ELP) assessment, 5 years for the writing portion of the ELP assessment, 5 years for all domains of the ELP assessment, and between 6 and 7 years for meeting all reclassification criteria simultaneously).

Boys, native Spanish speakers, and students whose parents have lower levels of education are all less likely to be reclassified than their peers, after controlling for other factors. In addition, students' likelihood of reclassification varies dramatically based on their initial academic language proficiencies both in English and their primary language. The construct of academic language proficiency here refers to students' use of the types of language valued and used in school contexts. Problematic aspects of language proficiency assesments are discussed later in

the article. Students who enter kindergarten with beginning levels of academic language proficiency are 24% less likely to reclassify after 9 years than their peers who enter with high levels of both. Because students' initial academic language proficiencies both in English and their primary language have such strong relationships to likelihood of reclassification, it would be helpful if targets for the time frames in which students are expected to develop English proficiency took these factors into account.

Finally, after 9 years in the district, one-fourth of students had not yet been reclassified as proficient in English. More than 30% of these students qualified for special education. This points to a pressing need for research and innovation in appropriate identification, placement, assessment, and services for ELs who may have disabilities.

As states overhaul their assessment and accountability systems to align with rigorous new college- and career-ready standards, it is vital that analyses of empirical data, including data on ELs' language acquisition trajectories, inform the design of these systems.

Background

Approximately one in five children in the United States speak a language other than English at home (Ryan, 2013), and approximately half of this group have not yet developed proficiency in English (Boyle, Taylor, Hurlburt, & Soga, 2010). Although 5 million students are currently classified as ELs, stark discrepancies exist between the achievement of ELs and their native-English speaking peers across a variety of outcomes, including achievement test scores and graduation rates (Fry, 2007; Aud et al., 2010). Most students initially classified as ELs upon school entry eventually meet the necessary criteria to be reclassified and shed the EL label. Although there is a growing body of literature on the time necessary for ELs to be reclassified that uses rigorous empirical methods and longitudinal student-level data (e.g., Conger, 2009;

Parrish, Perez, Merickel, & Linquanti, 2006; Slama, 2014; Umansky & Reardon, 2014), key questions remain. In particular, criteria used by education agencies to determine when ELs should be reclassified vary widely (e.g., Abedi, 2008; Hill, Weston, & Hayes, 2014; Wolf et al., 2008). Yet limited research exists about the time necessary for ELs to attain separate components of these reclassification criteria, limiting the generalizability of much existing research on time to reclassification. In addition, the role of students' initial academic language proficiencies—both in English and in their primary language—on the timing of reclassification has not been explored in detail.

More rigorous analysis of the normative time required to attain particular reclassification criteria is especially vital because accountability systems establish targets for ELs' English proficiency development. At present, many of these targets are not empirically grounded. As the Elementary and Secondary Education Act (ESEA) comes up for reauthorization and as states implement a new generation of standards and assessments, research on time to proficiency has the potential to influence policymakers as they establish and revise targets for ELs' English proficiency development and design systems to monitor ELs' progress. This study draws on 9 years of longitudinal, student-level data from LAUSD, which enrolls more ELs than any other district in the nation, to provide updated, empirically-based estimates of the time necessary for ELs to be reclassified as proficient in English. In addition, to illustrate how different aspects of proficiency develop, this study provides estimates of the time necessary for ELs to attain proficiency in six separate reclassification criteria, including listening, speaking, reading, and writing sections of ELP assessments and scores on content-area English language arts assessments. Factors related to variation in time to reclassification are also analyzed, with a

particular focus on the role of students' initial academic language proficiencies, both in English and in their primary language, on later outcomes.

What Are the Legal Underpinnings of the EL Classification and Reclassification System?

Much about the educational experience of an EL rests on whether she experiences a single dichotomous event: reclassification. To understand the significance of reclassification, it is useful to understand how students initially get classified as ELs and what this label means. Under the 1974 Supreme Court ruling in *Lau v. Nichols*, school systems must "take affirmative steps" to teach English to those not yet fluent in the language while also providing access to the general curriculum. Title III of ESEA provides local education agencies with supplemental funding for the special services ELs are entitled to under *Lau* and requires schools to identify students who require language support services. When a student enters school, her parents complete a Home Language Survey. If the parents indicate that they speak a language other than English at home, the student must then take an ELP assessment. If the student scores below the established English proficiency criteria, the student is considered an EL.

Under the federal No Child Left Behind Act (2001), all local education agencies must administer ELP assessments to ELs each year. States must define English-proficient performance standards on this assessment. When a student attains that performance standard on the ELP assessment, and, in the case of most states, also attains additional state- and/or district-defined criteria, the student is reclassified and sheds the EL label. The specific actions triggered by reclassification vary by state, district, and school but typically include the end of placement in courses designed for ELs.

How Is EL Reclassification Determined?

NCLB defines an EL as:

> an individual . . . whose difficulties in speaking, reading, writing, or understanding the English language . . . may deny the individual the ability to . . . successfully achieve in classrooms where the language of instruction is English. (Title XI, § 9101(25)(D))

The confusion surrounding the reclassification of ELs stems, in part, from disagreement about when students no longer fit this description. Although all states administer ELP assessments to ELs, no agreed-upon definition of proficiency exists. NCLB simply declares that states must measure "students' oral language, reading, and writing skills in English" (Title I, Section 1111(B)(3)). An initiative is underway to establish a common definition of ELs (Linquanti & Cook, 2013), but no consensus currently exists.

In their analysis of state practices related to English proficiency assessment, Wolf and colleagues (2008) determined that during the 2006-2007 school year, states used 30 different ELP assessments. In recent years, many states have joined two large consortia working to develop new ELP assessments aligned to the Common Core State Standards. However, nine states, including California, Texas, and New York, are not part of either ELP assessment consortium and plan to continue administering their own ELP assessments (Linquanti & Cook, 2013).

In addition to wide variation in ELP assessments, other reclassification criteria vary widely, as well. A student who would be considered an EL in one district might be reclassified as English proficient in another district simply because of differences in reclassification criteria (e.g., Abedi, 2008; Hill et al., 2014; Parrish et al., 2006). Wolf and colleagues (2008) found that while 12 states considered only students' results on the ELP assessment when determining students' eligibility for reclassification, the majority of states considered additional factors, most

often content-area achievement scores, which were considered as part of the reclassification decisions in 28 states. Other factors considered in some contexts include school personnel input, parent/guardian input, and student grades. The normative time needed for ELs to be reclassified in one school system may be quite different from the normative time needed in another school system because of differences in reclassification criteria, thus limiting the generalizability of much prior research on time to reclassification. In contrast, this study analyzes the time necessary for ELs to attain six separate reclassification criteria, demonstrating how expected time frames may vary depending on the particular reclassification criteria education agencies apply.

What Expectations About the Time Necessary for English Learners to Be Reclassified Are Established by Current Policies?

NCLB requires states to establish three distinct Annual Measurable Achievement Objectives (AMAOs) to track ELs' linguistic and academic achievement. First, states must set annually increasing performance targets for the percent of ELs *making progress* toward English proficiency, as measured by the states' ELP assessment. Second, states must set targets for the percent of ELs who will *attain* English proficiency on the state ELP assessment. Third, states must set targets for the percent of ELs who will score proficient on the states' content-area assessments. Although some states, including California, used empirical data to establish AMAO targets, a recent evaluation of Title III notes, "Most states lacked empirical data—and indeed, even standards-based ELP assessments—with which to determine progress criteria and performance targets" (Cook, Linquanti, Chinen, & Jung, 2012, p. 1).

Meanwhile, voters in several states have approved ballot measures that attempt to define the time necessary for ELs to attain proficiency. For example, in California, Proposition 227,

approved by voters in 1998, severely restricted bilingual education and required ELs to be placed in special "structured English immersion classrooms" with other ELs for "a transition period not normally intended to exceed one year." After this transition period, students are expected to have acquired "reasonable English fluency" such that they can successfully achieve in "English language mainstream classrooms" (Cal. Ed Code § 305-306). These ballot initiatives, while shaping the types of instructional programs that districts offer, have no direct relationship to the AMAOs required by Title III and do not supersede federal legal requirements for the education of ELs.

What Factors Are Associated With Variation in the Time Necessary to Develop Proficiency in a Second Language?

Both linguists and scholars within education have explored differences in the speed with which individuals acquire a second language. Not surprisingly, those with higher levels of initial English proficiency (Conger, 2009; Cook, Boals, Wilmes, & Santos, 2007) tend to develop proficiency in a second language more quickly than their peers. Researchers have also found a strong yet complex relationship between students' language and literacy skills in their primary language (L1) and their acquisition of English (L2). Students may be able to transfer some skills from their L1 when acquiring English (e.g., Genesee, Geva, Dressler, & Kamil, 2008). However, the degree of linguistic similarity between students' home language and English may affect the extent of the transfer process (Genesee et al., 2008). Although existing studies provide important information about the cross-linguistic relationships between specific language and literacy skills, no existing large-scale, longitudinal study was identified that tests the relationship between students' initial L1 academic language skills and time to reclassification.

Given the strong relationship between socioeconomic status (SES) and a wide variety of literacy outcomes, it is not surprising that researchers have found a strong relationship between SES and second language acquisition (August & Shanahan, 2006; Conger, 2009; Hakuta, Goto Butler, & Witt, 2000). The home language spoken by students' families also appears as a factor related to the rate of second language acquisition in a variety of studies. Native Spanish speakers often appear to progress toward English proficiency more slowly than students from other home language backgrounds (Hill, 2004; Parrish et al., 2006; Slama, 2014), which may be the result of higher levels of poverty and social disadvantage experienced by Latino families (Grissom, 2004). Girls (Conger, 2009; Grissom, 2004) also tend to develop proficiency in English more quickly than their male peers, although the reasons for this have not been fully explored.

A vast literature has examined the impact of some amount of primary language instruction on second language acquisition. Although a thorough review of this literature is beyond the scope of this article, multiple meta-analyses have found modest positive effects for bilingual education on a variety of second language outcomes (August & Shanahan, 2006; Greene, 1997; Slavin & Cheung, 2005; Willig, 1985). Although students in bilingual programs may, in some cases, take longer to develop English proficiency in the short term, this may be due to negative selection mechanisms rather than the effect of the program itself (Conger, 2010). Results from a major recent randomized controlled trial suggest that students learn to read in English equally well in bilingual programs and in English-only programs, indicating that quality of instruction may matter more than language of instruction (Slavin, Madden, Calderón, Chamberlain, & Hennessy, 2011).

In this study, available data allow exploration of the relationship between time to reclassification in LAUSD and all of the above factors: initial academic English proficiency,

initial academic L1 proficiency, SES, home language, gender, and participation in bilingual education. Researchers have suggested a variety of other factors that lie outside the scope of this study but may affect ELs' opportunity to learn and their English acquisition. Studies have documented that younger children (Birdsong, 1999; Conger, 2009) and those with stronger general cognitive abilities (Genesee et al., 2008) all tend to develop proficiency in a second language more quickly than their peers. In addition, studies indicate that ELs typically attend under-resourced schools, are enrolled in classrooms with less-experienced teachers, and may encounter discrimination and low expectations from teachers and counselors (e.g., Gándara, Maxwell-Jolly, & Rumberger, 2008). Although available data do not allow this study to test the relationship of these factors and time to reclassification, it is important to remember the wide variety of contextual factors that may influence ELs' outcomes.

What Does Prior Research Say About the Time Necessary for Individuals to Develop Proficiency in a Second Language?

Researchers have undertaken multiple analyses of how long it takes individuals to develop proficiency in a second language. In perhaps the most well-known study, Hakuta et al. (2000) use data from four school districts to explore how long acquisition of English proficiency takes. They found that acquisition of oral English proficiency takes 3 to 5 years, but acquisition of academic English, as measured by a variety of standardized assessments, takes 4 to 7 years. However, students were not followed over time but were given a variety of assessments at one time point and results were disaggregated based on students' length of residence in the country.

A growing body of research has used student-level, longitudinal data to analyze time to reclassification, testing Hakuta et al.'s (2000) findings. Conger applied survival analysis to longitudinal data from New York City to investigate the role of age (2009) and the role of

bilingual education (2010) on time to reclassification, while Slama (2014) applied the same methods to analyze time to reclassification using statewide data from Massachusetts. Umansky and Reardon (2014) used data from a large California district to analyze the role of instructional program models on time to reclassification, and Parrish et al. (2006) analyzed time to reclassification across California as part of a broader analysis of the impact of Proposition 227 on student outcomes in the state. Both Conger (2009) and Slama (2014) found that approximately half of ELs were reclassified within 3 years of school entry in New York City and Massachusetts, respectively. However, Parrish et al. (2006) and Umanksy and Reardon (2014) found that reclassification typically took much longer for students in California. For example, Umansky and Reardon (2014) found that the median time necessary for reclassification to occur for the Latino ELs in their sample was 8 years, and 25% of Latino ELs had not been reclassified by 12th grade.[1]

The differences in estimates of time to reclassification across these studies are likely explained by differences in reclassification criteria used by states. Both New York and Massachusetts rely on the state ELP assessment for reclassification decisions and do not consider content-area assessments in English language arts (Wolf et al., 2008); thus, students reclassify more quickly in these states than in California and the 27 other states in which content-area assessments are considered as part of the reclassification decision. This highlights the need for additional research that analyzes the time necessary for students to attain separate reclassification criteria, including separate criteria on state ELP assessments and criteria related to state content-area assessments.

Additional studies examining time to reclassification have either analyzed fewer years of data or relied on less rigorous methodologies. Cook and colleagues (2007) used 3 years of

longitudinal student-level data from three states and 2 years of longitudinal data from an additional nine states to estimate time required to acquire English proficiency. In their sample, only 3% of students who started with beginning proficiency levels had achieved full proficiency 3 years later. Although this study provides important information about the early stages of English acquisition, its short time frame prevents the authors from analyzing the full acquisition process.

Contributions of the Current Study

Because current accountability systems require states to establish targets for the number/percentage of ELs attaining English proficiency, an analysis of longitudinal data, focused on determining how long it is currently taking ELs to attain separate reclassification criteria, has the potential to inform the target-setting process as next-generation assessments come online. This study uses 9 years of longitudinal, student-level data from LAUSD to examine three inter-related research questions:

Research Question 1: How long does it take for students entering school as ELs in kindergarten to be reclassified as proficient in English, according to district reclassification criteria?

Research Question 2: What factors are related to variation in time to reclassification for students entering as ELs in kindergarten?

Research Question 3: How long does it take students who enter school as ELs in kindergarten to attain individual reclassification criteria?

Many previous studies rely primarily on data collected before the implementation of NCLB in 2002, which was the first federal legislation to mandate annual assessment of English

proficiency for ELs and to mandate that states set targets for progress in English proficiency across years (Conger, 2009, 2010; Hakuta et al., 2000; Parrish et al., 2006). By using post-NCLB data, this study provides an updated estimate of time to reclassification under these new regulations. In addition, this study is unique in examining both the role of initial academic English proficiency and initial academic L1 proficiency on time to reclassification. Finally, this study analyzes how long it takes students to attain separate components of reclassification, including minimum listening, speaking, reading, and writing scores on the ELP assessment and minimum scores on the content-area assessment in English language arts, building a detailed picture of how English proficiency develops over time.

Reclassification in LAUSD

The criteria that ELs in LAUSD must meet to be reclassified as proficient in English have remained relatively stable over time, with two key exceptions. Like many districts across the state, LAUSD employed reclassification guidelines established by the California State Board of Education. To be reclassified during the time period examined here, students needed to attain an overall score of 4 or 5 (out of 5) on the California English Language Development Test (CELDT). In addition, students needed to attain scores at Level 3 (out of 5) or higher in each of the domains assessed by the CELDT: Listening, Speaking, Reading, and Writing. Finally, students needed to score at the Basic level or above on the English Language Arts section of the California Standards Test, the state's content assessment (California Department of Education, 2013). LAUSD added one additional criterion that students needed to meet to be reclassified. Students at the elementary grades needed to attain a progress report score of 3 or higher in each domain of English Language Arts, while students at the secondary level needed to attain a grade of C or higher in their English or Advanced ESL

(English as a second language) course. (In earlier years, the district also required minimum grades in Mathematics, but that criterion was dropped at the beginning of the 2006-2007 school year.)

Although reclassification criteria in LAUSD did not change substantially during the time period under study, two changes in the CELDT may affect estimates presented here. First, in 2006-2007, Listening and Speaking were split into two distinct domains with separate scale scores. Also in 2006-2007, a common CELDT scale was established that allowed more reliable score comparisons from year to year. In the process, cut scores were reset, with the effect of making the test more challenging (California Department of Education, 2007). Because CELDT scale scores before and after the rescaling are not comparable, CELDT scores from the pre-2007 period were converted to the "common scale" (i.e., new scale) using the concordance tables in CELDT 2007 Technical Manual (California Department of Education, 2007). These equated scores were used to explore the relationship between initial English proficiency and time to reclassification.

As discussed previously, reclassification criteria vary dramatically across states and across districts within states (e.g., Abedi, 2008; Hill et al., 2014; Wolf et al., 2008). Exploring when students attain specific components of the reclassification criteria rather than just when students are reclassified addresses some of the generalizability concerns that arise from the variation in reclassification criteria. LAUSD set the CELDT and CST cutpoints at the lowest level suggested by the state. Therefore, estimates for how long it takes students to attain these criteria should be considered a floor. ELs in districts with higher CELDT and CST cutpoints might take longer to be reclassified. Similarly, districts that include more or fewer additional reclassification criteria compared with LAUSD could expect to see corresponding shifts in the precise amount of time students need to attain all criteria. However, by separately examining

how long ELs take to achieve reclassification criteria related to the state ELP assessment and how long ELs take to achieve reclassification criteria related to the state content-area assessment, we can develop a better understanding of how time to reclassification may vary between states that primarily use ELP assessments to make decisions about reclassification compared with states that use both ELP and content-area assessments (Wolf et al., 2008).

Data

This study uses 9 years of longitudinal student-level data from LAUSD. The analytic sample consists of longitudinal student-level data for 202,931 students who entered LAUSD as ELs in kindergarten. The dataset spans the years 2001-2002 through 2009-2010 and contains information on students from eight different cohorts, defined by the year that students entered kindergarten. During these 9 years, data on students are available for as long as they were enrolled in the district. Table 1 reports mean characteristics of the analytic sample separately by cohort. The sample overwhelmingly consists of native Spanish speakers (94%) who come from low-income families (95%, as measured by eligibility for free or reduced-price lunch). Approximately 10% of students were ever in a bilingual program. A similar percentage ever qualified for special education services. The modal value for parent education is "not a high school graduate."

LAUSD assesses students in their primary language upon entry to the district. The L1 assessment used for the vast majority of students was the Pre-LAS (used for 94% of the sample), an assessment whose construct validity has been criticized (MacSwan, Rolstad, & Glass, 2002). Researchers have argued that labeling students as non-proficient speakers of their primary language is both incorrect and damaging, signifying flaws with the assessments rather than a lack of linguistic knowledge on students' part (MacSwan, 2005; MacSwan et al., 2002). These

assessments may be best understood as assessments of students' academic L1 proficiency. Similarly, the CELDT is explicitly designed to measure students' proficiency in "language used in academic settings" (California Department of Education, 2007). Therefore, this study refers to assessments at kindergarten entry as measuring students' initial academic L1 proficiency and their initial academic English proficiency, meaning students' use of the types of language valued and used in school contexts. These assessments do not capture students' full range of language use in other contexts or their knowledge of non-dominant language varieties. Students tend to enter the district between the Beginning and Early Intermediate levels of academic English proficiency (with a mean proficiency level of 1.628 on a scale of 1 to 5, using scores that have been equated to be reported on the same scale throughout the study period), as measured by the CELDT in the fall of kindergarten. The mean value for students' initial academic L1 proficiency at kindergarten entry is 2.119 on a scale from 1 to 4.

INSERT TABLE 1 HERE

The primary outcome of interest is students' date of reclassification as proficient in English. To understand when students meet separate reclassification criteria, this study uses students' scale scores and proficiency levels on the Listening, Speaking, Reading, and Writing sections of the CELDT, as well as their overall CELDT scale scores and proficiency levels. In addition, students' scale scores and proficiency levels on the California Standards Test in English Language Arts (CST-ELA) are used.

The sample is restricted to students who entered LAUSD as ELs in kindergarten. Although understanding how age at school entry relates to the time necessary to attain English proficiency is of interest, several factors led to the removal of that question from consideration in this study. Students who enter at other grades may have either attended school in other districts within the United States

or in other countries prior to enrolling in LAUSD. Limiting the sample to ELs who enter LAUSD in kindergarten ensures that the district was responsible for students' schooling until students disappear from the dataset (either because the student left the district or data collection ended). As noted earlier, prior research suggests that students tend to attain English proficiency more quickly when they enter U.S. schools at earlier grade levels (Conger, 2009; Cook et al., 2007). Therefore, estimates of time to reclassification from this study could be considered lower bounds because they are based on students entering in kindergarten.

As with most datasets, information for some variables is missing. The variable with by far the highest percentage of missing information is parent education, which is missing for 33% of the sample. To address the issue of missing data, this study uses multiple imputation. Specifically, the iterated chain equations (ICE) procedure in Stata was used to impute five different values for the missing information, using all the variables in the model in chained regression equations to predict the missing values (Van Buuren, Boshuizen, & Knook, 1999). The models were run on all five datasets with imputed values, and then the resulting estimates were averaged using Stata's multiple imputation commands. (The models were also estimated using only observations with complete information on all variables, and results were very similar to the results obtained using the datasets with imputed information.)

Methods

Research Question 1: How Long Does It Take for Students Entering School as ELs in Kindergarten to be Reclassified as Proficient in English?

To estimate how long it takes students classified as ELs at kindergarten entry to be reclassified, this study uses discrete-time survival analysis (Singer & Willett, 2003). This technique is superior to standard regression methods because it is explicitly designed to answer

questions about the likelihood that a particular event will occur over time. Survival analysis accounts for a key complication of analyzing how long it takes students to be reclassified: Some students never experience reclassification. In survival analysis, such students are deemed *censored*. Students may be censored either because they leave the district or data collection ends. Rather than ignoring students who never experience reclassification, survival analysis uses information from all students in the dataset up until the point at which they experience reclassification or are censored.

Using survival analysis, two functions of interest are calculated. First, the hazard function provides the conditional probability that an individual will be reclassified during a particular year, given that the individual had not been reclassified already. This enables identification of particular time points during which reclassification is especially likely to occur. Second, the survivor function represents the probability that an individual will *not* be reclassified after the passage of a range of time. This enables determination of how likely it is, for example, that a student could remain in the district for 9 years without being reclassified as proficient in English.

To estimate the probability that individual *i* was reclassified during time period *j*, first the baseline discrete-time hazard model is estimated:

$$\ln\left(\frac{h(t_{ij})}{1-h(t_{ij})}\right) = \left[\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \ldots + \alpha_J D_{Jij}\right] + \varepsilon_{ij}, \text{(Equation 1)}$$

where D_1 through D_J represents a series of time indicators for each year in which the student was observed, through *j* number of years, and the parameter estimates α_1 through α_J can be converted into the estimated probability of being reclassified in each of these years provided that the student was not reclassified in the previous years (the hazard function). These estimated

parameters can then be used to estimate the cumulative percentage of students who have not been reclassified after a particular number of years has elapsed (the survivor function):

$$\hat{s}(t_j) = \hat{s}(t_{j-1})\left[1 - \hat{h}(t_j)\right], \qquad \text{(Equation 2)}$$

in which the survivor function for a particular time period, j, is calculated by multiplying the survival probability for the previous time period by one minus the hazard probability for the current time period. For ease of interpretation, the complement of the survivor function, $1 - s(t_j)$, is reported. This is known as the cumulative failure function and represents the cumulative probability that a student *will* experience reclassification.

Research Question 2: What Factors Are Related to Variation in Time to Reclassification for Students Entering as ELs in Kindergarten?

From these baseline discrete-time hazard models, this study then explores factors associated with variation in time to reclassification by introducing predictors into the model. To account for the nesting of observations at the school level, standard errors are clustered by school.[2] The main equation for the fitted models is,

$$\begin{aligned}\ln\left(\frac{h(t_{ij})}{1-h(t_{ij})}\right) = &\left[\alpha_1 D_{1ij} + \alpha_2 D_{2ij} + \ldots + \alpha_J D_{Jij}\right] \\ &+\beta_1 X_i + \beta_2 P_i + \beta_3 E_i + \beta_4 L_i + \beta_5 P_i \times \\ &T_j + \beta_6 E_i \times T_j + \beta_7 L_i \times T_j + \beta_8 E_i \times L_i + C_i + \varepsilon_{ij},\end{aligned} \qquad \text{(Equation 3)}$$

where D_1 through D_J represents a series of time indicators for each year in which the student was observed, through j number of years, X_i represents a vector of student demographic characteristics, P_i is an indicator variable denoting whether the student was ever enrolled in a

bilingual program, E_i is a continuous variable representing the student's initial academic English proficiency at kindergarten entry (as measured by scale score on the CELDT overall, equated to be comparable throughout the duration of the study), L_i is an ordinal variable representing the student's initial academic L1 proficiency (as measured by district assessments), T_j is a continuous variable representing the number of years in which the student has already been enrolled in LAUSD, $P_i \times T_j$ represents the interaction between whether the student was ever enrolled in a bilingual program and the number of years in which the student has already been enrolled in the district, $E_i \times T_j$ represents the interaction between the student's initial academic English proficiency and the number of years in which the student has already been enrolled in the district, $L_i \times T_j$ represents the interaction between the student's initial academic L1 proficiency and the number of years in which the student has already been enrolled in the district, $E_i \times L_i$ represents the interaction between the student's initial academic English proficiency and their initial academic L1 proficiency, and C_i represents fixed effects for cohort.

It is important to understand the way that time functions in these models and the role that information from different cohorts plays. Time is defined in terms of the duration of a student's enrollment in LAUSD. Also, the number of years of data available for each cohort varies. For the cohort who entered kindergarten in 2001-2002, 9 years of data are available, through 2010. However, for the cohort who entered kindergarten in 2008-2009, only 2 years of data are available, again through 2010. Because at least 2 years of data are available from all cohorts, data from all cohorts are used to estimate the likelihood of reclassification after 1 or 2 years in the district. However, because 9 years of data are available only from the earliest cohort, only information from this earliest cohort is used to estimate the likelihood of reclassification after 9

years. As the number of years in the district increases, estimates of the likelihood of reclassification are based on fewer cohorts.

Survival analysis rests on two key assumptions. The first assumption relates to censoring. Recall that censoring occurs when a student is not reclassified during the time period under study, either because data collection ended before a student was reclassified or because the student left the district before being reclassified. For survival analysis to be valid, censoring must be non-informative. The point at which censoring occurs must not relate to students' likelihood of being reclassified. While proving that censoring is non-informative is difficult, we can determine the proportion of students who were censored because data collection ended; for these students, censoring is by definition non-informative. Of the 125,718 students in the dataset who were censored, a majority (67%) was censored because data collection ended.

Another way to examine whether censoring appears non-informative is to examine if covariates predict whether students left the district before data collection ended. If there is no relationship between covariates and attrition, this serves as additional evidence that censoring is non-informative. If there is a significant relationship between particular covariates and attrition, analysis of the direction of these relationships can provide information about the direction of potential bias. Results of a discrete-time hazard model predicting attrition suggest that significant relationships between several covariates and attrition exist. Students with higher CELDT scores are significantly less likely to leave the district before the end of data collection, suggesting that estimates of time to reclassification might contain positive bias (because students with higher CELDT scores are more likely to be reclassified). However, students in special education and students receiving free or reduced-price lunch are also significantly less likely to leave the district. Because students in special education and students receiving free or reduced-price lunch

are less likely to be reclassified (Table 3), this suggests negative bias in estimates of time to reclassification. Taken together, although we cannot ignore possible bias in estimates of time to reclassification caused by censoring, analysis suggests that these estimates contain both positive and negative bias (full results available from the author).

The second key assumption for survival analysis is the proportionality assumption, which stipulates that the effect of a predictor on the outcome is the same in every time period. For example, gender is a predictor in our model. Under the proportionality assumption, we assume that being female has the same impact on a student's likelihood of reclassification in the first year a student is enrolled in the district as in all subsequent years. Guided by theory and prior research, we can relax the proportionality assumption by including interactions between time and specific predictors. The main model above (Equation 3) includes three interactions between predictors and time. First, an interaction between students' initial academic English proficiency and time is included because theory suggests that although this variable might play a strong role in students' likelihood of reclassification at first, this might diminish over time. Second, for similar reasons, the model includes an interaction between time and students' initial academic L1 proficiency. An interaction between time and whether students were ever in a bilingual program is also included because both theory and prior research suggest that students in bilingual programs might be less likely to be reclassified in early years when they are receiving greater amounts of instruction in their primary languages but might be more likely to be reclassified in later years if bilingual education scaffolds students' English acquisition and content learning (Umansky & Reardon, 2014). Because prior research suggests that students' literacy knowledge in one language may affect the development of literacy in another language (e.g., Genesee et al.,

2008), the model also includes an interaction between students' initial academic English proficiency and initial academic L1 proficiency.

Research Question 3: How Long Does It Take Students Who Enter School as ELs in Kindergarten to Attain Individual Reclassification Criteria?

To explore variation in when students first attain each reclassification criteria, this study again employs discrete-time survival analysis. For this analysis, because the research question relates to timing, hazard and survivor functions from the sample data are used. Five separate models identical to the model presented in Equation 1 are run, except attaining an individual reclassification criterion serves as the outcome for each model. Separate models calculate when students first score 3 (out of 5) on the Listening and Speaking, Reading, and Writing sections of the CELDT, plus when students first score Basic or above on the CST-ELA.

Results

Research Question 1: How Long Does It Take for Students Entering School as ELs in Kindergarten to Be Reclassified as Proficient in English?

A useful tool for understanding how likely events are to occur over time is the life table (Table 2). This table displays the number of students present in the dataset at the beginning of each year who had not yet been reclassified (column A), the number of students reclassified during that year (column B), and the number of students who left the district by the end of that year (column C). Using these basic counts, the life table also displays information about the hazard of reclassification in each year (column D), the cumulative probability that students were not reclassified (i.e., the survivor probability, column E), the cumulative probability of reclassification (column F), and a confidence interval for the cumulative probability of reclassification (column G).

Examining the fluctuations of hazard numerically in column D and visually in Figure 1 reveals that students' likelihood of reclassification rises steadily during elementary school, peaking after students have spent 6 years in the district (when the vast majority of students are in fifth grade). After this peak at the end of elementary school, students' likelihood of reclassification drops. Thus, there appears to be a *reclassification window* during the upper elementary grades. Students not reclassified by this point in time become less likely ever to do so.

In looking at students' cumulative probability of reclassification (listed in column F and illustrated in Figure 2), we see that after 9 years in the district, students have a 74% likelihood of being reclassified. This is somewhat higher than the 62% of Latino ELs reclassified by the end of middle school in Umansky and Reardon's (2014) recent California-based study. This difference is likely due to the fact that the district in Umansky and Reardon's (2014) study required students to attain a score of 325 or above on the CST-ELA to be eligible for reclassification, while LAUSD required a score of 300 or above on this assessment. Again, the particular reclassification criteria that education agencies employ have important consequences for how long it takes students to be reclassified.

INSERT TABLE 2 AND FIGURES 1 AND 2 HERE

Research Question 2: What factors are related to variation in time to reclassification for students entering as English learners in kindergarten?

Table 3 reports regression coefficients for relationships between predictors and likelihood of reclassification. To facilitate interpretation, these estimates are presented as odds ratios (ORs). An OR of 1 indicates that two groups have the same probability of experiencing reclassification at each time point. ORs greater than 1 indicate that a particular group is more likely to

experience an event, while ORs less than 1 indicate that a particular group is less likely to experience the event. When ORs are greater than 1, the difference between the OR and 1 is equal to the difference in the likelihood of reclassification for the two groups. Results indicate that at each time point, girls are 16.2% more likely than boys to be reclassified as proficient in English (as shown by the OR of 1.162). For home language, the reference category is Spanish. We see that students who speak languages other than Spanish are more likely to be reclassified, as evidenced by the fact that the ORs for each of the indicator variables for home languages in the model are greater than one. Speakers of Cantonese, Korean, and Filipino are about twice as likely as Spanish speakers to be reclassified at each time point, consistent with Slama's (2014) recent results in Massachusetts. In addition, a one-unit difference in parent education (on a scale of 1 to 5, with 1 = *not a high school graduate* and 5 = *some graduate school*) is associated with an 8.6% greater likelihood of reclassification.

When ORs are less than 1, the ORs are most easily understood by calculating the reciprocal of the original OR. Using this method, we see that students who do not qualify for special education are almost five times more likely to be reclassified than their peers in special education, after controlling for other factors (1/0.215 = 4.651). In addition, we see that students who do not qualify for free/reduced-price lunch are about 20% (1/0.840 = 1.190) more likely than their peers to be reclassified.

INSERT TABLE 3 HERE

Findings about the role of participation in a bilingual program, the role of initial academic L1 proficiency, and the role of initial academic English proficiency merit further discussion. As noted earlier, models include interactions for these variables. Therefore, the coefficients on these variables must be considered in combination with the coefficients on

the interaction terms. For example, the fact that the coefficient on the main effect for ever having been in a bilingual program is less than one suggests that students ever in bilingual programs are less likely than their peers to be reclassified. However, the fact that the coefficient on the interaction between whether students were ever in a bilingual program and time is greater than one suggests that students ever in bilingual programs become increasingly likely to be reclassified in later years. These results are most easily interpreted visually (Figure 3).[3] We see that, as the theoretical design of bilingual programs would predict, students ever in bilingual programs are initially less likely than their peers to be reclassified, but this shifts over time. After controlling for other factors, students ever in a bilingual program have no statistically significant difference in cumulative likelihood of reclassification after 9 years in the district. Importantly, in this sample, the mean number of years students ever in a bilingual program participated in such a program was 2.6 years, and existing data do not allow for disaggregation of outcomes for students participating in different types of bilingual programs. These results should not in any way be interpreted as causal assessments of bilingual programs' effectiveness. Rather, they simply suggest that in this district during these years, students who received a small amount of L1 instruction had a cumulative likelihood of reclassification after 9 years that was indistinguishable from their peers.

Additional graphs show fitted cumulative probabilities by students' initial academic L1 proficiency (Figure 4), by initial academic English proficiency (Figure 5), and by differing levels of the two variables (Figure 6). Students who enter kindergarten with high levels of academic L1 proficiency are approximately 12 percentage points more likely to be reclassified than those who enter with beginning levels of academic L1 proficiency,

controlling for other factors, including initial academic English proficiency. Meanwhile, students with lower levels of initial academic English proficiency are much less likely to be reclassified in earlier years, but this gap closes somewhat over time. After 9 years in the district, students who enter kindergarten with high levels of academic English proficiency are approximately 13 percentage points more likely to be reclassified than students who enter with beginning levels of academic English proficiency, again after controlling for other factors.[4]

Figure 6 illustrates how students' initial academic L1 proficiency and initial academic English proficiency interact to produce dramatic differences in their likelihood of reclassification. Students who enter kindergarten with high levels of academic L1 proficiency and high levels of academic English proficiency are approximately 24% more likely to be reclassified after 9 years in the district than students who enter with beginning levels of academic L1 proficiency and beginning levels of academic English proficiency (comparing the top and bottom lines in Figure 6). Meanwhile, students who enter kindergarten with high levels of academic L1 proficiency but beginning levels of academic English proficiency are slightly less likely to be reclassified in the early years than their peers who enter with beginning levels of academic L1 proficiency but high levels of academic English proficiency. However, after 9 years in the district, this gap closes, with the two groups having statistically indistinguishable likelihoods of reclassification (comparing the middle two lines in Figure 6). Again, these estimates are descriptive rather than causal.

INSERT FIGURES 3, 4, 5, AND 6 HERE

Research Question 3: How long does it take students who enter school as English learners in kindergarten to attain individual reclassification criteria?

As noted earlier, to be reclassified as proficient in English, students must meet a variety of reclassification criteria. Key criteria are (a) having an overall score of Proficient on the CELDT, (b) having scores of Intermediate or higher on each of the CEDLT sections (Listening, Speaking, Reading, and Writing), and (c) scoring Basic (300) or above on the CST-ELA. (Students also must have minimum grades in ELA, but data about students' grades are not available, so that criterion is not considered in this analysis. Analysis showed that 9% of students who were never reclassified met all the CST-ELA and CELDT criteria during a particular year but were never reclassified.)

To understand the timing for attaining individual reclassification criteria, discrete-time survival analysis is again used to calculate hazard and survivor functions for each reclassification criterion. For this analysis, because the research question focuses on timing, hazard and survivor functions from the sample data, rather than from fitted models, are displayed. Figure 7 displays the cumulative probabilities of meeting each of the reclassification criteria. When the CELDT was rescaled in 2006-2007, cut scores on the CELDT also changed. When considering how long it took students to attain particular reclassification criteria for Research Question 3, this study considers students to have attained a particular criterion in a particular year if their period-specific CELDT scale score fell at or above the cut score that was used to determine reclassification eligibility during that particular year (i.e., the period-specific cut score). Additional analyses using the cut scores in place after the CELDT was rescaled were conducted. In all cases, all conclusions presented here about the number of years it took students to attain particular criteria held across both analysis methods. This is largely because students did not typically attain most of these criteria until 4 or 5 years in the district, during which time all students were being held to the more difficult cut scores after the CELDT rescaling.[5]

INSERT FIGURE 7 HERE

Results indicate that students meet the Listening and Speaking criteria very early, with 90% of students attaining a score of Intermediate or above on these sections of the CELDT in their third year in the district, when the vast majority of students are in second grade.[6] Students do not take the CST or the Reading and Writing sections of the CELDT until second grade (as reflected in Figure 7). Nonetheless, students take much longer to meet these literacy-based criteria, as evidenced by the shallower slopes for the lines pertaining to the CST-ELA, CELDT Reading, CELDT Writing, and CELDT Overall criteria. Looking vertically at the lines for each criterion above the mark at Year 3, it is evident that students are least likely to meet the CELDT Reading criterion at this time point. By the time students have been in the district for 9 years, however, the CST-ELA has become the most difficult criteria for students to meet, as indicated by the fact that the line corresponding to the CST-ELA criterion is the lowest of all the lines at the Year 9 mark (excluding the line for reclassification, which is a composite measure). After Year 6, the slope for the CST-ELA line is extremely shallow, suggesting that students who have not scored Basic or above on the CST-ELA by the end of elementary school are unlikely ever to do so.

Figure 7 can be interpreted as providing lower and upper bounds on estimates of how long it takes students to attain English proficiency. Most students attain proficiency in listening/speaking after 1 to 2 years, in line with the targets laid out by Proposition 227 in California and with prior research using listening/speaking as outcomes. However, literacy in a second language appears to develop considerably more slowly, as shown by the cluster of lines in Figure 7 corresponding to cumulative probabilities for attaining proficiency on literacy-based measures (CELDT Reading, CELDT Writing, CELDT Overall, and CST-ELA). After 3 years in

the district, 86% of students have met the CELDT Listening/Speaking criterion. In contrast, after 3 years, 27% of students have met the CELDT Reading criterion, 38% have met the CELDT Writing criterion, 43% have met the CELDT Overall criterion, and 58% have met the CST-ELA criterion. The cumulative probability of reclassification lags far behind the cumulative probabilities for meeting each of the separate reclassification criteria. This lag is likely due to the fact that to be reclassified, students need to meet all the reclassification criteria in the same year.

Discussion and Implications

These findings corroborate prior research suggesting that English proficiency, when defined to encompass text-based literacy practices, does not develop quickly. A recent technical report suggested that one method for defining an expected time period for English proficiency attainment would be to determine, using survival analysis, the point at which at least 60% of students had reached proficiency targets (Cook et al., 2012). Applying this method, the time necessary for students who enter school as ELs in kindergarten to reach proficiency on literacy-based measures ranges from 4 to 7 years (4 years for the CST-ELA and CELDT Writing criteria, 5 years for the CELDT Reading and CELDT Overall criteria, and between 6 and 7 years for meeting all reclassification criteria), in line with estimates from prior cross-sectional research (Hakuta et al., 2000). Although the 60% threshold is subject to debate, basing targets on empirical data would represent a substantial advance from current practices.

This analysis also provides information about the factors associated with variation in time to reclassification. Boys, native Spanish speakers, students with lower levels of initial academic English proficiency, students with lower levels of initial academic L1 proficiency, students in special education, and students whose parents have lower levels of education all have lower probabilities of reclassification than their peers, controlling for the other factors. Differences in

cumulative probabilities of reclassification vary quite dramatically along many of these dimensions. For example, estimates suggest that students who enter the district with beginning levels of initial academic L1 proficiency and beginning levels of academic English proficiency are 24% less likely to reclassify after 9 years than their peers who enter with high levels of both.

Because we see that students' academic L1 proficiency and academic English proficiency at kindergarten entry are associated with large differences in students' likelihood of reclassification by the end of middle school, expected time frames for attaining English proficiency should take these key variables into account (Hopkins, Thompson, Linquanti, Hakuta, & August, 2013). In addition, the low likelihood of reclassification for students who enter kindergarten with beginning levels of academic language proficiency in English and in their home language suggests a need for interventions targeting this group of students, building on their existing literacy practices. Furthermore, although these findings are descriptive rather than causal, they are consistent with other evidence suggesting a need to expand access to high-quality preschool programs in students' primary language, in English, or in both languages (e.g., Cooper & Costa, 2012; Espinosa, 2013).

In addition, the relatively abrupt closing of the *reclassification window* that occurs at the end of elementary school has important implications, as well. Content-area assessments represent the largest barrier to reclassification for ELs at the secondary level. The low probability that ELs will attain content-area assessment targets for the first time in middle school may indicate that students are not receiving appropriate curriculum, instruction, and support and suggest that students who remain ELs in middle school may need additional enrichment services and increased access to rigorous core curriculum (e.g., Estrada, 2014). However, this finding may also raise questions about the inclusion of content-area assessment scores as part of the

reclassification criteria, a topic that could be explored in future research by comparing outcomes for students in states with differing policies.

This study also has implications for the role of primary language assessments. Because current assessment and accountability systems do not require assessment of ELs in their home languages, few districts systematically conduct such assessments. (This likely explains why no prior large-scale longitudinal study has examined the role of initial academic L1 proficiency on time to reclassification.) Furthermore, both researchers and practitioners have raised serious questions about the validity and reliability of existing primary language assessments (e.g., MacSwan, 2005; MacSwan et al., 2002). These concerns led LAUSD to suspend primary language assessments in its new Master Plan for English Learners (Los Angeles Unified School District, 2012). However, both theory and practice, as well as results from this study, suggest that a student's academic language and literacy skills in her home language play an important role in her acquisition of a second language. Thus, assessments of students' academic language and literacy knowledge in their primary language may have a useful role to play in next-generation assessment and accountability systems—if we frame the results not as providing information about what students lack but as providing information about the resources students bring to the classroom.

Finally, we must remember that after 9 years in LAUSD, one-fourth of students had not yet been reclassified as proficient in English. As this analysis suggests, these students are much more likely to qualify for special education than their peers. More than 30% of LAUSD students who entered the district as ELs in kindergarten and remained classified as ELs 9 years later qualified for special education. Students who remained ELs 9 years after entering the district were also more likely to be male, qualify for free or reduced-price lunch, have parents with lower

levels of education, speak Spanish as their primary language, and have entered school with lower levels of initial academic language proficiency in their primary language and in English. With the implementation of new, more challenging content-area and ELP assessments aligned to the Common Core State Standards, attaining reclassification criteria may become even more challenging. Thus, there is a pressing need for curriculum and professional development targeted toward supporting ELs—and all students—in developing the academic language necessary for success on next-generation assessments.

Accountability targets for how long it should take ELs to attain proficiency in English have too often been based on wishful thinking rather than empirical data. In addition, conversations about the timing of reclassification have sometimes failed to examine differences in the reclassification criteria that education agencies employ. A major overhaul of assessment and accountability systems is now underway across the nation as states implement the Common Core State Standards, along with ELP assessments and content-area assessments aligned to these new standards. As policymakers establish expected time frames for English proficiency under the new assessment and accountability systems, ongoing examination of longitudinal data pertaining to ELs must be a more integral part of the policy-making process.

References

Abedi, J. (2008). Classification system for English language learners: Issues and recommendations. *Educational Measurement: Issues and Practice*, *27*(3), 17–31.

August, D., & Shanahan, T. (Eds.) (2006). *Developing literacy in second-language learners*. Mahwah, NJ: Erlbaum.

Birdsong, D. (1999). *Second language acquisition and the critical period hypothesis.* Lawrence Erlbaum Associates, Mahwah, NJ.

Boyle, A., Taylor, J., Hurlburt, S., & Soga, K. (2010). *Title III: Behind the numbers*. Washington DC: U.S. Department of Education.

California Department of Education. (2007). Technical report for the California English Language Development Test (CELDT) 2006-07 edition. Retrieved from http://www.cde.ca.gov/ta/tg/el/documents/formftechreport.pdf

California Department of Education. (2013). California English Language Development Test (CELDT): 2012-13 CELDT information guide. Retrieved from http://www.cde.ca.gov/ta/tg/el/documents/celdt13guide.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.

Conger, D. (2010). Does bilingual education interfere with English-language acquisition? *Social Science Quarterly, 91*(4), 1103-1122.

Cook, G., Boals, T., Wilmes, C., & Santos, M. (2007). *Issues in the development of Annual Measurable Achievement Objectives (AMAOs) for WIDA consortium states.* Madison, WI: Wisconsin Center for Education Research.

Cook, G., Linquanti, R., Chinen, M., & Jung, H. (2012). *National evaluation of Title III implementation supplemental report: Exploring approaches to setting English Language Proficiency performance criteria and monitoring English learner progress*. Washington, DC: U.S. Department of Education.

Cooper, D. & Costa, K. (2012). Increasing the effectiveness and efficiency of existing public investments in early childhood education: Recommendations to boost program outcomes and efficiency. Washington, DC: Center for American Progress. Retrieved from http://www.americanprogress.org/wp-content/uploads/issues/2012/06/pdf/earlychildhood.pdf

Espinosa, L. (2013). Pre-K-3rd: Challenging common myths about dual-language learners – An update to the seminal 2008 report. New York: Foundation for Child Development. Retrieved from http://fcd-us.org/sites/default/files/Challenging%20Common%20Myths%20Update.pdf

Estrada, P. (2014). English learner curricular streams in four middle schools: Triage in the trenches. *The Urban Review, 46*(5). Retrieved from http://link.springer.com/article/10.1007/s11256-014-0276-7

Fry, R. (2007). "How far behind in math and reading are English language learners?" Washington, DC: Pew Hispanic Center.

Gándara, P., Maxwell-Jolly, J., & Rumberger, R. W. (2008). Resource needs for English learners: Getting down to policy recommendations. UC Santa Barbara: Linguistic Minority Research Institute.

Genesee, F., Geva, E., Dressler, C., & Kamil, M. (2008). Cross-linguistic relationships in second-language learners. In D. August & T. Shanahan (Eds.), *Developing reading and writing in second-language learners: Lessons from the report of the National Literacy Panel on language-minority children and youth* (pp. 61-93). New York: Routledge.

Ghali, W. A., Quan, H., Brant, R., van Melle, G., Norris, C. M., Faris, P. D., … & Knudtson, M. L. (2001). Comparison of 2 methods for calculating adjusted survival curves from proportional hazards models. *Journal of the American Medical Association, 286*(12), 1494-1497.

Greene, J.P. (1997). A meta-analysis of the Rossell & Baker review of bilingual education research. *Bilingual Research Journal, 21* (2/3).

Grissom, J. B. (2004). Reclassification of English learners. *Education Policy Analysis Archives, 12*(36). Retrieved from http://epaa.asu.edu/epaa/v12n36/

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

Hill, E. (2004). A look at the progress of English learner students. Sacramento, CA: Legislative Analyst's Office. Retrieved from http://www.lao.ca.gov/2004/english_learners/021204_english_learners.htm

Hill, L. E., Weston, M. & Hayes, J. (2014). Reclassification of English learner students in California. San Francisco, CA: Public Policy Institute of California. Retrieved from http://www.ppic.org/content/pubs/report/R_114LHR.pdf

Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

Lau v. Nichols, 414 U.S. 563 (1974).

Linquanti, R., & Cook, G. (2013). Toward a "common definition of English learner": Guidance for states and state assessment consortia in defining and addressing policy and technical issues and options. Washington DC: Council of Chief State School Officers. Retrieved from http://www.ccsso.org/Documents/2013/Toward_a_Common_Definition_2013.pdf

MacSwan, J., Rolstad, K., & Glass, G. V. (2002). Do some school-age children have no language? Some problems of construct validity in the Pre-LAS Español. *Bilingual Research Journal, 26*(2), 213-238.

MacSwan, J. (2005). The "non-non" crisis and academic bias in native language assessment of linguistic minorities. In Cohen, J., McAlister, K., Rolstad, K., & MacSwan, J. (Eds.) ISB4: Proceedings of the 4th International Symposium on Bilingualism (pp. 340-375). Somerville, MA: Cascadilla Press.

Parrish, T., Perez, M., Merickel, A., & Linquanti, R. (2006). Effects of the implementation of Proposition 227 on the education of English learners, K-12: Findings from a five-year evaluation. Palo Alto, CA: American Institutes for Research.

Ryan, C. (2013). Language use in the United States: 2011. Washington, DC: U.S. Census Bureau. Retrieved from: http://www.census.gov/prod/2013pubs/acs-22.pdf

Sálazar, J. J. (2007). *Master Plan evaluation report for English learners—2005/2006*. Los Angeles: LAUSD Program Evaluation and Assessment Branch.

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.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction. *Review of Educational Research, 75* (2), 247-284.

Slavin, R. E., Madden, N., Calderón, M., Chamberlain, A., & Hennessy, M. (2011). Reading and language outcomes of a multiyear randomized evaluation of transitional bilingual education. *Educational Evaluation and Policy Analysis 33*(1), 47-58.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Umansky, I. M. & Reardon, S. F. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal 51*(5), 879-912.

U.S. Department of Education, National Center for Education Statistics. (2010). *The condition of education 2010* (NCES 2010–028).

Van Buuren, S., Bushuizen, H. C, & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine*, 18, 681-694.

Wolf, M. K., Kao, J. C., Griffin, N., Herman, J. L., Bachman, P. L., Chang, S., & Farnsworth, T. (2008). Issues in assessing English language learners: English language proficiency measures and accommodation uses. Practice review. CRESST Report 732. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Willig, A. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research, 55*(3), 269-317.

Notes

[1] Results from these rigorous California-based empirical studies are broadly consistent with findings from other California studies that either followed students for shorter time periods or did not employ student-level longitudinal data (Grissom, 2004; Hill, 2004; Sálazar, 2007).

[2] To further control for school-level factors in time to reclassification, models employing school fixed effects were also used. Parameter estimates and significance levels remained very similar with the inclusion of school fixed effects. (Full results available from the author.)

[3] The best method for graphing survivor functions to show the role of particular predictors in fitted models has been debated in the literature. Although many researchers have created graphs by using the mean-values of all other predictors besides the predictor of interest, this method has serious flaws (e.g., Ghali, et al. 2001). Rather than using mean values, I use a method created by Sean Reardon (personal communication, May 21, 2013), using actual values for all variables except for the predictor of interest. Each student is imagined to have remained in the dataset for the maximum number of years. Survival curves are then estimated for each individual represented in the dataset, and then these survival curves are averaged to generate

mean survival probabilities for each time period. This procedure is repeated for each level of the predictor of interest. The graphs shown here (Figures 3, 4, 5, and 6) all show estimated results for students who entered LAUSD in 2001-02. Because of large, positive cohort effects, estimates of the cumulative likelihood of reclassification for later cohorts would be higher. However, the difference between the various levels of predictors would remain similar. A full description of this graphing method and associated results are available from the author.

[4] The highest CELDT score students can attain at kindergarten entry and still be considered an EL is Level 3 (Intermediate). Therefore, I consider ELs who score Level 3 on the CELDT at kindergarten entry to have high levels of academic English proficiency relative to their peers.

[5] These additional analyses are available from the author by request.

[6] In 2006-07, Listening and Speaking became separate domains on the CELDT. Following this change, I consider students to have met the Listening/Speaking criteria if they scored at Level 3 (out of 5) or above on both Listening and Speaking.

Table 1. Mean characteristics of analytic sample by cohort.

	Cohort								
	All	2002	2003	2004	2005	2006	2007	2008	2009
Female	0.492	0.492	0.492	0.491	0.486	0.491	0.496	0.494	0.492
Home language									
Spanish	0.942	0.948	0.948	0.948	0.943	0.946	0.939	0.934	0.927
Korean	0.012	0.010	0.009	0.010	0.011	0.012	0.013	0.015	0.015
Armenian	0.011	0.010	0.010	0.010	0.011	0.010	0.010	0.012	0.012
Filipino	0.007	0.006	0.007	0.007	0.007	0.006	0.007	0.007	0.008
Cantonese	0.005	0.006	0.006	0.005	0.006	0.005	0.004	0.006	0.006
Other Language	0.024	0.020	0.020	0.021	0.022	0.022	0.027	0.028	0.032
Free/reduced-price lunch ever	0.949	0.981	0.983	0.979	0.978	0.966	0.944	0.893	0.852
Bilingual program ever	0.109	0.109	0.123	0.132	0.126	0.116	0.097	0.088	0.081
Special education ever	0.119	0.117	0.128	0.135	0.138	0.132	0.119	0.100	0.079
Parent education level									
(1 = Not hs grad, 5 = Grad school)	1.794	1.740	1.777	1.755	1.766	1.785	1.812	1.839	1.870
Initial L1 proficiency									
(1=Beginning, 4=Proficient)	2.118	2.279	2.165	2.131	2.087	2.077	2.042	2.046	1.996
Equated initial English proficiency									
(1=Beginning, 5=Advanced)	1.628	1.434	1.487	1.527	1.552	1.595	1.762	1.848	1.865
Equated initial CELDT score (200-600)	331.697	319.001	328.464	322.621	323.778	326.764	336.145	348.617	350.992
Number of students	202,931	29,608	27,449	24,058	24,363	24,272	25,941	24,183	23,057

Note: CELDT=California English Language Development Test. The CELDT was rescaled in 2006-07. To provide equated CELDT proficiency levels and scale scores during the entire time period under study, initial English proficiency levels and initial CELDT scale scores from the pre-2007 period were converted to the "common scale" (i.e., new CELDT scale) using the concordance tables in the CELDT 2007 Technical Manual (Appendix O).

Table 2. Life table for reclassification of students entering LAUSD as ELs in K, using data from 2001-02 through 2009-10.

	A	B	C	D	E	F	G	
Year	Beginning Total	Reclassified	Left District	Hazard	Survival	Cumulative Likelihood of Reclassification	Confidence Interval	
1	202,931	5	8,740	0.0000	1.0000	0.0000	0.0000	0.0001
2	194,186	9,085	29,055	0.0468	0.9532	0.0468	0.0459	0.0478
3	156,046	10,146	24,790	0.0650	0.8912	0.1088	0.1073	0.1103
4	121,110	18,881	19,027	0.1559	0.7523	0.2477	0.2455	0.2499
5	83,202	14,318	14,125	0.1721	0.6228	0.3772	0.3745	0.3798
6	54,579	16,000	10,905	0.2922	0.4408	0.5592	0.5561	0.5622
7	27,854	5,297	6,755	0.1902	0.3570	0.6430	0.6398	0.6462
8	15,802	2,435	6,224	0.1541	0.3020	0.6980	0.6946	0.7014
9	7,143	1,046	6,097	0.1464	0.2578	0.7422	0.7384	0.7460

Note: In Year 9, the students in Column C were considered censored because they had not been reclassified by the time data collection ended.

Table 3. Results for logistic regression models predicting reclassification hazard for students entering LAUSD as ELs in K, using data from 2001-02 through 2009-10.

	Odds ratios b/se
Female	1.162***
	[0.010]
Korean	1.978***
	[0.246]
Armenian	1.053
	[0.087]
Filipino	1.935***
	[0.120]
Cantonese	2.152***
	[0.241]
Other language	1.750***
	[0.077]
Parent education	1.086***
	[0.007]
Special education ever	0.215***
	[0.005]
Free/reduced-price lunch ever	0.840**
	[0.047]
Initial English proficiency	2.567***
	[0.059]
Initial English proficiency * time	0.885***
	[0.004]
Initial L1 proficiency	1.335***
	[0.021]
Initial L1 proficiency * time	0.976***
	[0.004]
Bilingual program	0.615*
	[0.126]
Bilingual program * time	1.086*
	[0.042]
Initial L1 Proficiency * initial English proficiency	0.931***
	[0.005]
N	851,648

Notes. *p<0.05; **p<0.01; ***p<0.001. Cluster-robust SEs in brackets below estimates, clustered by school. Parameter estimates for indicator variables for years and cohorts are omitted. The coefficients on the indicator variables for years suggest that hazard rises through Year 6 and then falls, as in the baseline hazard model. The coefficients on the indicator variables for cohort indicate that students in more recent cohorts are more likely to be reclassified. For home language, the reference category is Spanish. Initial English proficiency scores have been equated to address changes in the CELDT scale as of 2006-07. CELDT scores from the pre-2007 period were converted to the "common scale" (i.e., new CELDT scale) using the concordance tables in CELDT 2007 Technical Manual (Appendix O).

Figure Captions

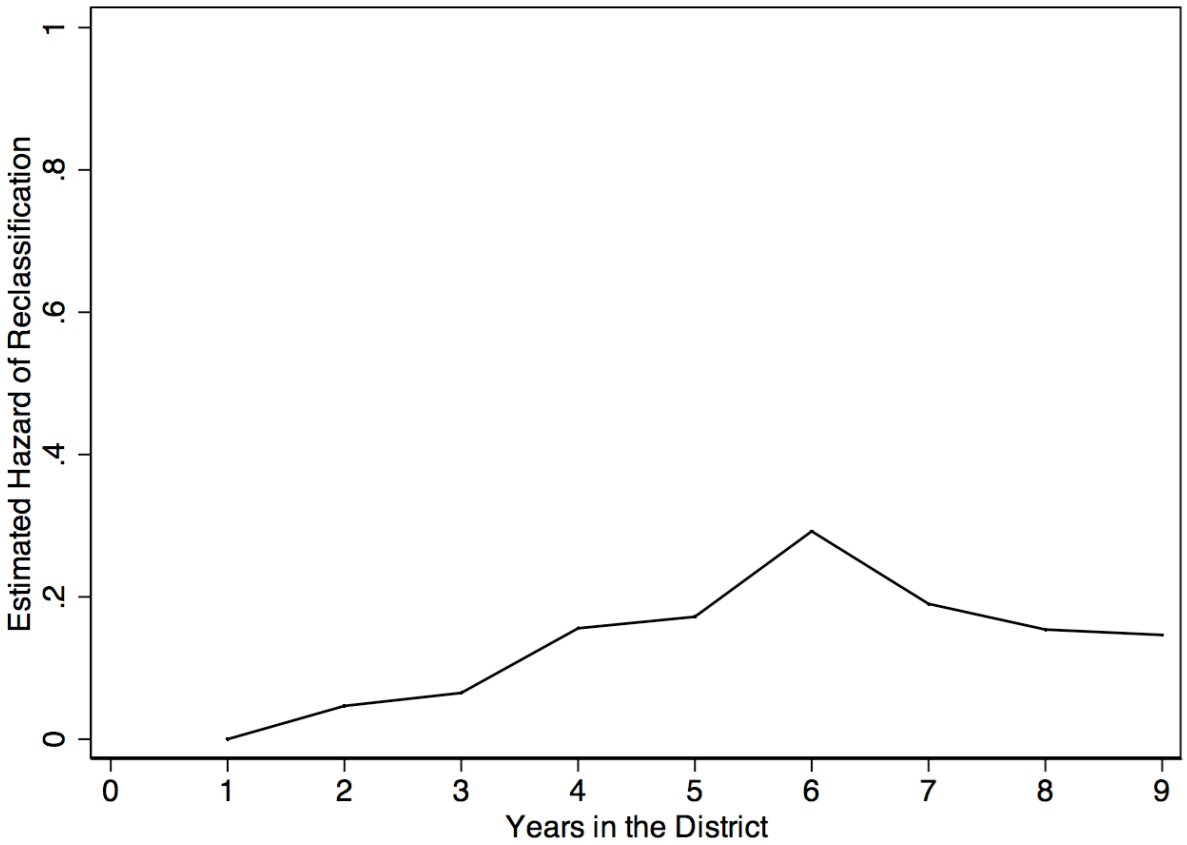

Figure 1. Hazard of reclassification for students entering LAUSD as ELs in K, using data from 2001-02 through 2009-10.

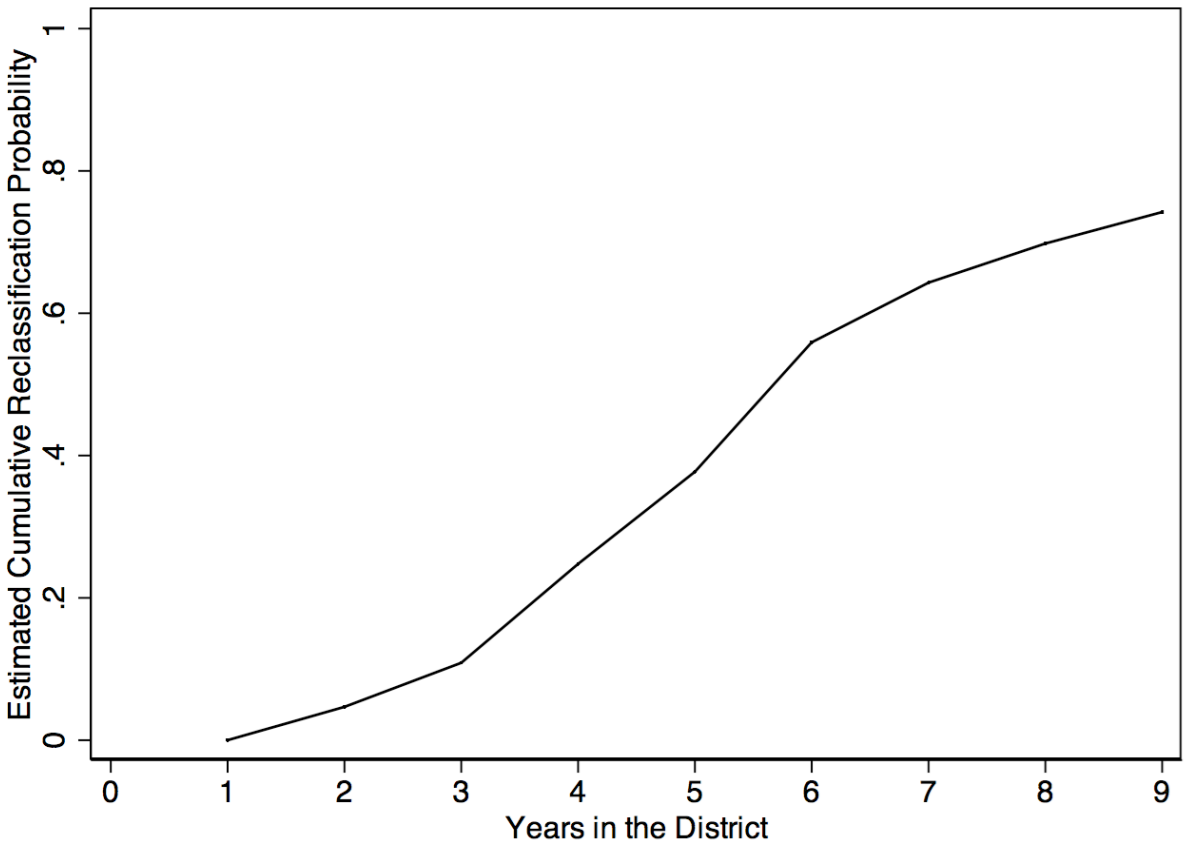

Figure 2. Cumulative probability of reclassification for students entering LAUSD as ELs in K using data from 2001-02 through 2009-10.

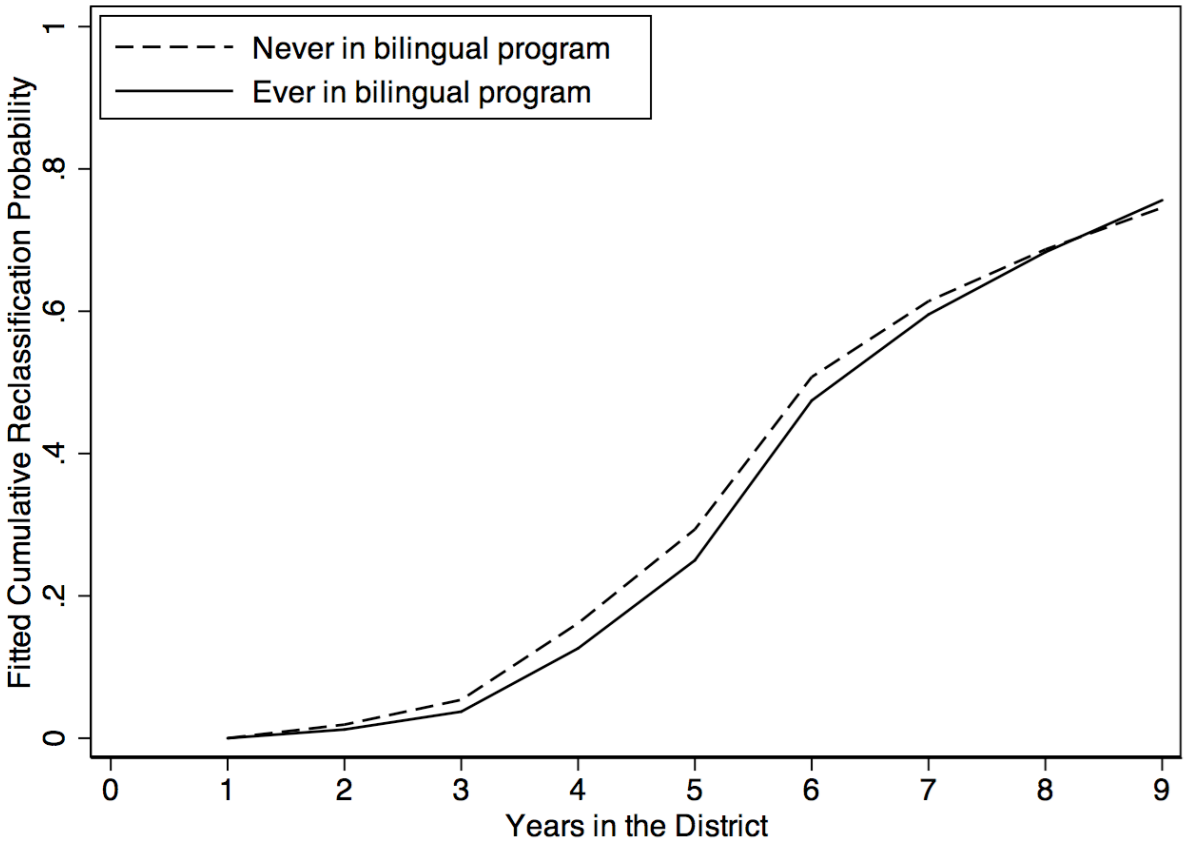

Figure 3. Fitted cumulative probability of reclassification for students entering LAUSD as ELs in K, by participation in a bilingual program.

Note. The graph shows fitted results for students who entered the district in 2001-02.

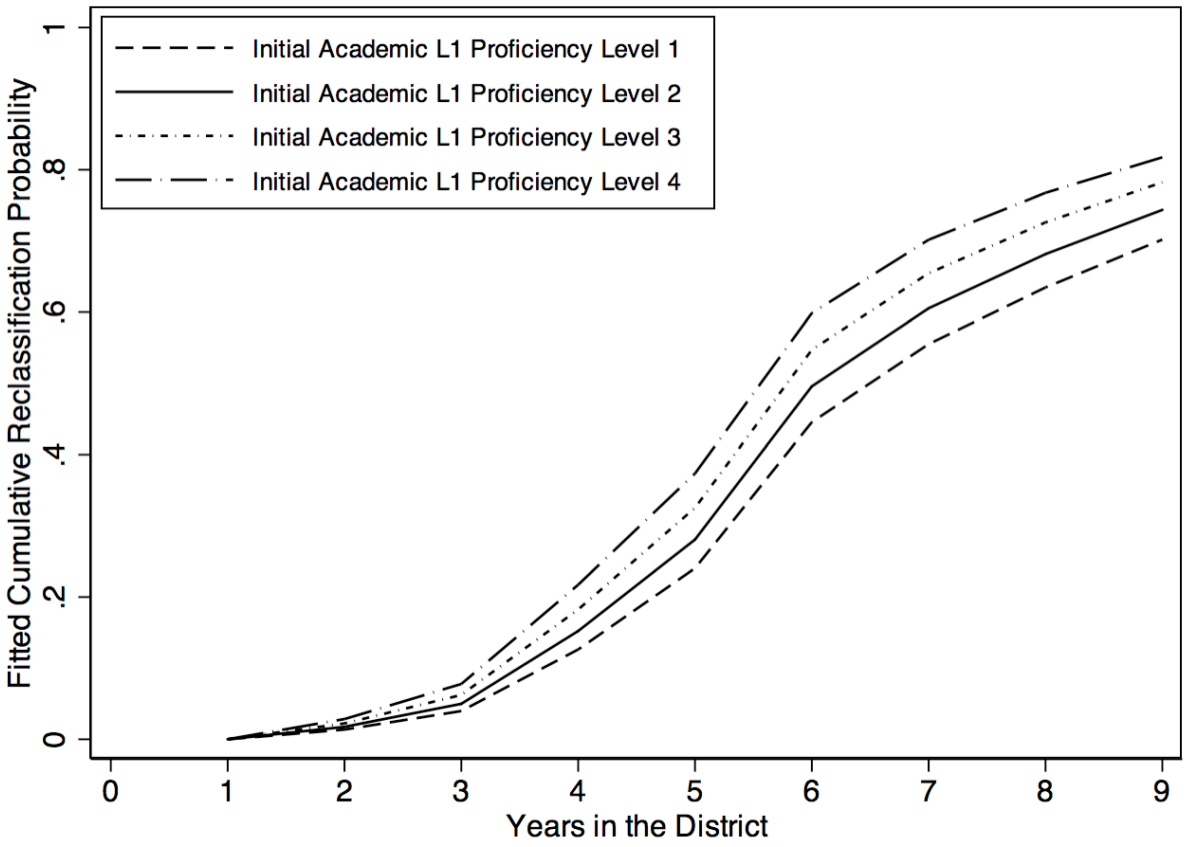

Figure 4. Fitted cumulative probabilities of reclassification for students entering LAUSD as ELs in K, by initial academic L1 proficiency level.

Note. The graph shows fitted results for students who entered the district in 2001-02.

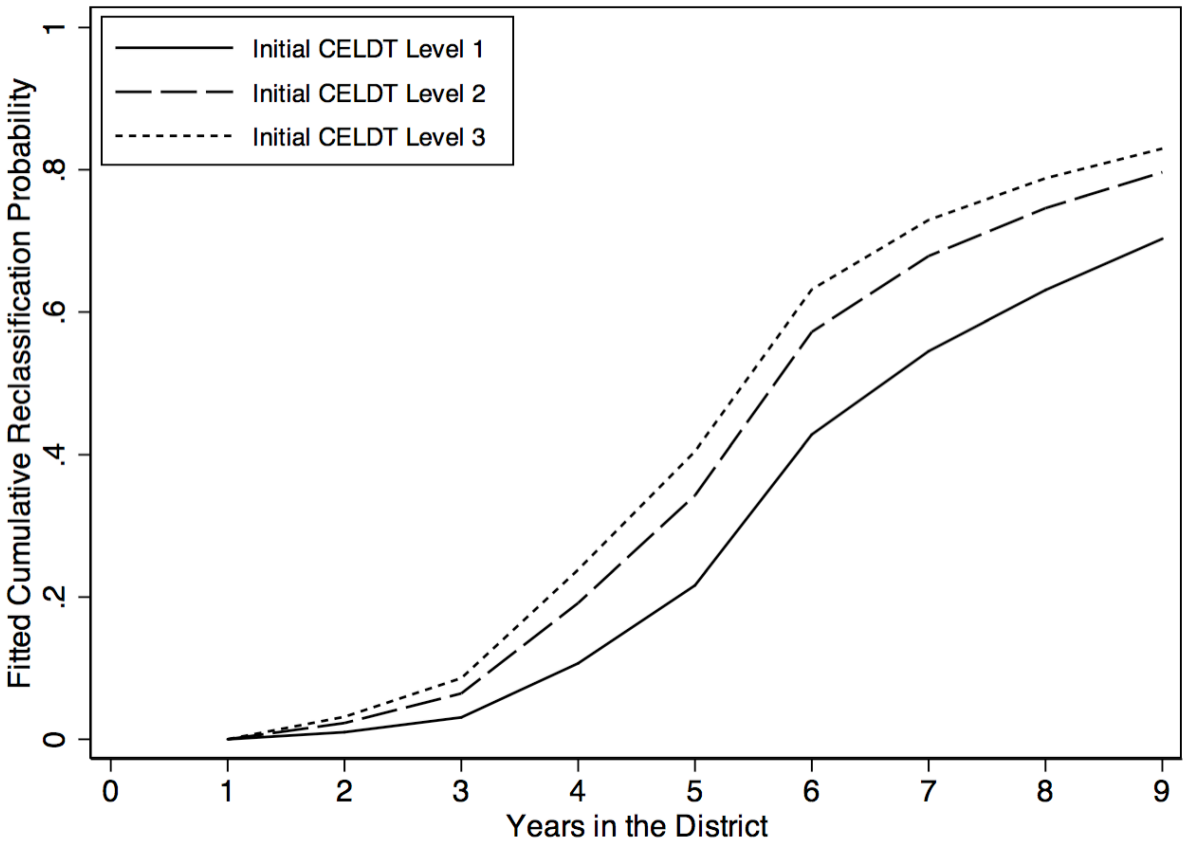

Figure 5. Fitted cumulative probability of reclassification for students entering LAUSD as ELs in K, by initial academic English proficiency level.

Note. CELDT=California English Language Development Test. The graph shows fitted results for students who entered the district in 2001-02.

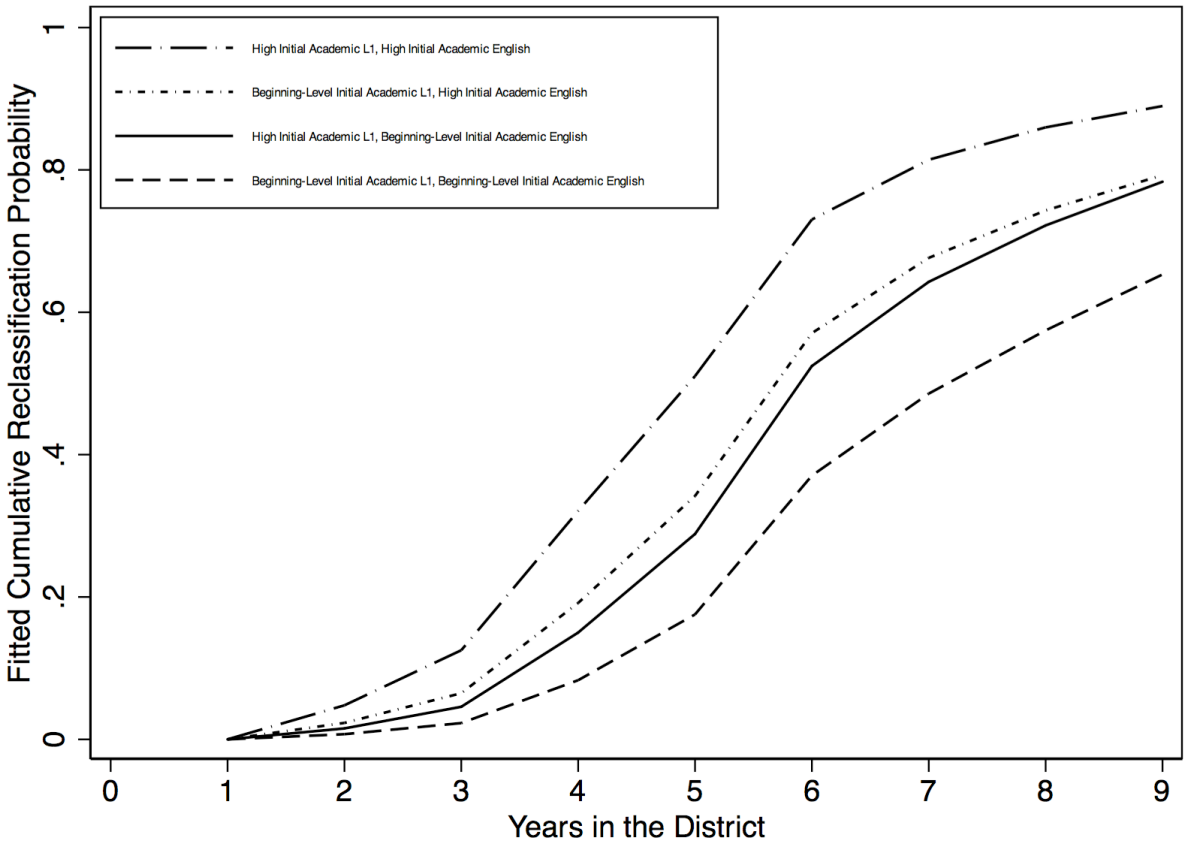

Figure 6. Fitted cumulative probability of reclassification for students entering LAUSD as ELs in K, by initial level of academic L1 proficiency and initial academic English proficiency.

Note. The graph shows fitted results for students who entered the district in 2001-02.

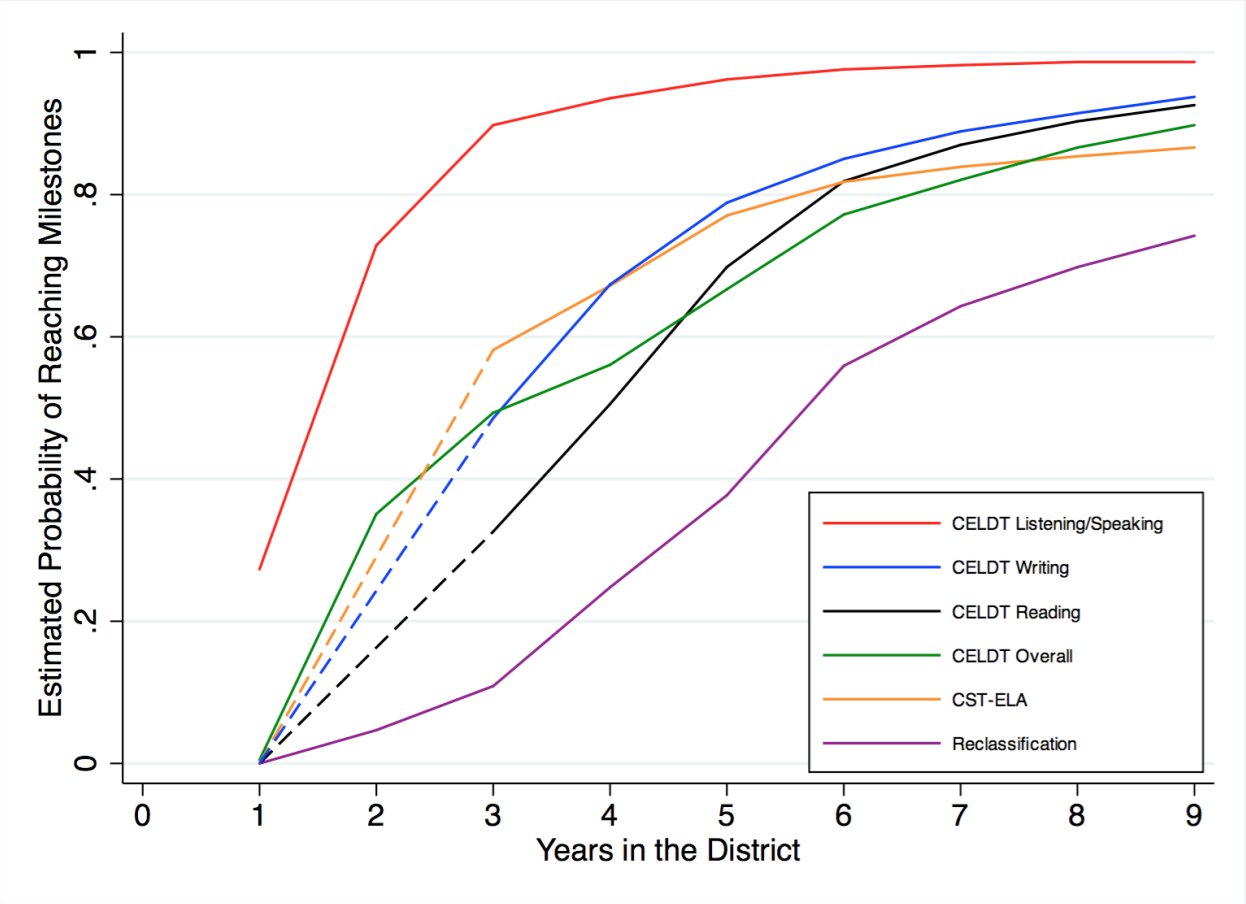

Figure 7. Cumulative probability of reaching English proficiency milestones for students entering LAUSD as ELs in K.

Note. CELDT=California English Language Development Test. CST-ELA=California Standards Test in English Language Arts. Lines for the CELDT Reading, CELDT Writing, and CST-ELA tests between Years 1 and 3 are dotted because these tests were administered for the first time in second grade (during students' third year in the district).