Making Connections

Patterns of English learner student reclassification in New York City public schools

Michael J. Kieffer
New York University

Caroline E. Parker
Education Development Center, Inc.

Key findings

Approximately half the students who entered kindergarten in New York City public schools as English learner students were reclassified as former English learner students within four years (by the end of their expected grade 3 year). Median times to reclassification differed by grade of entry, initial English proficiency, and disability status. English learner students who entered in grade 6 or 7 took a year longer to become reclassified. English learner students with below average initial English proficiency took two years longer to become reclassified than did English learner students with above average initial English proficiency. English learner students with specific learning disabilities took four years longer to become reclassified than did students without such disabilities, and students with speech or language impairments took two years longer to become reclassified than students without these impairments.

ies **NATIONAL CENTER** FOR **EDUCATION EVALUATION** AND **REGIONAL ASSISTANCE**
Institute of Education Sciences
U.S. Department of Education

REL **NORTHEAST & ISLANDS**
Regional Educational Laboratory
At Education Development Center, Inc.

Summary

With the rapid growth in the number of English learner students served by schools in the United States, educators are increasingly concerned with how these students progress toward proficiency in English. The large and diverse English learner student population in New York City public schools, where more than 41 percent of students speak a language other than English at home, provides a unique opportunity to investigate how long it takes for English learner students to become reclassified as former English learner students.

This study was conducted at the request of the English Language Learners Alliance at the Regional Educational Laboratory Northeast & Islands to better understand variation in English language acquisition patterns among English learners. It followed seven student cohorts for periods ranging from two to nine years, through the 2011/12 school year. The first student cohort followed in the study entered school in 2003/04 and was followed for nine years; the last cohort entered in 2010/11 and was followed for two years. The study drew on longitudinal administrative data in order to provide descriptive evidence on the time that English learner students take to become reclassified and in order to shed light on student characteristics that were associated with longer or shorter times to reclassification as former English learner students.

Key findings include the following:
- Slightly more than half the students who entered kindergarten in New York City public schools as English learner students were reclassified within four years (by the end of their expected grade 3 year). A quarter of students were not reclassified within six years and thus became long-term English learner students.
- The median time to reclassification for English learner students who entered school in grade 6 or 7 was about a year longer than that of English learner students who entered in kindergarten.
- The median time to reclassification was three years for English learner students with above average initial English proficiency and nearly five years for English learner students with below average initial English proficiency, among students entering at any grade.
- The median time to reclassification for English learner students with specific learning disabilities was about four years longer than that of their peers without disabilities, and the median time to reclassification for students with speech or language impairments was two years longer than that of students without these impairments, among students entering at any grade.

Contents

Why this study?

As U.S. schools become more linguistically diverse, school and district leaders across the country face new challenges and opportunities in providing effective instruction to English learner students. Under the requirements of the Every Student Succeeds Act (2015), accountability for English learner students has been integrated into the system by which all schools and districts are held accountable, making progress on English proficiency central to school and district accountability. With the increasing availability of high-quality longitudinal datasets that include both student characteristics and assessment results, it is possible to examine patterns in the reclassification of English learner students as former English learner students (see box 1 for definitions of key terms). Such analyses can improve understanding of the typical trajectory of English learner students toward reclassification and of how these patterns vary by student characteristics. A better understanding of reclassification patterns can inform policy and practice decisions, from district and state policies on criteria for reclassification to school-level decisions about which student groups may require different types or intensity of instructional support.

In New York City public schools more than 41 percent of students speak a language other than English at home, and 14 percent are classified as English learner students (New York City Department of Education, 2013). The city's English learner students speak 151 home languages and come from a variety of educational and socioeconomic backgrounds. This extraordinary diversity provides a unique opportunity to build knowledge about English learner student reclassification patterns for educators and policymakers across the region and in other large cities across the country.

This study identified patterns of English learner student reclassification from elementary school through middle and high school in New York City and examines the relationships between student characteristics and reclassification patterns over time. Conducted by Regional Educational Laboratory Northeast & Islands in collaboration with its English Language Learners Alliance, the study used longitudinal administrative data from the New York City Department of Education, which were compiled and processed by the Research Alliance for New York City Schools.

This exploratory study examined longitudinal reclassification patterns to answer two related questions about outcomes for English learner students: what is the median time to reclassification as a former English learner student, and what is the probability that a student becomes a long-term English learner student

What the study examined

This study examined longitudinal reclassification patterns to answer two important, related questions about outcomes for English learner students: what is the median time to reclassification as a former English learner student, and what is the probability that a student becomes a long-term English learner student (defined in New York City as a student who receives English learner services for six or more years without becoming reclassified; New York City Department of Education, 2013). Median time to reclassification as a former English learner student is an estimate of the number of school years that a median English learner student (or a median English learner student with certain characteristics) is expected to take to reach English proficiency as determined by reaching a predetermined cut-score on the New York State English as a Second Language Achievement Test. Although seemingly a simple descriptive question, determining the median time to reclassification is challenging in the absence of high-quality longitudinal data. The second question on the probability that a student will become a long-term English learner student under the New York City definition also relates to median time to reclassification. However, the outcome

Box 1. Key terms

Cohort. A given group of students in a longitudinal study, defined based on the years for which their data was used. This study included seven cohorts of students who entered New York City public schools as English learner students in each school year between the 2003/04 and 2010/11 school years, which the exception of 2008/09. Data for the cohort of students who entered as English learner students in the 2008/09 school year were excluded because of problems with the data quality.

Disability category. A category that indicates the type of disability with which a student has been identified. These categories are determined by the Individuals with Disabilities Act (2004). This study focuses on two disability categories for which identification procedures involve language assessments: specific learning disabilities and speech/language impairments.

English learner student. A student who comes from a home in which a language other than English is spoken and who has not yet developed sufficient English proficiency to benefit from instruction without special linguistic support, as determined by their score on an initial English proficiency test. In New York City public schools during the years studied, this test was the Language Assessment Battery–Revised (see appendix D). The study sample included 229,249 students who were initially classified as English learner students when they entered the New York City public school system based on their scores on the Language Assessment Battery–Revised. The report further distinguishes English learner students by three levels of initial English proficiency.

- *Below average initial English proficiency.* English learner students who performed substantially below the average for students entering New York City public schools as English learner students. On average, their scores are close to zero on this initial English language assessment.
- *Average initial English proficiency.* English learner students who scored close to the average score for students entering New York City public schools as English learner students.
- *Above average initial English proficiency.* English learner students who scored substantially above the average for students entering New York City public school as English learner students. On average, their scores are close to, but do not yet meet, the standard for reclassification as former English learner students.

Former English learner student. A student who was formerly classified as an English learner student but who later developed sufficient English proficiency to benefit from instruction without special linguistic support. The criteria and process by which English learner students are reclassified as former English learner students differ by state and sometimes by district within states but typically involve reaching established benchmarks on an English proficiency test based on the state's English language proficiency standards and may involve other criteria. In New York City public schools, reclassification as a former English learner student requires reaching a predetermined cutscore on the New York State English as a Second Language Achievement Test (see appendix D).

Long-term English learner student. A student who receives English learner services for six or more years without becoming reclassified as a former English learner student (New York City Department of Education, 2013).

Median time to reclassification as a former English learner student. The time by which half the English learner students no longer require special language support based on reaching a predetermined cutscore on the New York State English as a Second Language Achievement Test. It provides an estimate for the number of school years that a median English learner student (or a median English learner student with certain characteristics) is expected to take to reach English proficiency.

provides additional information about how many English learner students are struggling for a substantial period to achieve English proficiency.

Student characteristics

This study also addressed student characteristics associated with time to reclassification as a former English learner student. Specifically, the study team investigated three student characteristics that may relate to reclassification patterns:

- *Grade of entry to New York City public schools.* Although the grade level at which English learner students enter a U.S. school is likely to be important for predicting their time to reclassification, the direction of this relationship is unclear. Research supports conflicting hypotheses about whether older or younger students learn English more quickly (Conger, 2009; Gass & Selinker, 2008; Marinova-Todd, Marshall, & Snow, 2000). This study uses a student's grade of entry to a New York City public school as a proxy for a student's grade of entry to any U.S. school because data on students' grade of entry to any U.S. school were not available.
- *Initial English proficiency.* English learner students who enter school with lower initial English proficiency are likely to take longer to become reclassified (Thompson, 2015).
- *Disability category.* A recent study conducted by English Language Learners Alliance researchers in a New England district found that English learner students with disabilities had substantially lower performance on English language achievement assessments and content tests than English learner students without disabilities (Parker, O'Dwyer, & Irwin, 2014). Researchers note particular concerns around disabilities for which identification procedures involve language assessments, such as specific learning disabilities and speech or language impairments[1] (Klingner, Artiles, & Barletta, 2006; Orozco, Almanza, de Onis, Klingner, & Hoover, 2008). Nonetheless, there is little research on how disability category, particularly the categories of specific learning disabilities and speech or language impairments, relates to English learner student reclassification.

This study addressed student characteristics associated with time to reclassification as a former English learner student, specifically grade of entry to New York City public schools, initial English proficiency, and disability category

See appendix A for a more extensive review of the literature on time to reclassification.

Research questions

To better understand patterns in reclassification of English learner students in New York City public schools, this exploratory study addressed the following research questions:

- Among students who enter New York City public schools as English learner students in kindergarten, how does the probability of reclassification as former English learner students change from kindergarten through grade 7?
  - At what grade do half of these students become reclassified?
  - What is the probability that students will remain classified as English learner students for six or more years, thus becoming long-term English learner students?
- Among students who enter New York City public schools in kindergarten or after kindergarten, which of the following student characteristics are associated with time to reclassification and the probability that students will become long-term English learner students?
  - Grade of entry to New York City public schools.
  - Initial English language proficiency at school entry.
  - Disability category.

This study was designed to describe reclassification patterns, how they changed as students spend more time in New York City public schools, and how reclassification patterns differed by specific student characteristics (the data and methods are described in box 2 and discussed in more detail in appendixes B–F). Because correlational analysis was used to address these questions, causal inferences cannot be drawn from the findings. In addition, the analyses did not examine the effect of reclassification on later outcomes (such as high school achievement or graduation).

Box 2. Data and methods

Data

This study used existing administrative data for seven cohorts of students who entered New York City public schools as English learner students between the 2003/04 and 2010/11 school years. Data for the cohort of students who entered as English learner students in the 2008/09 school year were excluded because of problems with the data quality; see appendix B. The study focused on data collected after implementation of the No Child Left Behind Act began in the 2002/03 school year because district documentation and preliminary analyses suggested that this implementation led to major changes in English learner student assessment, classification, and data reporting practices.

Methods

The seven student cohorts were followed for periods ranging from two to nine years, through the 2011/12 school year. The first student cohort examined for this study entered school in 2003/04 and was followed for nine years; the last cohort entered in 2010/11 and was followed for two years. The analytic sample included 229,249 students who were initially classified as English learner students based on their scores on the Language Assessment Battery–Revised. Nearly 44 percent of these students entered New York City public schools in kindergarten, though sample sizes were large for students entering in grades 1–10 (more than 9,000 students in each grade) and adequate for students entering in grades 11 and 12 (more than 1,500 students in each grade). See appendix B for additional descriptive information on the sample.

In New York City public schools reclassification as a former English learner student during the study period occurred in the spring (and sometimes continued into the summer). Therefore, if a student was enrolled in New York City public schools but was classified as a former English learner student for the first time in the fall of a given school year, the study team indicated that she or he was reclassified by the end of the previous school year. Students from cohorts after 2003/04 contributed fewer than nine years of data. If those students were not reclassified in the years observed, they were treated as censored for the years for which they did not have data. They were neither treated as reclassified nor as continuing as English learner students, and their probability of being reclassified in each subsequent year was assumed to be the same as that for students who contributed data. In addition, if a student was classified as an English learner student in one school year and then was not enrolled in New York City public schools in subsequent years (because he or she transferred to another school system or dropped out), they were treated as censored in each subsequent year. Because students who dropped out of high school may differ from those who did not, results are reported only for students who entered in grade 7 or earlier and thus have a potential minimum of two years to be observed being reclassified before they reach high school, at which point dropping out

(continued)

4

Box 2. Data and methods *(continued)*

becomes more frequent. Results are not reported for students entering in grade 8 or later who have too few potential years observed before the effects of dropping out may substantially bias results. In addition, students who transferred to another school system may differ from those who remained. Thus, the results may not generalize to students who dropped out of school or transferred to another school system before being reclassified. See appendix C for more information.

The first research question used the subset of data for students who entered New York City public schools as English learner students in kindergarten, with the goal of comparing the probability of reclassification as it changed over grade levels through the end of grade 7. The second research question combined these data with the data on students who entered after kindergarten to facilitate comparisons in time to reclassification between students who entered at different grade levels, among other comparisons. To address the three student characteristics of interest (grade of entry, initial English proficiency, and disability category), three parallel models investigated the relationship between time to reclassification and each characteristic individually. Discrete-time survival analyses were used to estimate the probability of reclassification as it changed over time. Statistical significance was determined using robust standard errors that accounted for clustering of students within schools. See appendix C for complete details on the data-analytic approach and appendix D for psychometric information on the measures.

The results for the second research question were based on analyses of each student characteristic without accounting for the other characteristics or additional demographic characteristics. Analyses conducted to describe the extent to which results differed when taking into account the nonprimary student characteristics (grade of entry, initial English proficiency, and disability category) and other student characteristics (gender, race/ethnicity, home language, family poverty, retention in grade, and cohort) indicated that the key findings were robust to the inclusion or exclusion of these controls (see appendix F). The differences in median times to reclassification were within half a year of those discussed in the main report, and estimated percentages of students who became long-term English learner students were within approximately four percent (see tables F1–F3 in appendix F).

What the study found

This section describes the key findings related to each research question. See appendix E for the detailed results of the analysis.

Slightly more than half of students who entered kindergarten as English learner students were reclassified within four years

Of students who entered kindergarten as English learner students, 52 percent were reclassified as former English learner students by the end of their fourth year in New York City public schools, which equates to the end of grade 3 for students who were not retained in grade (figure 1 and table E1 in appendix E). In other words, the median time to reclassification for students who entered kindergarten as English learner students was approximately four years. Of students who were English learner students at the beginning of kindergarten, 11 percent were reclassified by the end of kindergarten, 29 percent by the end of expected grade 1, 41 percent by the end of grade 2, and so on. After six years 75 percent

of students who entered kindergarten as English learner students were reclassified. The remaining 25 percent are considered long-term English learner students by the New York City Department of Education. (See appendix C for details on the analytic approach used to address this question and the questions that follow concerning grade of entry, initial English proficiency, and disability category.)

The median time to reclassification was about a year longer for English learner students who entered New York City public schools in grade 6 or 7 rather than in kindergarten

The time for English learner students to become reclassified differed by the grade at which they entered New York City public schools (figure 2 and table E2 in appendix E). Estimates of the median years to reclassification varied from roughly three years (for students entering in grade 2) to more than five years (for students entering in grade 7). Overall, estimates of time to reclassification were shortest for students who entered in the primary grades but after kindergarten (grades 1–3) and longest for students who entered in grade 6 or grade 7. For students who entered in kindergarten and in the upper-elementary grades (grades 4 or 5), estimated time to reclassification ranged between these two extremes.

The percentages of students who became long-term English learner students also differed by the grade level at which they entered school (figure 3). The percentages were highest

Figure 1. For students entering New York City public schools as English learner students in kindergarten during 2003/04–2010/11, the cumulative probability of becoming reclassified as a former English learner student increases gradually, with a majority of students reclassified by the end of their expected grade 3 year

Cumulative probability of reclassification

Expected grade level[a]

Note: *n* = 229,249. Data are based on a discrete-time survival analysis model in which the hazard probability of reclassification in each year was freely estimated. The cohort of students who entered as English learner students in the 2008/09 school year was excluded because of data quality issues.

a. The grade level that a student who was never retained in a grade reached by the given number of school years after entering kindergarten.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Figure 2. The estimated median time to be reclassified as a former English learner student varied by grade of entry to New York City public schools during 2003/04–2010/11

Years

Note: The results are based on a discrete-time survival analysis model in which the hazard probability of reclassification in each year was freely estimated and allowed to differ by grade of entry. The cohort of students who entered as English learner students in 2008/09 was excluded because of data quality issues.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Figure 3. The estimated percentage of students who became long-term English learner students varied by grade of entry to New York City public schools during 2003/04–2010/11
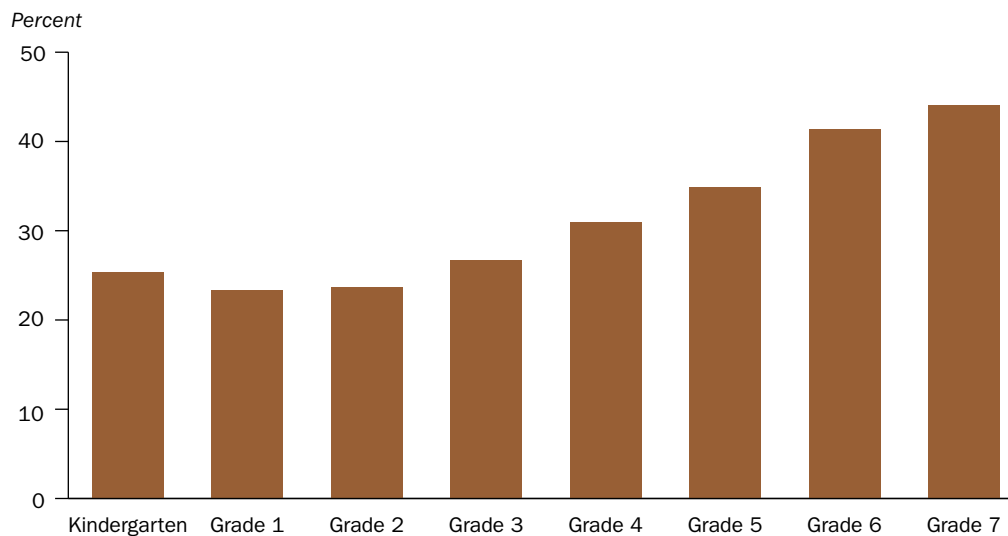
Percent

Note: Results are based on a discrete-time survival analysis model in which the hazard probability of reclassification in each year was freely estimated and allowed to differ by grade of entry. The cohort of students who entered as English learner students in 2008/09 was excluded because of data quality issues. New York City Department of Education defines long-term English learner students as students who have been classified as English learner students for six or more years.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

for students who entered in grade 6 (41 percent) or grade 7 (44 percent) and lowest for students who entered during the primary grades (23–27 percent). In addition, 33 percent of students who entered in grade 6 and 44 percent who entered in grade 7 did not become reclassified by the end of their expected year of high school graduation (seven years after entering grade 6 or six years after entering grade 7).

The median time to reclassification was three years for English learner students with above average initial English proficiency and nearly five years for students with below average initial English proficiency

English learner students' initial English proficiency (when they entered New York City public schools) had a positive relationship with their later probability of being reclassified: students with lower initial English proficiency had longer times to reclassification (figure 4 and table E3 in appendix E). The initial English proficiency of English learner students ranged from virtually no proficiency to proficiency close to that needed to benefit from mainstream instruction without special language support. To illustrate, consider three typical English learner students: with below average initial English proficiency (scoring close to zero on the test), average initial English proficiency (similar to the average English learner student in New York City public schools), and above average initial English

English learner students' initial English proficiency had a positive relationship with their later probability of being reclassified: students with lower initial English proficiency had longer times to reclassification

Figure 4. The estimated median time to become reclassified as a former English learner student in New York City public schools was shorter for English learner students with higher initial English proficiency than for students with lower initial English proficiency, 2003/04–2011/12

Note: Results are based on a discrete-time survival analysis model in which the hazard probability of reclassification in each year was freely estimated and allowed to differ by initial English proficiency. Initial English proficiency is a continuous variable. The figure illustrates typical students with average initial English proficiency (sample mean was 12.1–13.2 raw score points, depending on cohort, out of a total of 70 points), below average proficiency (one sample standard deviation below the sample mean, which was approximately 0 raw score points), and above average proficiency (one sample standard deviation above the sample mean, which was 25.7–26.9 raw score points, depending on cohort). The cohort of students who entered as English learner students in 2008/09 was not included because of data quality issues.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

proficiency (closer to the standard for reclassification). English learner students with above average initial English proficiency took approximately three years to become reclassified, while students with below average initial English proficiency took nearly five years. Thirty-seven percent of students with below average initial English proficiency became long-term English learner students, compared with 19 percent of students with above average initial English proficiency (figure 5).

The median time to reclassification was about four years longer for English learner students with specific learning disabilities and two years longer for English learner students with speech or language impairments than for English learner students without disabilities

The median time to reclassification was approximately 8 years for students with specific learning disabilities, 6 years for students with speech or language impairments, and 3.5 years for students without any disabilities (figure 6 and table E4 in appendix E). Students with other disabilities (disabilities that are not typically considered to be directly related to language) also demonstrated longer times to reclassification than students without disabilities.

Thirty-seven percent of students with below average initial English proficiency became long-term English learner students, compared with 19 percent of students with above average initial English proficiency

Figure 5. The estimated percentage of English learner students who became long-term English learner students was greater for students with lower initial English proficiency when they entered New York City public schools than for students with higher initial English proficiency, 2003/04–2011/12

Percent

Note: Results are based on a discrete-time survival analysis model in which the hazard probability of reclassification in each year was freely estimated and allowed to differ by initial English proficiency. Initial English proficiency is a continuous variable. The figure illustrates typical students with average initial English proficiency (sample mean was 12.1–13.2 raw score points, depending on cohort, out of a total of 70 points), below average proficiency (one sample standard deviation below the sample mean, which was approximately 0 raw score points), and above average proficiency (one sample standard deviation above the sample mean, which was 25.7–26.9 raw score points, depending on cohort). The cohort of students who entered as English learner students in 2008/09 was not included because of data quality issues. New York City Department of Education defines long-term English learner students as those students who have been classified as English learner students for six or more years.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Figure 6. The estimated median time to become reclassified as a former English learner student in New York City public schools was longer for English learner students with specific learning disabilities or speech or language impairments than for English learner students without disabilities, 2003/04–2011/12
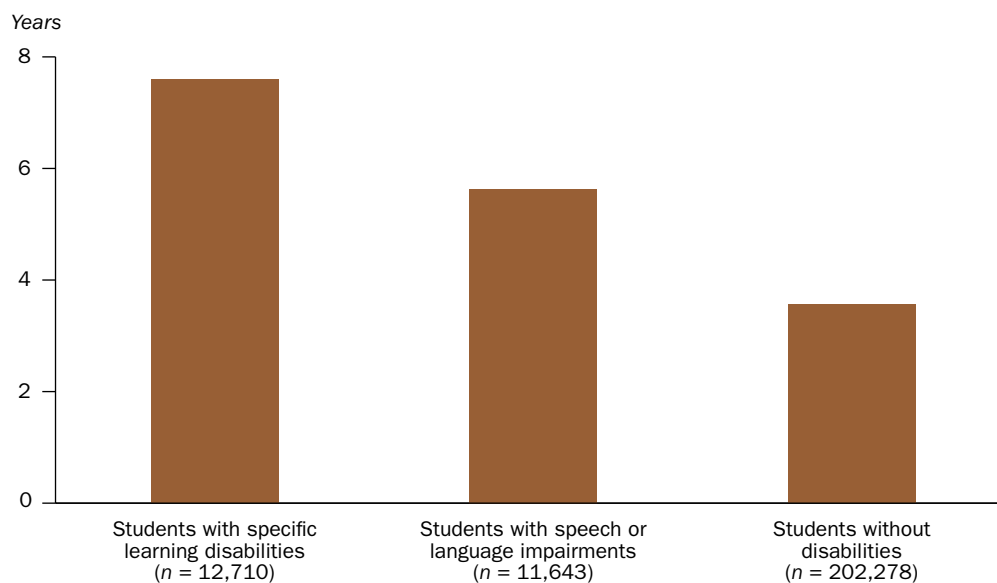
Years

The median time to reclassification was approximately 8 years for students with specific learning disabilities, 6 years for students with speech or language impairments, and 3.5 years for students without any disabilities

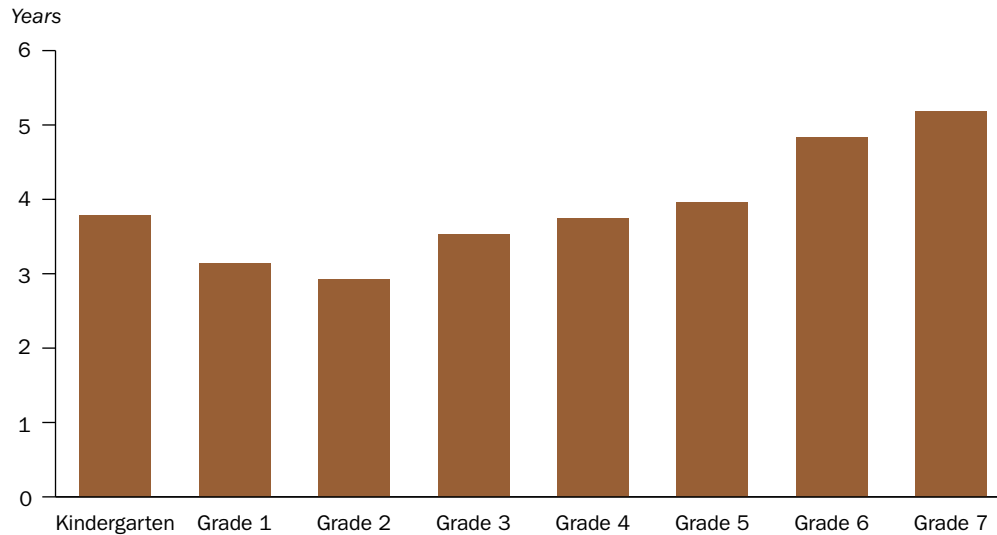**Note:** Results are based on a discrete-time survival analysis model in which the hazard probability of reclassification in each year was freely estimated and allowed to differ by disability category. The cohort of students who entered as English learner students in 2008/09 was not included because of data quality issues.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Sixty-three percent of English learner students with specific learning disabilities became long-term English learner students, while 46 percent of students with speech or language impairments became long-term English learner students (figure 7). Note that these are the percentages of students in these disability categories who attended New York City public schools long enough (six or more years) to be assessed as becoming long-term English learner students. For example, an English learner student who entered in kindergarten would be considered a long-term English learner student at the beginning of grade 6. Among long-term English learner students in grades 6–12, 23 percent were identified with specific learning disabilities and 12 percent were identified with speech or language impairments. For comparison, the districtwide K–12 average is 6.2 percent of students identified with specific learning disabilities and 5.1 percent identified with speech or language impairments (New York City Department of Education, 2011; New York State Education Department, 2011).

Figure 7. Nearly half to two-thirds of English learner students in New York City public schools with speech or language impairments or with specific learning disabilities became long-term English learner students, 2003/04–2011/12
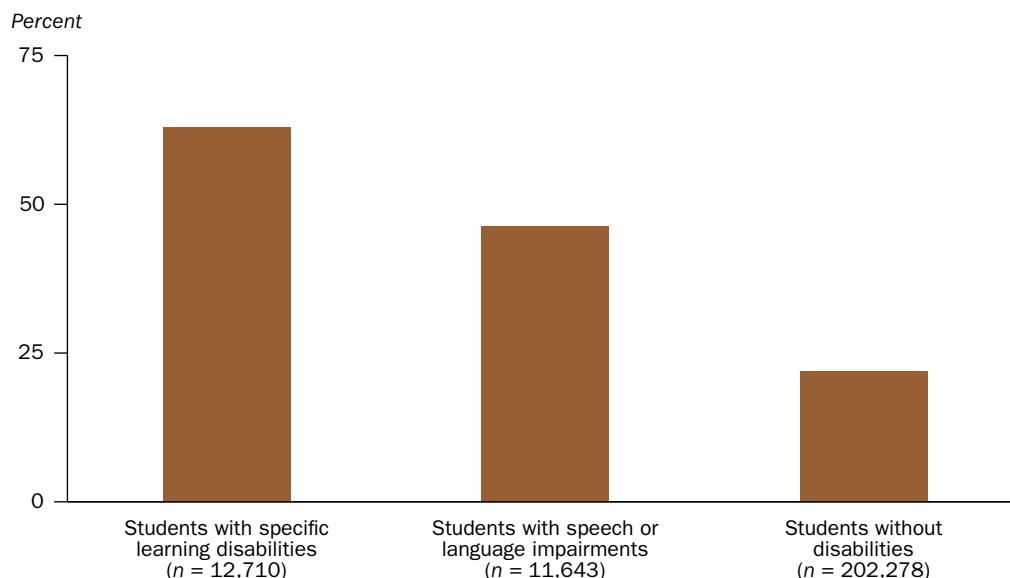
Percent

Students with specific learning disabilities (*n* = 12,710)	Students with speech or language impairments (*n* = 11,643)	Students without disabilities (*n* = 202,278)

Sixty-three percent of English learner students with specific learning disabilities became long-term English learner students, while 46 percent of students with speech or language impairments became long-term English learner students

Note: Results are based on a discrete-time survival analysis model in which the hazard probability of reclassification in each year was freely estimated and allowed to differ by disability category. The cohort of students who entered as English learner students in 2008/09 was not included because of data quality issues. New York defines long-term English learner students as those students who have been classified as English learner students for six or more years.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.
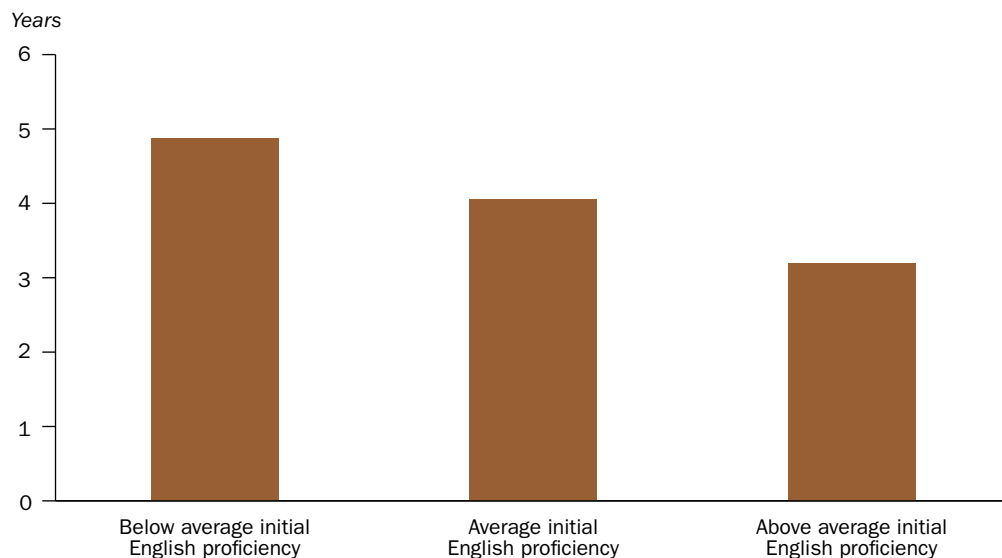
Implications of the study findings

This section presents implications of these findings for policymakers, educators, and researchers.

Students who enter with less English proficiency may need extra supports to catch up to their peers

This study confirms other studies' findings that the time students take to meet reclassification criteria is related to students' initial level of English proficiency (Cook, Boals, & Lundberg, 2011; Thompson, 2015). The median time for English learner students to become reclassified is nearly five years for students who enter with below average initial English proficiency compared with three years for students who enter with above average initial English proficiency. This finding is particularly concerning because late reclassification may be associated with fewer opportunities to learn sophisticated grade-level content (Callahan, 2005). Nonetheless, because this study did not examine the effects of early or late reclassification, future experimental and quasi-experimental research should evaluate whether and under what circumstances late reclassification impedes students' subsequent success. In addition, future studies should investigate whether particular school or classroom interventions are effective in helping English learner students who enter school with lower initial English proficiency develop English skills.

English learner students who enter New York City public schools in grade 6 or 7 are at elevated academic risk and may need more supports than those who enter during the elementary school grades

The median time to reclassification was longer for English learner students who entered New York City public schools in grade 6 or 7 than for students who entered in elementary grades. It is of particular concern that students who entered in grade 6 or grade 7 were not reclassified until they were well into high school, leaving them little time to take advantage of advanced coursework. In addition, many students who enter school as English learner students in grade 6 or 7 were not reclassified by the end of their expected high school graduation year (33 and 44 percent). These findings suggest that districts and states might want to focus their policies and practices on efforts to improve services for English learner students who enter during middle school.

Multiple factors that were not investigated in this study may account for these findings. For instance, English learner students who enter in the middle school grades may take longer to become reclassified because they are entering a more challenging educational and linguistic environment than students who enter school in the primary grades. Middle school English learner students face the double challenge of learning both English and more complex academic content (Short & Fitzsimmons, 2007). It may also be that middle school provides fewer or less effective supports for new English learner students than elementary school. Future research is necessary to explore whether middle schools are less supportive than elementary schools for new English learner students and, if so, to evaluate policies and practices to improve middle school for new English learner students.

Future research is necessary to explore whether middle schools are less supportive than elementary schools for new English learner students and could examine instructional practices for English learner students with disabilities, particularly the way special education and English learner student services are provided

Long-term English learner students who are identified with specific learning disabilities or speech or language impairments may need targeted instruction throughout their academic careers

Among long-term English learner students 23 percent are identified with specific learning disabilities and 12 percent with speech or language impairments. This study was not designed to determine how many of these students were appropriately identified with disabilities that led them to struggle to meet English proficiency standards. Given the challenges in distinguishing disabilities from English language acquisition challenges (Klingner et al., 2005, 2006), further research should investigate disability identification procedures for English learner students and former English learner students. The New York City Department of Education provides resources for teachers on best practices for distinguishing disability from language acquisition challenges (Klingner, n.d.). In addition, further research could examine instructional practices for English learner students with disabilities, particularly the way special education and English learner student services are provided.

Research on English learner student classification systems should explore differences in the median time to reclassification across states

This study estimated a median time to reclassification of four years for English learner students in New York City public schools. Findings differ for other studies that employed different English learner student classification systems and examined students with different demographic characteristics. This study's estimate is three to four years shorter than estimates from similar studies conducted in districts in California (Umansky & Reardon,

2014, Thompson, 2015), and one to two years longer than another study conducted in New York City using data collected before implementation of the No Child Left Behind Act of 2002 (Conger, 2009).

These differences likely stem at least in part from differences in English learner student classification assessments and criteria (Bailey & Kelly, 2013; Linquanti & Cook, 2013). However, these differences may also be due in part to differences in student demographic characteristics or to the influence of other state policies and practices. Further research is needed to investigate the sources of these differences.

It will be especially important to understand the reasons for these differences as many states and districts are moving toward adopting new classification systems (Linquanti & Cook, 2013). These new systems align with a trajectory of English proficiency designed to move students toward the Common Core State Standards. Many states are considering whether and how to adopt the common assessment systems offered by multistate consortia, including the World-class Instructional Design and Assessment Consortium and the English Language Proficiency Assessment for the 21st Century consortium.

Limitations of the study

Several limitations of this study should be considered when interpreting these findings.
- This observational study cannot support causal inferences; rather, it presents descriptive patterns and generates hypotheses for future research employing designs that can better support such inferences.
- The study's scope and design did not allow for investigation of instructional policy and practices that may influence English learner students' progress toward English proficiency.
- These results may not be applicable to English learner students who drop out of school or transfer to another school system outside of New York City public schools before they are reclassified.
- Grade of entry into New York City public schools may not always be the same as the grade at which students entered any school in the United States, so the findings for grade of entry cannot be directly interpreted as the grade at which students began English instruction. Reliable data were not available on the grade at which immigrant students first moved to the United States.

This observational study cannot support causal inferences; rather, it presents descriptive patterns and generates hypotheses for future research employing designs that can better support such inferences

Historically, systems for identifying English learner students and for reclassifying them as former English learner students have varied considerably across states and districts (August & Hakuta, 1997). Under the requirements of the No Child Left Behind Act of 2002, states were required to administer annual English language proficiency tests to inform decisions about English learner student identification and reclassification. All students whose parents reported speaking a home language other than English when the students entered school were assessed on an English language proficiency test. Students who scored not yet proficient in English were classified as English learner students. The No Child Left Behind Act specified that these tests had to include reading, writing, listening, and speaking subtests. The results were used to determine annual measurable achievement objectives, which included monitoring the percentage of English learner students who make progress as well as the percentage who reach English proficiency (reclassification as former English learner students; Cook, Chinen, & Jung, 2012).

The changes in the assessment systems in response to the No Child Left Behind Act provided higher quality longitudinal information about English learner students' progress toward English proficiency than had previously been available. However, the systems continued to vary considerably from state to state and sometimes district to district (Ragan & Lesaux, 2006). The higher-quality longitudinal datasets have made it possible to explore trajectories of English learner students toward reclassification, as well as variations in those trajectories by student characteristics. This information can inform policy and practice decisions, from district- and state-level policies about the criteria for reclassification to school-level decisions about which student groups may require different types or intensity of instructional support.

Time to reclassification

One basic, descriptive question is whether and when English learner students become reclassified as former English learner students by meeting standards for English language proficiency in New York City (NYC) public schools by reaching a specified cutscore on the New York State English as a Second Language Achievement Test. Reclassification is a critical step for English learner students because it can determine the types of services they receive and their access to advanced content (Umansky & Reardon, 2014). Studies before the No Child Left Behind Act found that it takes up to seven years for students to gain English proficiency, but those studies did not have reliable measures of English proficiency or high-quality longitudinal datasets (Hakuta, Butler, & Witt, 2000).

Only recently have studies examined time to reclassification using longitudinal data and methods that are similar to those used in the current study. In an analysis of data from the Los Angeles Unified School District, Thompson (2015) found a window for reclassification in the upper elementary grades that closes after grade 5. Among students who entered kindergarten as English learner students, those who did not meet reclassification criteria by the end of elementary school (between kindergarten and grade 5) were much less likely to be reclassified in middle school (between grade 6 and 8), compared with those who were reclassified before the end of elementary school. The median time to reclassification for Latino kindergarteners in a large urban district is seven years according to Umansky &

Reardon (2014). These studies used data from districts with assessments and reclassification criteria that differ from those in NYC public schools, and similar analyses with NYC data may thus yield different results.

Predictors of time to reclassification: Grade of entry, initial English proficiency, and disability category

In addition to the characteristics typically considered predictive of achievement for English learner students (socioeconomic status, race/ethnicity, and home language backgrounds), investigators should consider less commonly studied student characteristics that may differentiate reclassification patterns among English learner students. These include grade of school entry, initial English proficiency, and disability category. Understanding variation in reclassification rates by student characteristics can help identify which type or intensity of instructional support different student groups may require.

Students' grade of entry in NYC public schools is likely to be important to their time to reclassification, but the direction of this relationship is unclear. This uncertainty stems in part from conflicting evidence about the role of age in second-language acquisition. Although some studies demonstrate benefits for younger children in acquiring a second language (Gass & Selinker, 2008), others find that older children can learn a second language more quickly than younger children when they are provided sufficient opportunities to use the language (Marinova-Todd, Marshall, & Snow, 2000). This uncertainly also stems in part from the fact that English learner students who are immigrants are likely to differ from English learner students who are born in the United States but who are not yet proficient in English when they arrive at kindergarten (August & Shanahan, 2006). A study that investigated the relationship between age of entry and time to reclassification for English learner students in NYC public schools found that younger English learner students reclassify more quickly than older ones, but the study used data before the No Child Left Behind Act (Conger, 2009). Another study found that "students of equivalent levels of language proficiency at higher grade levels take longer to attain the standard than their counterparts in lower grade levels" (Cook et al., 2012, p. 30). Given that English learner students enter U.S. schools at all grade levels, understanding variation in their trajectory toward reclassification can help determine whether age or grade of entry should be considered when designing instructional programs.

English learner students' initial English proficiency may be valuable in identifying students who are likely to be reclassified earlier (or at all; Kieffer, 2008). For students in California initial English proficiency in kindergarten had a strong relationship with students' probability of being reclassified within nine years (Thompson, 2015). Using data from 27 states that participate in the World-class Instructional Design and Assessment Consortium, Cook et al. (2011) found that time to reclassification depended strongly on initial English proficiency and that the median years to proficiency was four years for students with the highest initial English proficiency but up to seven years for English learner students with the lowest initial English proficiency. Confirming these findings for NYC public school students who face different reclassification criteria is valuable for decisions in the district and for clarifying potential differences between states in these patterns.

A student's disability category also warrants further research, particularly concerning specific learning disabilities and speech or language impairments (Klingner et al., 2006; Orozco et al., 2008). Educators across the country are increasingly concerned about how to

meet the needs of English learner students with disabilities. In NYC public schools nearly 22 percent of English learner students (approximately 34,000 students) are classified with a disability and receive some kind of special education services. A recent study conducted by the Regional Educational Laboratory Northeast & Islands English Language Learners Alliance in a New England district found that English learner students with disabilities performed substantially worse on English language achievement assessments and content tests than English learner students without disabilities (Parker et al., 2014). Although one might expect that English learner students with disabilities would take longer to be reclassified than English learner students without disabilities (or might not be reclassified at all before graduating from high school), the extent of these differences is unclear. Describing this pattern will be of use to district officials in understanding how well they are serving this dually identified population.

Characteristics of those who are not reclassified (long-term English learner students)

The characteristics associated with time to reclassification might also be related to becoming a long-term English learner student (defined in New York as not being reclassified after six years). One statewide study of all grade 9 English learner students found that 60 percent of them had been born in the United States, indicating that they had been in English learner programs for at least eight years (Slama, 2012). A similar study found that more than 60 percent of English learner students in one district became long-term English learner students, and 25 percent were not reclassified as former English learner students by grade 12 (Umansky & Reardon, 2014). While more attention is being paid to understanding the characteristics and experiences of long-term English learner students (Olsen, 2010), there has been limited research describing specific characteristics of long-term English learner students (such as grade of entry, initial English proficiency, and disability category).

Appendix B. Study sample

This appendix describes the sample in detail, including the sampling strategy and resulting sample sizes, descriptive statistics for student characteristics, and descriptive statistics for student disability categories.

Sampling strategy and sample sizes by grade and by cohort

This study used administrative data for seven cohorts of students who entered New York City (NYC) public schools as English learner students between 2003/04 and 2010/11. (The cohort of students who entered as English learner students in 2008/09 was not included in the study because of data quality problems; see below.) The earliest cohort of students was followed for nine years (from 2003/04 through 2011/12), and the latest cohort was followed for two years (from 2010/11 through 2011/12). The cohorts in between were followed for periods ranging from three to eight years. Students with fewer than nine years of data (those in cohorts after the 2003/04 cohort) contributed data for the years they were observed and were treated as censored for subsequent years when they were not observed (see appendix C for more information about censoring.)

Starting with the 2003/04 school year allows for a relatively long span of time while also including only data that were collected after implementation of the No Child Left Behind Act beginning in the 2002/03 school year in New York City. District documentation as well as preliminary analyses of data from earlier years indicated that implementation of the No Child Left Behind Act led to substantial changes in English learner student assessment and classification practices, including both initial identification and reclassification practices in NYC public schools. In addition, implementation of the No Child Left Behind Act also led to changes in data reporting practices, which might also compromise the ability to conduct analyses that incorporate data before and after the 2003/04 school year.

The students in each cohort were identified as newly entering English learner students based on information from the enrollment dataset and the Bilingual Education Student Information Survey (BESIS) dataset provided by the New York City Department of Education.[2] Specifically, students were identified as newly entering English learner students if enrollment data indicated that they had not been actively enrolled in a given year, BESIS data indicated that they were classified as English learner students in the same year, and enrollment data indicated that they were not actively enrolled the previous year. Thus, this approach used BESIS data starting with the 2003/04 school year and enrollment data starting a year before, from the 2002/03 school year. Students who were both enrolled and classified as English learner students in 2002/03 were not included in the sample because their time of entry into NYC public schools (and initial classification as English learner students) was not observed. Given this strategy, students who entered NYC public schools as English learner students before fall 2002, left the district, and then returned as English learner students during or after fall 2003 were treated as new English learner students.

The analytic sample excludes students who entered as English learner students in the 2008/09 school year. Preliminary analyses indicated that the BESIS data used for identifying new English learner students were not reliable for that particular school year. Specifically, there were substantial inconsistencies between the students identified as English learner students in the BESIS data and in two other independent datasets that identify

English learner students (the enrollment dataset, which included a variable representing English learner status in school years 2008/09 and after, and the New York State English as a Second Language Achievement Test dataset), whereas no such inconsistencies appeared for other school years. For consistency in the identification of new English learner students across cohorts, and given the already large sample sizes provided by the other seven cohorts, the decision was made to exclude this cohort from the analyses. As described below, additional efforts were made to address the unreliability of the 2008/09 BESIS data in identifying the time of reclassification.

The resulting analytic sample included 229,249 students who entered NYC public schools between fall 2003 and fall 2010 and were initially classified as English learner students. Nearly 44 percent of these students entered NYC public schools in kindergarten, though sample sizes were large for all grade levels (table B1). The sample sizes were similar across the seven cohorts (table B2).

The first research question focused on students who entered NYC public schools as English learner students in kindergarten, with the goal of comparing the probability of reclassification as it changes over grade levels. Analyses were conducted with the complete dataset (including students who entered NYC public schools after kindergarten), but the findings relied on probabilities of reclassification fitted for students who entered NYC public schools in kindergarten. Data from the earliest cohort allowed the study team to investigate the probability of reclassification from kindergarten through the end of grade 7 (for students who were not retained in grade), while data from later cohorts contributed to estimates for the probability of reclassification in earlier grades.

The second research question leveraged the data on students who entered after kindergarten to facilitate comparisons in time to reclassification between students who entered at

Table B1. Sample size and percentage of English learner students, by grade of entry to New York City public schools, 2003/04–2010/11

Grade of entry	Students in sample	
	Number	Percent
Kindergarten	100,206	43.7
1	18,988	8.3
2	11,135	4.9
3	9,648	4.2
4	9,373	4.1
5	9,239	4.0
6	9,288	4.1
7	10,132	4.4
8	10,773	4.7
9	20,203	8.8
10	13,781	6.0
11	4,812	2.1
12	1,671	0.7
Total	229,249	

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table B2. Sample size and percentage of English learner students, by year of entry to New York City public schools, 2003/04–2010/11

Year of entry	Students in sample	
	Number	Percent
2003/04	34,015	14.8
2004/05	32,099	14.0
2005/06	33,604	14.7
2006/07	35,438	14.6
2007/08	29,127	12.7
2009/10	30,990	13.5
2010/11	35,976	15.7
Total	229,249	

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

different grade levels, among other comparisons. Because the earliest cohort was followed for nine years, these analyses could investigate spans up to nine years for students who entered at different grades. For instance, among students in the earliest cohort those who entered NYC public schools as English learner students in grade 1 were observed through grade 10, students who entered in grade 2 were observed through grade 11, and students who entered in grade 3 were observed through grade 12. By combining students who entered at different grade levels across cohorts, the study team examined reclassification probabilities across the entire grade span, though a given student was only observed for up to nine years.

Student characteristics of the analytic sample

Descriptive statistics for gender, race/ethnicity, home language group, and family poverty for the analytic sample are listed in table B3. For the gender variable male students are coded with a 1, and female students are coded with a 0. Family poverty is indicated by federal school lunch program status (eligible or not eligible); students eligible for the school lunch program had a value of zero for this variable. Neighborhood average per capita income (the average annual income in the census track in which the student resided) for students in the analytic sample, an additional indicator of socioeconomic status, had a mean of $20,718, a median of $18,133, and a standard deviation of $11,535. Among all students in the analytic sample 14.8 percent (n = 33,985) were retained in grade once across the years studied, and 4.05 percent (n = 9,293) were retained in grade twice or more.

Descriptive statistics for disability categories are listed in table B4. The disability category variables refer to the disability category that students were identified with at any time during the years observed; if a student's records for any of the years available indicated a disability category (regardless of whether it was a year in which the student was classified as an English learner student or a year in which the student was classified as a former English learner student), the student was treated as part of that disability category. The largest categories were specific learning disabilities and speech or language impairments. No other category included more than 0.5 percent of students in the analytic sample.

Table B3. Demographic characteristics of English learner students entering New York City public schools in 2003/04–2010/11

Characteristic	Students in sample	
	Number	Percent
Gender		
Male	119,672	52.2
Race/ethnicity		
Hispanic	132,863	58.0
Asian/Pacific Islander	60,980	26.6
White, Non-Hispanic	20,742	9.1
Black	13,304	5.8
Native American	556	0.2
Unreported	718	0.3
Multiracial/multiethnic	86	0.04
Home language group		
Spanish	132,629	57.9
Chinese	35,084	15.3
Another language	32,900	14.3
Bengali	8,226	3.6
Russian	6,538	2.9
Haitian Creole	4,117	1.8
French	3,172	1.4
Korean	3,027	1.3
English[a]	2,567	1.1
Southeast Asian Languages	989	0.4
Family poverty		
Eligible for the federal school lunch program	154,659	70.1

Note: n = 229,249. Percentages may not sum to 100 because of rounding.

a. The home language group information is based on a time-invariant demographic dataset that combines demographic information across multiple years. Although this decision yields the most complete data on this variable, it also produces the somewhat surprising result that a small proportion of English learner students list English as their home language. It is not entirely clear from district data and records why this would be the case; the most likely explanation is that these students' parents reported a home language other than English when they first entered school but then reported English as their home language in subsequent years.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table B4. Descriptive statistics for disability categories for students entering New York City public schools in 2003/04–2010/11

Disability category	Students in sample	
	Number	Percent
Specific learning disabilities	12,710	5.54
Speech or language impairments	11,643	5.08
Other disabilities		
Autism	64	0.03
Deaf	9	0.00
Emotionally disturbed	554	0.24
Hard of hearing	375	0.16
Intellectual disabilities	37	0.15
Multiple disabilities	32	0.01
Other health impaired	1,032	0.45
Traumatic brain injury	21	0.01
Visually impaired	51	0.02
No disability	202,721	88.43

Note: $n = 229,249$. Percentages may not sum to 100 because of rounding.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Appendix C. Analytic approach

This appendix describes the analytic approach, including an overview of the methods used, details on the hypothesized models for each research question, and treatment of missing data due to attrition and other reasons.

Overview of discrete-time survival analyses

Discrete-time survival analytic (DTSA) methods (Singer & Willett, 2006) were used to estimate the probability of reclassification as it changed over time. DTSA methods are considered the most appropriate for longitudinal questions about whether and when an event occurs, in cases where the periods studied are discrete and reasonably small in number (like the years in this study). These methods have the advantage of modeling the probability of reclassification over time without making assumptions regarding students who were censored (that is, were not observed experiencing the event). In this study censoring can occur when the event meets one of the following criteria:

- Never occurs (students graduate from high school still classified as English learner students).
- Occurs after the observation period ends (students were reclassified after 2011/12, when data were not available).
- Occurs after students left the study, meaning that they left New York City (NYC) public schools before becoming reclassified.

With this approach students who were not observed experiencing the event do not need to be casewise deleted from the analytic sample, thus avoiding the bias in time to reclassification (and relations to time to reclassification) that would likely result from such deletion. Rather, they contribute to the estimate of the probability of reclassification during those periods in which they are observed. For this study this approach means that students who were not observed as being reclassified due to censoring nonetheless do contribute to the estimation of time to reclassification and percent of long-term English learner students, because they contribute to the estimation of the probability of reclassification in the years they were observed and thereby to the cumulative probability in later years.

The beginning of time was specified as the beginning of a student's first year as an English learner student in NYC public schools. The event was defined as reclassification; the first potential period was defined as before the end of the first year, and the later periods were defined as between the end of each year, beginning with the first year, and the end of the subsequent year.

To determine whether and when reclassification occurred, a variable was created for the event of reclassification in each period. Reclassification processes in NYC public schools occurred in the spring (and sometimes continued into the summer). Given this, the timing of reclassification was defined based on data from the enrollment dataset indicating that a student continued to be enrolled in NYC public schools in a given year but was no longer classified as an English learner student in the Bilingual Education Student Information Survey (BESIS) dataset after having been classified as such in the previous year. If a student was a former English learner for the first time in the fall of a given school year, the student was indicated as reclassified by the end of the previous school year (and after the end of the school year before that). Because this approach uses one year's data to

determine whether a student was reclassified by the end of the previous school year, the years for which the reclassification variable is available are one fewer than the years that a cohort was observed. For instance, the earliest cohort was observed for nine years, providing eight periods in which reclassification could occur (for example, end of kindergarten to end of grade 7 for students entering the beginning of kindergarten or end of grade 4 to end of grade 12 for students entering in the beginning of grade 4).

Preliminary analyses indicated that the BESIS data reliably identified students classified as English learner students in a given school year for all years except 2008/09. Specifically, for the other years, the BESIS data were consistent with two independent datasets—the New York State English as a Second Language Achievement Test dataset and the enrollment dataset, which included a variable representing English learner status (in the 2008/09 school year and beyond). However, the BESIS data were not consistent with those two datasets (which were consistent with one another) for 2008/09. To address this problem, the variable representing English learner status in the enrollment file was used in place of the BESIS data to determine whether the event of reclassification had occurred in 2008/09 and in 2009/10.

In all analyses statistical significance was determined using robust standard errors that account for clustering of students within schools. Specifically, these robust standard errors account for the lack of independence in the outcome of reclassification by accounting for the school in which students were clustered, using a time-varying school identification variable. Because the analyses were conducted using a person-period or long-format dataset, the values for the school identification variable were allowed to differ for different periods for students who switched between schools in the years observed. In such cases the value for the school identification variable was consistent with the school in which the student was enrolled at the period associated with the reclassification variable (the school identification value and the reclassification value on the same row of the person-period dataset represent the same year, but students who switched between schools had different school identification variables for different rows in the person-period dataset). Alternate methods were considered, including cross-classified multilevel modeling, but given the lack of substantive interest in the school-level variance components, robust standard errors that account for clustering were chosen.

Hypothesized model for the first research question

The first research question, involving reclassification patterns for the median English learner student who entered NYC public schools in kindergarten, was addressed using a DTSA model with the subset of data for the 100,206 students who entered New York City public schools as English learner students in kindergarten during the years studied. The final hypothesized model for the first research question is given as follows:

$$\text{logit } h(t_j) = [\alpha_1 D1_{Y1} + \alpha_2 D2_{Y2} + \alpha_3 D3_{Y3} + \alpha_4 D4_{Y4} + \alpha_5 D5_{Y5} + \alpha_6 D6_{Y6} + \alpha_7 D7_{Y7} + \alpha_8 D8_{Y8}] \quad \text{(C1)}$$

In this model, logit $h(t_j)$ represents the logit hazard probability of reclassification for time period j. The eight α terms collectively capture the effects of time (specified as years) on the logit hazard probability of reclassification. Including eight α terms allows the hazard function to vary across all eight time periods without assuming a specific functional form (for example, linear or quadratic). Note that there is no intercept term, so the estimates for

each α term can be interpreted as the conditional logit hazard probability of reclassification for its respective time period (by the end of the specified year).

The resulting fitted logit hazard probabilities were then converted into a hazard function and then into a survival function to estimate the median lifetime (median time to reclassification) and the fitted survival probability after six years (the probability of becoming a long-term English learner student) that are discussed in the main report.

Hypothesized models for the second research question

Given that the second research question involves three student characteristics of interest—grade of entry, initial English proficiency, and disability category—three separate, parallel DTSA models were fitted to investigate associations between reclassification patterns and each characteristic individually. These models were fitted to data with the entire dataset (the 229,249 students who entered NYC public schools as English learner students at any grade in the years studied). In each of the three separate models on which the key findings were based, the other characteristics of interest (for example, grade of entry when initial English proficiency was the focus) were not included as controls nor did the models control for student demographic characteristics. Appendix F provides additional information on supplemental analyses conducted to explore the extent to which estimates changed when controls were included.

To investigate relationships between grade of entry and reclassification patterns, grade of entry was added as a categorical time-invariant predictor with time-varying effects to the DTSA model (see equation C1). A set of interactions between dummy variables representing each grade of entry from grade 1 to grade 12 and each time variable ($D1$ to $D8$) was included (with entry in kindergarten specified as the omitted reference category). This model produced coefficients for the hazard probabilities of reclassification in each period for each grade of entry on a logit scale. To convert these logit hazard probabilities into interpretable results (time to reclassification in years and percentage of long-term English learners), the coefficients were used to generate a set of eight hazard functions, one for each grade of entry between kindergarten and grade 7. Corresponding survival functions were generated for each grade of entry between kindergarten and grade 7, which yielded estimates of the median time to reclassification and the probability that students would become long-term English learner students, by grade of entry (see figures 2 and 3 in the main report). Hazard and survival functions were also generated for students who entered in grades 8–12, but these results are not reported, for two reasons. First, for each of these grades, a median lifetime was not reached before the end of their expected grade 12 year (when they are expected to leave high school), so the median time to reclassification is not meaningful. Second, the potential for systematic (informative) censoring due to students dropping out of school is relatively high for these students, so results may be misleading (see below).

To investigate the relationships between initial English proficiency and reclassification patterns, a separate DTSA model was fitted with initial English proficiency included as a continuous time-invariant predictor with time-varying effects to the model in equation C1. (Note that grade of entry was not included as a control in this model.) Initial English proficiency was based on the Language Assessment Battery–Revised total score (which incorporates information on the reading, writing, listening, and speaking subtests) at

school entry. The assessment uses different forms at different grade levels, and it did not appear from district and state records that the different forms were vertically equated (see appendix D). Given this, the raw total scores were converted to *z*-scores within cohort and grade, so that the mean was 0 and the standard deviation was 1 within each cohort and grade.

Language Assessment Battery–Revised scores from the year of entry were missing for 21.2 percent of the students; the missing data were taken into account using a model-based approach in which a dummy variable represented the missing data and an arbitrary value was substituted for the missing value; including these together enabled retaining the complete analytic sample. A set of interactions between the total *z*-score score for initial English proficiency score and the time variables ($D1$ to $D8$) was included to allow the relationship between initial English proficiency and the probability of reclassification to be different over time, as one would expect. To facilitate interpretation, the resulting coefficients from this model (in which initial English proficiency is treated as a continuous variable) were used to generate hazard and survival functions for three prototypical students (at different values along the continuum of initial English proficiency). Specifically, three hazard functions were generated for three prototypical students with values of initial English proficiency at the average (sample mean or 12.1–13.2 raw score points, depending on cohort, out of a total of 70 points); below average (one sample standard deviation below the sample mean or approximately 0 raw score points); and above average (one sample standard deviation above the sample mean, which is 25.7–26.9 raw score points, depending on cohort). Corresponding survival functions were generated by initial English proficiency, yielding estimates of the median time to reclassification and the probability that students would become long-term English learner students (see figures 4 and 5 in the main report).

To investigate the relationships between disability categories and reclassification patterns, disability category was added to the model in equation C1 as a categorical, time-invariant predictor with time-varying effects. In a single model using the entire dataset, a set of interactions between the time variables ($D1$ to $D8$) and three dummy variables representing disability category groups (specific learning disabilities, speech or language impairments, and other disabilities, with no disabilities specified as the reference category) were included. Note that neither grade of entry nor initial English proficiency were included as controls in this model. The resulting coefficients from this model were used to generate three hazard functions for prototypical students without disabilities, students with specific learning disabilities, and students with speech or language impairments. Corresponding survival functions were generated, yielding estimates of the median time to reclassification and the probability that students would become long-term English learner students, by disability category (see figures 6 and 7 in the main report). Hazard and survival functions for students with other disabilities were fitted, but estimates for this group were not reported in the key findings, given a lack of substantive interest in reclassification patterns for English learner students in these disability categories.

Treatment of missing data due to attrition and other reasons

There are multiple sources of missing data in a study with this design. As mentioned, data were missing on initial English proficiency, possibly due to data reporting errors and other sources. The missing data were taken into account using a model-based dummy variable approach, which maintains the full sample size.

In addition, missing data related to censoring (not observing the event of reclassification) had multiple sources in this study, as in any longitudinal study. In this study censoring occurred for four reasons:

- Students never having been reclassified.
- Students being reclassified after the years of data available.
- Students attriting from the study because they left the NYC public schools.
- Students attriting from the study because they dropped out of school entirely at age 16 or older.

By design, DTSA handles censoring due to students never having been reclassified without evoking additional assumptions by estimating the logit hazard probability based on risk sets in which students are observed either experiencing the event or not. With respect to censoring due to the other three reasons, DTSA assumes that censoring is noninformative; that is, it assumes that the cases in a given risk set (nonmissing cases in a given period such as beginning of year 2 to end of year 2) are representative of the population of students that would have entered that risk set if they had not been censored. In this sense, the assumption is not that the students in the overall analytic sample are representative of the population of NYC students, but rather that the students in each risk set are representative of the students who would have been in that risk set if censoring were unrelated to the variables of interest. This assumption is most reasonable for students who are reclassified after having years of available data because the source of this censoring is the years of data available to the researchers. Thus censoring is very likely to be exogenous to the relations of interest. It is nonetheless important to keep in mind that these results may not generalize to English learner students who entered NYC public schools before 2003/04 or after 2011/12, if policy or instructional changes related to reclassification patterns occurred. (The preliminary analyses suggest that this is the case for implementation of the No Child Left Behind Act in 2002/03.)

DTSA also assumes noninformative censoring due to attrition. Estimates of annual attrition by cohort for each year in which that cohort was observed are displayed in table C1. Note that year 2 (students' second year in NYC public schools) is the first possible year for attrition, because all students had to be observed in year 1 (their first year in NYC public schools) to be included in the sample. The cumulative attrition by year for each cohort is displayed in table C2.

The assumption of noninformative censoring due to students attriting from the study because they left the NYC public schools would bias the findings to some degree if the English learner students who left NYC public schools to attend other schools were less or more likely to be reclassified in the periods after they left, compared with the students who remained in NYC public schools during those periods. The direction of this bias is not entirely clear. The assumption of noninformative censoring may be particularly inappropriate for censoring due to students attriting from the study because they dropped out of school entirely at age 16 or older, because students who drop out of school may be different in many ways from students who do not drop out of school. If these differences are related systematically to these students' probability of reclassification in the high school years (after they dropped out), the hazard probabilities for reclassification for students in their high school years may be biased. If English learner students who drop out of school are less likely to become reclassified later than English learner students who do not drop out, this bias would yield higher probabilities of reclassification for students in their high school

Table C1. Estimates of annual attrition of students in the study, by number of years since entry to New York City school as English learner student and by cohort year of entry, 2003/04–2010/11 (percent)

Years since entry	Cohort year of entry						
	2003/04	2004/05	2005/06	2006/07	2007/08	2009/10	2010/11
Year 2	9.85	11.18	10.26	10.15	8.01	8.33	8.09
Year 3	9.83	9.31	8.53	7.80	6.86	6.73	na
Year 4	7.98	7.6	7.31	6.05	5.59	na	na
Year 5	5.31	4.99	4.13	4.04	3.48	na	na
Year 6	4.05	3.84	3.66	3.67	na	na	na
Year 7	3.93	3.74	3.84	na	na	na	na
Year 8	2.85	2.83	na	na	na	na	na
Year 9	2.74	na	na	na	na	na	na
Average	5.82	6.21	6.29	6.34	5.99	7.53	8.09

na is not applicable.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table C2. Estimates of cumulative attrition of students in the study, by years since entry to New York City school as English learner student and by cohort year of entry, 2003/04–2010/11 (percent)

Years since entry	Cohort year of entry						
	2003/04	2004/05	2005/06	2006/07	2007/08	2009/10	2010/11
Year 2	9.85	11.18	10.26	10.15	8.01	8.33	8.09
Year 3	18.71	19.45	17.91	17.16	14.32	14.50	na
Year 4	25.20	25.57	23.92	22.17	19.11	na	na
Year 5	29.17	29.29	27.06	25.31	21.92	na	na
Year 6	32.04	32.00	29.73	28.06	na	na	na
Year 7	34.71	34.54	32.43	na	na	na	na
Year 8	36.57	36.40	na	na	na	na	na
Year 9	38.31	na	na	na	na	na	na

na is not applicable.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

years. For this reason the key findings regarding grade of entry do not include results for students who entered school in grade 8 or later (as noted in box 2 in the main report). There is a reasonable chance that missing-data bias due to students dropping out in their high school years would have yielded a more negatively biased time to reclassification estimate for these students than for students who enter in earlier grades. For this reason time to reclassification may also be negatively biased to some extent for students who entered in grade 6 or 7. However, this bias would not call into question the key finding that students who enter in these grades take longer to become reclassified than students in the elementary grade. Rather, it would suggest that this difference is underestimated, such that grades 6 and 7 are an even riskier time to enter NYC public schools as English learner students than study team's current estimates indicate.

Appendix D. Measures used to create initial English proficiency and reclassification variables

This appendix provides psychometric details on the measures used to create the initial English proficiency and reclassification variables. All other variables indicate demographics reported by parents or students or disability categories determined by school processes, so psychometric details are not relevant.

Initial English proficiency

The Language Assessment Battery–Revised is the instrument used for initial classification of students who speak a home language other than English as English learner students; however, as described in appendix A, the study used Bilingual Education Student Information Survey data on students' English learner student classification rather than Language Assessment Battery–Revised data to identify students who were initially classified as English learner students. For the question concerning the relationship between initial English proficiency and reclassification patterns, the study used the total raw score from the Language Assessment Battery–Revised, which represents students' overall English proficiency in reading, writing, listening, and speaking. Although the Language Assessment Battery–Revised also provides broad proficiency levels, these are too coarse to be used in the analyses. For students classified as English learner students (not including the level for students initially classified as non-English learner students), there are two levels (beginning/intermediate and advanced) for kindergarten to grade 8 and three levels (beginning, intermediate, and advanced) for grades 9–12. The test is administered at four levels: grades K–2, 3–5, 6–8, and 9–12, and includes reading, writing, speaking, and listening. According to the Clearinghouse on Assessment and Education, "norms were established on both English-proficient and limited-English-proficient students, with between 150 and 450 subjects per grade level for each form" (Center for Equity and Excellence in Education & Clearinghouse on Assessment and Evaluation, 1982, abstract). The Kuder-Richardson 20 estimates of reliability range from .88 to .96 on individual subtests and from .95 to .98 on the total test. Some validity evidence is also provided. Like many language assessments from the last decade, it does not appear from the technical records that this assessment is vertically scaled. Given this, within-grade and within-cohort z-scores based on the total raw scores were used.

Reclassification

The event of reclassification was defined based on appearing in district data as continuing to be enrolled in New York City (NYC) public schools in a given school year but not being classified as an English learner student after having been classified as an English learner student in the prior school year. This variable is thus based on actual classification, rather than on students' test scores on the English proficiency test used for reclassification purposes. Nonetheless, preliminary analyses indicated that nearly all students who reach the overall English proficient level on the New York State English as a Second Language Achievement Test when taken in the spring of a given year were reclassified as former English learner students by the end of that year. The New York State English as a Second Language Achievement Test is taken annually by all students currently classified as English learner students and is used to determine if a limited English proficiency or English learner student continues to be limited English proficient, based on student scoring at a

state designated level of English proficiency on the assessment. The test is administered in five grade spans (K–1, 2–4, 5–6, 7–8, and 9–12) and consists of reading, writing, listening, and speaking subtests. The tests are vertically scaled. As described in the 2008 technical manual for the test: "The tests are designed to assess the English language skills of English language learners in grades K–12 and to capture their progress toward achieving full English language proficiency" (Pearson, 2009, p. 6). Technical reports indicate adequate reliability, with ranges across the grade spans from .67 to .89 for reading, .78 to .89 for writing, .73 to .90 for listening, and .93 to .97 for speaking (New York State Department of Education, 2015). The reliability for the combined reading and writing score ranged from .84 to .94, and for the combined listening and speaking score .89 to .96. The reports describe validity evidence, including internal correlations between the subscores and external correlations with the New York English Language Assessment.

Appendix E. Fitted model estimates for key findings

This appendix provides the full fitted model estimates on which the key findings in the main report were based (tables E1–E4). The final fitted model estimates reported include selected estimated coefficients, robust standards errors (accounting for nesting of students within the school attended at the time of reclassification for each person-period entry in the dataset), and *p*-values for statistical significance tests.

Table E1 presents reclassification patterns over time for students who enter New York City public schools as English learner students in kindergarten.

Table E1. Fitted estimates for the relationship between logit hazard probability of reclassification and time for students entering kindergarten as English learner students, 2003/04–2010/11

Variable (main effects of time)	Coefficient (in logits)	Robust standard error
D1	–2.10	0.04
D2	–1.37	0.04
D3	–1.65	0.03
D4	–1.38	0.03
D5	–1.11	0.03
D6	–0.90	0.03
D7	–1.11	0.04
D8	–1.34	0.06

Significant at *p* = <.001.

Note: *n* = 100,206. *D* refers to dummy variables representing time period, with the number referring to year since school entry.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table E2 presents the findings for the relationships between grade of entry and probability of reclassification over time.

Table E2. Fitted estimates for the relationship between grade of entry and logit hazard probability of reclassification in each period, 2003/04–2011/12

Variable (main effects of time)	Coefficient (in logits)	Robust standard error	*p* value
Entering in kindergarten (reference category)			
D1	–2.10	0.04	<.001
D2	–1.37	0.04	<.001
D3	–1.65	0.03	<.001
D4	–1.38	0.03	<.001
D5	–1.11	0.03	<.001
D6	–0.90	0.03	<.001
D7	–1.11	0.04	<.001
D8	–1.34	0.06	<.001

(continued)

Table E2. Fitted estimates for the relationship between grade of entry and logit hazard probability of reclassification in each period, 2003/04–2011/12 *(continued)*

Variable (main effects of time)	Coefficient (in logits)	Robust standard error	*p* value
Entering in grade 1			
xD1	0.33	0.04	<.001
xD2	−0.33	0.04	<.001
xD3	0.34	0.03	<.001
xD4	0.31	0.04	<.001
xD5	0.22	0.04	<.001
xD6	−0.22	0.06	0.001
xD7	−0.33	0.11	0.003
xD8	0.25	0.14	0.066
Entering in grade 2			
xD1	−0.18	0.05	0.001
xD2	−0.09	0.04	0.033
xD3	0.59	0.04	<.001
xD4	0.41	0.05	<.001
xD5	−0.01	0.07	0.859
xD6	−0.52	0.09	<.001
xD7	−0.14	0.11	0.208
xD8	0.11	0.19	0.550
Entering in grade 3			
xD1	−0.21	0.05	<.001
xD2	−0.08	0.04	0.075
xD3	0.51	0.04	<.001
xD4	−0.06	0.06	0.387
xD5	−0.38	0.07	<.001
xD6	−0.08	0.08	0.298
xD7	−0.29	0.12	0.015
xD8	0.03	0.18	0.852
Entering in grade 4			
xD1	−0.15	0.06	0.006
xD2	−0.24	0.05	<.001
xD3	0.11	0.07	0.131
xD4	−0.14	0.06	0.864
xD5	0.01	0.07	<.001
xD6	−0.43	0.08	0.901
xD7	−0.01	0.11	0.447
xD8	0.15	0.19	0.011
Entering in grade 5			
xD1	−0.14	0.05	0.011
xD2	−0.63	0.07	<.001
xD3	−0.07	0.07	0.286
xD4	0.13	0.06	0.045
xD5	−0.39	0.08	<.001
xD6	−0.49	0.09	<.001
xD7	−0.31	0.12	0.013
xD8	0.63	0.20	0.002

(continued)

Table E2. Fitted estimates for the relationship between grade of entry and logit hazard probability of reclassification in each period, 2003/04–2011/12 *(continued)*

Variable (main effects of time)	Coefficient (in logits)	Robust standard error	*p* value
Entering in grade 6			
xD1	−0.38	0.07	<.001
xD2	−0.82	0.08	<.001
xD3	0.11	0.06	0.065
xD4	−0.26	0.08	0.001
xD5	−0.65	0.07	<.001
xD6	−0.84	0.10	<.001
xD7	−0.24	0.17	0.162
xD8	0.45	0.32	0.158
Entering in grade 7			
xD1	−0.50	0.07	<.001
xD2	−0.50	0.08	<.001
xD3	−0.20	0.08	0.011
xD4	−0.49	0.07	<.001
xD5	−0.70	0.07	<.001
xD6	−0.84	0.13	<.001
xD7	−0.07	0.22	0.737
xD8	0.43	0.50	0.398
Entering in grade 8			
xD1	−0.15	0.07	0.032
xD2	−0.73	0.08	<.001
xD3	−0.11	0.07	0.126
xD4	−0.50	0.08	<.001
xD5	−0.72	0.11	<.001
xD6	−0.66	0.18	<.001
xD7	−0.21	0.44	0.636
Entering in grade 9			
xD1	−0.04	0.10	0.651
xD2	−0.56	0.07	<.001
xD3	0.21	0.07	0.004
xD4	−0.46	0.09	<.001
xD5	−0.27	0.12	0.028
xD6	−0.49	0.31	0.110
xD7	−0.14	0.74	0.847
Entering in grade 10			
xD1	−0.01	0.07	<.001
xD2	−0.11	0.09	<.001
xD3	0.10	0.11	<.001
xD4	−0.28	0.17	<.001
xD5	0.10	0.25	<.001
xD6	−0.02	0.84	<.001

(continued)

Table E2. Fitted estimates for the relationship between grade of entry and logit hazard probability of reclassification in each period, 2003/04–2011/12 *(continued)*

Variable (main effects of time)	Coefficient (in logits)	Robust standard error	*p* value
Entering in grade 11			
xD1	0.62	0.08	<.001
xD2	0.24	0.10	0.018
xD3	0.53	0.18	0.004
xD4	0.40	0.48	0.406
Entering in grade 12			
xD1	1.11	0.13	<.001
xD2	0.39	0.27	0.149
xD3	2.50	0.43	<.001

Note: $n = 229{,}249$. *D* refers to dummy variables representing time, with the number referring to year since school entry.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table E3 presents the findings for the relationships between initial English proficiency and probability of reclassification over time.

Table E3. Fitted estimates for the relationships between initial English proficiency and logit hazard probability of reclassification in each period, 2003/04–2011/12

Variable	Coefficient (in logits)	Robust standard error	*p* value
Main effects of time			
D1	–2.46	0.03	<.001
D2	–1.60	0.03	<.001
D3	–1.59	0.03	<.001
D4	–1.37	0.02	<.001
D5	–1.15	0.02	<.001
D6	–0.96	0.03	<.001
D7	–1.11	0.04	<.001
D8	–1.22	0.06	<.001
Centered initial English proficiency by time			
xD1	0.55	0.02	<.001
xD2	0.40	0.02	<.001
xD3	0.31	0.02	<.001
xD4	0.24	0.02	<.001
xD5	0.22	0.02	<.001
xD6	0.24	0.02	<.001
xD7	0.23	0.04	<.001
xD8	0.16	0.07	0.016

Note: $n = 229{,}249$. *D* refers to dummy variables representing time, with the number referring to year since school entry. This model also included dummy variables representing missing data in initial English proficiency.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table E4 presents the findings for the relationships between disability category and probability of reclassification over time.

Table E4. Fitted estimates for the relationship between disability category and logit hazard probability of reclassification in each period, 2003/04–2011/12

Variable	Coefficient (in logits)	Robust standard error	*p* value
Main effects of time (reference category = students without disabilities)			
D1	–2.06	0.03	<.001
D2	–1.45	0.03	<.001
D3	–1.39	0.02	<.001
D4	–1.00	0.03	<.001
D5	–0.97	0.03	<.001
D6	–0.78	0.03	<.001
D7	–0.88	0.04	<.001
D8	–0.92	0.06	<.001
Specific learning disability by time			
xD1	–0.82	0.05	<.001
xD2	–1.36	0.05	<.001
xD3	–1.36	0.05	<.001
xD4	–1.53	0.06	<.001
xD5	–1.32	0.05	<.001
xD6	–1.21	0.06	<.001
xD7	–0.95	0.07	<.001
xD8	–0.99	0.11	<.001
Speech or language impairment by time			
xD1	–0.23	0.04	<.001
xD2	–0.66	0.04	<.001
xD3	–0.94	0.05	<.001
xD4	–0.97	0.05	<.001
xD5	–0.78	0.05	<.001
xD6	–0.74	0.06	<.001
xD7	–0.80	0.09	<.001
xD8	–0.76	0.14	<.001
Other disability by time			
xD1	–0.02	0.07	0.749
xD2	–0.46	0.07	<.001
xD3	–0.70	0.09	<.001
xD4	–0.85	0.10	<.001
xD5	–0.91	0.11	<.001
xD6	–0.95	0.13	<.001
xD7	–1.57	0.24	<.001
xD8	–0.87	0.25	<.001

Note: *n* = 229,249. *D* refers to dummy variables representing time, with the number referring to year since school entry.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Appendix F. Supplemental analyses and results incorporating additional control variables

Supplemental analyses were conducted to investigate how robust the results are to the inclusion of control variables. In addition to controlling for demographic characteristics, retention in grade, and cohort, these models also included all three of the characteristics of interest for the second research question simultaneously. Thus, these results also clarify the extent to which the key findings for one of these characteristics were robust when controlling for the other two characteristics (for example, findings for initial English proficiency, when controlling for grade of entry and disability category) as well as other controls. These supplemental analyses indicated that results were robust to the inclusion of these control variables. Specifically, as noted in box 2, differences in median time to reclassification from these analyses were within half of a year of those discussed in the main report, and estimated percentages of students who become long-term English learner students were within approximately 4 percent of those discussed in the main report.

Demographic controls included family poverty, neighborhood poverty, race/ethnicity, home language, gender, and age. Family poverty (as proxied by student eligibility for the federal school lunch program) and neighborhood poverty (as proxied by average annual income in the census track in which students lived) were both treated as time-invariant variables with time-varying effects. Their values were based on the earliest available value, which corresponded to the year of entry to New York City (NYC) public schools for many but not all students. For students who did not have values on these variables for their first year of entry, data from later years were used. After this process, 3.7 percent of students still had missing values for family poverty and 7.4 percent of students had missing values for neighborhood poverty. These missing data were taken into account using a model-based approach in which a dummy variable represented the missing data and an arbitrary value was substituted for the missing value; including these together allowed the study team to retain the complete analytic sample. Gender, race/ethnicity, and home language were specified as time-invariant, categorical variables, with time-varying effects. There were no missing data for these variables. An additional variable was created to indicate students' chronological age of entry.[3] To capture variability in age of entry beyond that associated with grade of entry, a continuous variable was created that indicated students' age relative to the average age of entry for their grade of entry. Students' age was centered at the average for their grade of entry so that students entering at the average age for their grade had a value of 0, students entering older than the average for their grade had a positive value, and students entering younger than the average for their grade had a negative value. The mean was 0 years and the standard deviation was 0.55 years. There were no missing data for these variables. Beyond these demographic variables, additional controls included retention in grade once or twice (specified as time-invariant, categorical variables with time-varying effects) and cohort (defined by school year of entry into NYC public schools and specified as a time-invariant, categorical variable with time-varying effects). Descriptive statistics on these variables are included in appendix B.

To evaluate the statistical significance of the relationships between these control variables and the probability of reclassification, a discrete-time survival analytic model (DTSA model) was fitted in which these variables were added to equation C1 in appendix C. Subsequent DTSA models were fitted for each of the characteristics of interest (grade of entry, initial English proficiency, and disability category), following the approach described

in appendix C, but also controlling for these variables. As with the approach described in appendix C, hazard and survival functions were produced and used to estimate median time to reclassification and probability of becoming a long-term English learner student.

The tables in this appendix provide comparisons of results with and without controls for each student characteristic of interest, including grade of entry (table F1), initial English proficiency (table F2), and disability category (table F3). The fitted estimates for each of the models incorporating controls are also presented (tables F4–F7).

Tables F1–F3 compare the time to reclassification and probability of becoming a long-term English learner student between models with and without additional controls.

Table F1. Estimated median time to reclassification and percent of students who become long-term English learner students, by grade of entry to New York City public schools from models with and without controls, 2003/04–2011/12

Grade of entry as English learner student	Median time to reclassification (years)		Difference in median time to reclassification, relative to entry in kindergarten (years)		Percent of students who become long term English learner students[a]	
	With controls	Without controls	With controls	Without controls	With controls	Without controls
Kindergarten	3.89	3.79	na	na	29.5	25.4
1	3.36	3.14	−0.53	−0.65	23.4	23.4
2	3.15	2.93	−0.74	−0.86	23.7	23.7
3	3.42	3.53	−0.47	−0.26	26.7	26.7
4	4.13	3.74	0.24	−0.05	31.0	31.0
5	4.24	3.96	0.35	0.17	34.9	34.9
6	4.74	4.84	0.85	1.05	41.4	41.4
7	5.24	5.18	1.35	1.38	44.1	44.1

na is not applicable.

Note: The model with controls included all the control variables listed in table F4.

a. New York City Department of Education defines long-term English learner students as those students who have been classified as English learner students for six or more years.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table F2. Estimated median time to reclassification and percent of students who become long-term English learner students, by initial English proficiency from models with and without controls, 2003/04–2011/12

Initial English proficiency	Median time to reclassification (years)		Percent of students who become long term English learner students[a]	
	With controls	Without controls	With controls	Without controls
Below average	4.90	4.88	38.0	37.2
Average	3.99	4.06	29.7	27.9
Above average	3.09	3.19	21.6	19.1

Note: The model with controls included all the control variables listed in table F4.

a. New York City Department of Education defines long-term English learner students as those students who have been classified as English learner students for six or more years.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table F3. Median time to reclassification and percent of students that become long-term English learner students, by disability category from models with and without controls, 2003/04–2011/12

Disability category	Median time to reclassification (in years)		Percent of students who become long term English learner students[a]	
	With controls	Without controls	With controls	Without controls
Students without disabilities (n = 202,278)	3.49	3.56	22.2	21.9
Students with specific learning disabilities (n = 12,710)	8.07	7.60	63.4	62.9
Students with speech or language impairments (n = 11,643)	5.94	5.63	49.4	46.3

Note: The models with controls included all the control variables listed in table F4.

a. New York City Department of Education defines long-term English learner students as those students who have been classified as English learners for six or more years.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Tables F4–F6 show the fitted estimates from models incorporating additional controls.

Table F4. Fitted estimates for the relations between logit hazard probability of reclassification and control variables, 2003/04–2011/12

Variable	Coefficient (in logits)	Robust standard error	*p* value
Main effects of time			
D1	–2.46	0.04	<.001
D2	–1.77	0.04	<.001
D3	–1.08	0.03	<.001
D4	–1.23	0.03	<.001
D5	–1.20	0.04	<.001
D6	–0.97	0.04	<.001
D7	–1.21	0.05	<.001
D8	–1.16	0.06	<.001
Neighborhood average income	0.07	0.01	<.001
Not eligible for the federal school lunch program	0.10	0.02	<.001
Male	–0.20	0.01	<.001
Race/ethnicity			
Asian/Pacific Islander	0.09	0.04	.025
White, non-Hispanic	0.13	0.04	.002
African American	–0.05	0.04	.265
Native American	0.03	0.09	.747
Not reported	0.45	0.07	<.001
Multiracial/multiethnic	–0.18	0.20	.366
Home language			
Chinese	0.31	0.06	<.001
Another language	0.41	0.04	<.001
Bengali	0.51	0.05	<.001

(continued)

Table F4. Fitted estimates for the relations between logit hazard probability of reclassification and control variables, 2003/04–2011/12 (continued)

Variable	Coefficient (in logits)	Robust standard error	p value
Russian	0.77	0.06	<.001
Haitian Creole	0.23	0.07	.001
French	0.35	0.06	<.001
Korean	0.76	0.08	<.001
English	2.06	0.05	<.001
Southeast Asian languages	0.41	0.07	<.001
Cohort			
Cohort2005xD1	0.10	0.04	.014
Cohort2005xD2	0.44	0.04	<.001
Cohort2005xD3	−0.58	0.04	<.001
Cohort2005xD4	−0.09	0.04	.016
Cohort2005xD5	0.04	0.04	.332
Cohort2005xD6	0.06	0.05	.256
Cohort2005xD7	0.26	0.06	<.001
Cohort2006xD1	0.34	0.05	<.001
Cohort2006xD2	0.09	0.04	.014
Cohort2006xD3	−0.69	0.04	<.001
Cohort2006xD4	−0.22	0.04	<.001
Cohort2006xD5	0.33	0.04	<.001
Cohort2006xD6	0.31	0.05	<.001
Cohort2007xD1	0.41	0.06	<.001
Cohort2007xD2	0.02	0.04	.696
Cohort2007xD3	−0.73	0.04	<.001
Cohort2007xD4	0.07	0.04	.057
Cohort2007xD5	0.33	0.04	<.001
Cohort2008xD1	−0.34	0.05	<.001
Cohort2008xD2	0.17	0.05	<.001
Cohort2008xD3	−0.41	0.04	<.001
Cohort2008xD4	−0.01	0.04	.877
Cohort2010xD1	0.01	0.05	.877
Cohort2010xD2	0.26	0.04	<.001
Cohort2011xD1	0.17	0.05	<.001
RetainedOncexD1	−0.33	0.04	<.001
RetainedOncexD2	−0.89	0.04	<.001
RetainedOncexD3	−0.61	0.03	<.001
RetainedOncexD4	−0.94	0.03	<.001
RetainedOncexD5	−0.91	0.04	<.001
RetainedOncexD6	−0.72	0.04	<.001
RetainedOncexD7	−0.25	0.06	<.001
RetainedOncexD8	−0.27	0.09	.003
RetainedTwiceorMorexD1	−0.27	0.06	<.001
RetainedTwiceorMorexD2	−0.63	0.06	<.001
RetainedTwiceorMorexD3	−0.73	0.05	<.001
RetainedTwiceorMorexD4	−0.98	0.06	<.001

(continued)

Table F4. Fitted estimates for the relations between logit hazard probability of reclassification and control variables, 2003/04–2011/12 *(continued)*

Variable	Coefficient (in logits)	Robust standard error	*p* value
RetainedTwiceorMorexD5	−1.17	0.08	<.001
RetainedTwiceorMorexD6	−1.04	0.09	<.001
RetainedTwiceorMorexD7	−0.45	0.10	<.001
RetainedTwiceorMorexD8	−0.08	0.14	.572

Note: $n = 229{,}249$. Models also included dummy variables to represent missing data in the family poverty and neighborhood income variables. D refers to dummy variables representing time, with the number referring to year since school entry.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table F5. Fitted estimates for the relations between grade of entry and logit hazard probability of reclassification in each period after including additional controls, 2003/04–2011/12

Variable	Coefficient (in logits)	Robust standard error	*p* value
Entering in kindergarten (reference category)			
D1	−2.43	0.05	<.001
D2	−1.62	0.04	<.001
D3	−1.21	0.04	<.001
D4	−1.27	0.04	<.001
D5	−1.16	0.04	<.001
D6	−0.85	0.04	<.001
D7	−1.15	0.06	<.001
D8	−1.22	0.07	<.001
Entering in grade 1			
xD1	0.33	0.04	<.001
xD2	−0.29	0.04	<.001
xD3	0.38	0.03	<.001
xD4	0.38	0.04	<.001
xD5	0.32	0.04	<.001
xD6	−0.15	0.07	.020
xD7	−0.30	0.11	.008
xD8	0.27	0.14	.051
Entering in grade 2			
xD1	−0.21	0.05	<.001
xD2	−0.10	0.05	.021
xD3	0.64	0.04	<.001
xD4	0.48	0.05	<.001
xD5	0.04	0.07	.542
xD6	−0.52	0.09	<.001
xD7	−0.13	0.11	.256
xD8	0.15	0.19	.454

(continued)

Table F5. Fitted estimates for the relations between grade of entry and logit hazard probability of reclassification in each period after including additional controls, 2003/04–2011/12 *(continued)*

Variable	Coefficient (in logits)	Robust standard error	*p* value
Entering in grade 3			
xD1	−0.24	0.05	<.001
xD2	−0.10	0.04	.029
xD3	0.52	0.04	<.001
xD4	−0.04	0.06	.530
xD5	−0.38	0.07	<.001
xD6	−0.08	0.07	.263
xD7	−0.28	0.12	.024
xD8	−0.01	0.18	.979
Entering in grade 4			
xD1	−0.19	0.06	.001
xD2	−0.28	0.05	<.001
xD3	0.11	0.07	.084
xD4	−0.13	0.06	.027
xD5	0.08	0.07	.225
xD6	−0.36	0.08	<.001
xD7	0.05	0.12	.664
xD8	0.15	0.20	.456
Entering in grade 5			
xD1	−0.16	0.05	.004
xD2	−0.63	0.06	<.001
xD3	−0.05	0.06	.407
xD4	0.19	0.06	.002
xD5	−0.30	0.07	<.001
xD6	−0.40	0.09	<.001
xD7	−0.28	0.13	.031
xD8	0.73	0.20	<.001
Entering in grade 6			
xD1	−0.39	0.07	<.001
xD2	−0.83	0.07	<.001
xD3	0.16	0.06	.005
xD4	−0.19	0.07	.007
xD5	−0.55	0.07	<.001
xD6	−0.83	0.10	<.001
xD7	−0.15	0.18	.406
xD8	0.52	0.31	.096
Entering in grade 7			
xD1	−0.51	0.06	<.001
xD2	−0.47	0.07	<.001
xD3	−0.11	0.07	.139
xD4	−0.39	0.07	<.001
xD5	−0.64	0.07	<.001
xD6	−0.64	0.14	<.001

(continued)

Table F5. Fitted estimates for the relations between grade of entry and logit hazard probability of reclassification in each period after including additional controls, 2003/04–2011/12 *(continued)*

Variable	Coefficient (in logits)	Robust standard error	*p* value
xD7	−0.02	0.22	.930
xD8	0.47	0.54	.384
Entering in grade 8			
xD1	−0.15	0.07	.038
xD2	−0.69	0.07	<.000
xD3	−0.03	0.07	.715
xD4	−0.43	0.08	<.001
xD5	−0.35	0.12	.004
xD6	−0.34	0.19	.076
xD7	−0.15	0.46	.750
Entering in grade 9			
xD1	−0.08	0.09	.411
xD2	−0.52	0.06	<.001
xD3	0.30	0.06	<.001
xD4	−0.08	0.09	.373
xD5	0.20	0.14	.133
xD6	−0.23	0.32	.466
xD7	−0.26	0.74	.722
Entering in grade 10			
xD1	−0.14	0.07	.036
xD2	−0.07	0.09	.445
xD3	0.45	0.09	<.001
xD4	0.20	0.18	.260
xD5	0.69	0.28	.014
xD6	0.42	0.82	.609
Entering in grade 11			
xD1	0.41	0.07	<.001
xD2	0.32	0.11	.002
xD3	0.71	0.21	.001
xD4	0.67	0.57	.241
Entering in grade 12			
xD1	0.80	0.14	<.001
xD2	0.18	0.28	.526
xD3	2.54	0.37	<.001
Centered age of entry	−0.05	0.01	<.001

Note: $n = 229,249$. D refers to dummy variables representing time, with the number referring to year since school entry. This model also included all the control variables listed in table F4. Their coefficients in this model were roughly similar in magnitude to the estimates presented in table F4.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table F6. Fitted estimates for the relations between initial English proficiency and logit hazard probability of reclassification in each period after including additional controls, 2003/04–2011/12

Variable	Coefficient (in logits)	Robust standard error	*p* value
Main effects of time			
D1	−2.68	0.05	<.001
D2	−1.68	0.04	<.001
D3	−1.21	0.04	<.001
D4	−1.25	0.04	<.001
D5	−1.13	0.04	<.001
D6	−0.81	0.04	<.001
D7	−1.11	0.06	<.001
D8	−1.17	0.08	<.001
Centered initial English proficiency by time			
xD1	0.58	0.02	<.001
xD2	0.39	0.01	<.001
xD3	0.27	0.01	<.001
xD4	0.22	0.01	<.001
xD5	0.17	0.02	<.001
xD6	0.18	0.03	<.001
xD7	0.20	0.04	<.001
xD8	0.18	0.07	.009

Note: $n = 229,249$. D refers to dummy variables representing time, with the number referring to year since school entry. This model also included dummy variables representing missing data in initial English proficiency, all the control variables listed in table F4, and the grade of entry variables listed in table F5. The coefficients for the control variables and the grade of entry variables in this model were roughly similar in magnitude to the estimates presented in tables F4 and F5.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Table F7 shows the relationship between disability category and probability of reclassification over time, after including control variables.

Table F7. Fitted estimates for the relationship between disability category and logit hazard probability of reclassification in each period after including additional controls, 2003/04–2011/12

Variable	Coefficient (in logits)	Robust standard error	*p* value
Students without disabilities (reference category)			
D1	−2.39	0.05	<.001
D2	−1.50	0.04	<.001
D3	−1.06	0.04	<.001
D4	−1.07	0.04	<.001
D5	−0.93	0.04	<.001
D6	−0.55	0.04	<.001
D7	−0.80	0.06	<.001
D8	−0.85	0.08	<.001

(continued)

Table F7. Fitted estimates for the relationship between disability category and logit hazard probability of reclassification in each period after including additional controls, 2003/04–2011/12 *(continued)*

Variable	Coefficient (in logits)	Robust standard error	*p* value
Specific learning disability by time			
xD1	–0.73	0.05	<.001
xD2	–1.27	0.05	<.001
xD3	–1.31	0.05	<.001
xD4	–1.54	0.05	<.001
xD5	–1.36	0.05	<.001
xD6	–1.30	0.06	<.001
xD7	–1.04	0.07	<.001
xD8	–1.01	0.12	<.001
Speech or language impairment by time			
xD1	–0.20	0.05	<.001
xD2	–0.72	0.04	<.001
xD3	–0.87	0.05	<.001
xD4	–1.02	0.05	<.001
xD5	–0.92	0.05	<.001
xD6	–0.93	0.06	<.001
xD7	–0.90	0.09	<.001
xD8	–0.78	0.15	<.001
Other disability by time			
xD1	–0.002	0.07	0.934
xD2	–0.46	0.07	<.001
xD3	–0.68	0.09	<.001
xD4	–0.90	0.10	<.001
xD5	–1.02	0.11	<.001
xD6	–1.05	0.14	<.001
xD7	–1.62	0.24	<.001
xD8	–0.89	0.25	<.001

Note: *n* = 229,249. This model also included all the control variables listed in table F4 and the grade of entry variables listed in table F5. D refers to dummy variables representing time, with the number referring to year since school entry. The coefficients for these variables in this model that were roughly similar in magnitude to the estimates presented in tables F4 and F5.

Source: Authors' analysis based on New York City Department of Education enrollment data and Bilingual Education Student Information Survey data for 2003/04–2011/12.

Notes

1. "Specific learning disability means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia" (Individuals with Disabilities Act of 1975 [2004]). "Speech or language impairment means a communication disorder, such as stuttering, impaired articulation, a language impairment, or a voice impairment, that adversely affects a child's educational performance" (Individuals with Disabilities Act of 1975 [2004]).

2. According to the New York City Department of Education, BESIS is the official annual survey of students identified as English learner students and eligible for bilingual or English as a second language services.

3. Preliminary analyses indicated a small number of students with outlier values for this variable. Some of these outlier values appeared to be due to data entry mistakes, so values that were plus or minus 3 standard deviations from the mean of their grade level were recoded to the value at plus or minus 3 standard deviations. In addition, 5.5 percent of students (*n* = 11,439) entered NYC schools at ages above 18, which may indicate students entering NYC schools and receiving services despite being adults; given that there is little reason to believe that variability in age among adults is relevant to these analyses, these values were recoded to be 18 years of age, and an additional dichotomous variable indicating that the student entered as an adult of 18 years of age or over was also created. In subsequent analyses, this variable was not found to significantly relate to the probability of reclassification, after controlling for other variables, so it was not included in the models presented in appendixes C and E.

References

August, D. E., & Hakuta, K. E. (1997). *Improving schooling for language-minority children: A research agenda.* Washington, DC: National Academy Press. http://eric.ed.gov/?ID=ED408820

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth.* Mahwah, NJ: Lawrence Erlbaum Associates. http://eric.ed.gov/?ID=ED553370

Bailey, A. L., & Kelly, K. R. (2013). Home language survey practices in the initial identification of English learners in the United States. *Educational Policy, 27*(5), 770–804.

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

Center for Equity and Excellence in Education & Clearinghouse on Assessment and Evaluation. (1982). Language assessment battery—English. New York, NY: New York City Board of Education. Abstract retrieved from Education Resources Information Center Clearinghouse on Assessment and Evaluation. Retrieved April 7, 2016, from http://ericae.net/eac/eac0127.htm.

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. http://eric.ed.gov/?ID=ED501661

Cook, H. G., Boals, T., & Lundberg, T. (2011). Academic achievement for English learners: What can we reasonably expect? *Phi Delta Kappan, 93*(3), 66–69. http://eric.ed.gov/?ID=EJ964382

Cook, H. G., 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, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. http://eric.ed.gov/?ID=ED531979

Every Student Succeeds Act. (2015). Pub. L. No. 114–95, Stat. 1177.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course.* New York, NY: Routledge.

Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* (University of California Linguistic Minority Research Institute Policy Report 2000-01). Santa Barbara, CA: University of California–Santa Barbara. http://eric.ed.gov/?ID=ED443275

Individuals with Disabilities Act of 1975, 34 C.F.R. §300.8 (2004).

Kieffer, M. J. (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology, 100*(4), 851–868. http://eric.ed.gov/?ID=ED531979

Klingner, J. K., Artiles, A. J., & Barletta, L. M. (2006). English language learners who struggle with reading: language acquisition or LD? *Journal of Learning Disabilities, 39*(2), 108–128.

Klingner, J. K., Artiles, A. J., Kozleski, E., Harry, B., Zion, S., Tate, W., et al. (2005). Addressing the disproportionate representation of culturally and linguistically diverse students in special education through culturally responsive educational systems. *Education Policy Analysis Archives, 13*(38), 1–43. http://eric.ed.gov/?ID=EJ846743

Klingner, J. K. (n.d). Dist*inguishing language acquisition from learning disabilities.* New York, NY: New York City Department of Education. Retrieved, April 7, 2016, from http://schools.nyc.gov/NR/rdonlyres/DABEF55A-D155–43E1-B6CB-B689FBC9803A/0/Language AcquisitionJanetteKlingnerBrief_73015.pdf.

Linquanti, R., & Cook, H. G. (2013). *Toward a "common definition of English learner": A brief defining policy and technical issues and opportunities for state assessment consortia.* Washington, DC: Council of Chief State School Officers. http://eric.ed.gov/?ID=ED542705

Marinova-Todd, S. H., Marshall, D. B., & Snow, C. E. (2000). Three misconceptions about age and L2 learning. *Tesol Quarterly, 34*(1), 9–34. http://eric.ed.gov/?ID=EJ603514

New York City Department of Education. (2011). Demographic snapshots. New York, NY: Author. Retrieved April 7, 2016, from http://schools.nyc.gov/Accountability/data/default.htm.

New York City Department of Education. (2013). *Office of English language learners 2013 demographic report.* New York, NY: Author.

New York State Education Department. (2011). *Statewide totals as of October 5, 2011 by disability, county, and school district.* New York, NY: Author. Retrieved April 7, 2016, from http://p1232.nysed.gov/sedcar/state.html.

New York State Education Department. (2015). Technical reports. New York, NY: Author. Retrieved April 7, 2016, from http://www.p12.nysed.gov/assessment/reports/.

Olsen, L. (2010). *Reparable harm: Fulfilling the unkept promise of educational opportunity for California's long term English learners.* Long Beach, CA: Californians Together.

Orozco, M. J., Almanza, E., de Onis, C., Klingner, J. K., & Hoover, J. J. (2008). Distinguishing between language acquisition and learning disabilities among English language learners. In J. K. Klingner, J. Hoover & L. Baca (Eds.), *Why do English language learners struggle with reading? Distinguishing language acquisition from learning disabilities* (pp. 5–16). Thousand Oaks, CA: Corwin.

Parker, C. E., O'Dwyer, L. M., & Irwin, C. W. (2014). *The correlates of academic performance for English language learner students in a New England district* (REL 2014–020).

Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. http://eric.ed.gov/?ID=ED546480

Pearson. (2009). English as a second language achievement test (NYSESLAT) 2008 administration technical manual. New York, NY: Author. Retrieved April 7, 2016, from http://www.p12.nysed.gov/assessment/reports/nyseslat-techrep-08.pdf.

Ragan, A., & Lesaux, N. (2006). Federal, state, and district level English language learner program entry and exit requirements: Effects on the education of language minority learners. *Education Policy Analysis Archives, 14*(20), 1–32. http://eric.ed.gov/?ID=EJ806048

Short, D. J., & Fitzsimmons, S. (2007). *Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners: A report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.

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

Slama, R. B. (2012). A longitudinal analysis of academic English proficiency outcomes for adolescent English language learners in the United States. *Journal of Educational Psychology, 104*(2), 265–285. http://eric.ed.gov/?ID=EJ991176

Thompson, K.D. (2015). English learners' time to reclassification: An analysis. *Educational Policy*. Advance online publication.

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. http://eric.ed.gov/?ID=EJ1041160

The Regional Educational Laboratory Program produces 7 types of reports

Making Connections
Studies of correlational relationships

Making an Impact
Studies of cause and effect

What's Happening
Descriptions of policies, programs, implementation status, or data trends

What's Known
Summaries of previous research

Stated Briefly
Summaries of research findings for specific audiences

Applied Research Methods
Research methods for educational settings

Tools
Help for planning, gathering, analyzing, or reporting data or research