APPENDIX II METHODOLOGY SPECIFIC TO *A CLOSER LOOK AT HIGH FLYERS IN HIGH-POVERTY SCHOOLS*

While the main findings of the report focused on the analysis of Cohorts 1 and 2, the sidebar titled *A Closer Look at High Flyers in High-Poverty Schools* introduced a separate line of inquiry. This analysis used a new definition of high flyers, and centered on two distinct cohorts (Cohorts 3 and 4). Here, a high achiever was defined as a student who performed in the top 10 percent *of his particular grade and school*.

Students in **Cohort 3** were followed from third grade through fifth grade. This cohort consisted of 235,709 students in math, of whom 21,291 were high flyers, and 250,550 students in reading, of whom 22,868 were high flyers. The students were drawn from 952 schools in thirty states.

Students in **Cohort 4** were followed from sixth grade through eighth grade. This cohort consisted of 184,674 students in math, of whom 17,425 were high flyers, and 210,577 students in reading, of whom 20,309 were high flyers. The students were drawn from 410 schools in twenty-nine states.

To assure that these students reflected their larger school populations, the sample included only those students who attended a school in which 80 percent of enrolled students overall and a minimum of thirty students in each grade were tested with NWEA assessments. In order to increase the sample size, we combined students from three successive years into each cohort (Figure A-1). For example, if a school had eight high achievers in third grade in 2005-06, six high achievers in third grade in 2006-07, and nine high achievers in third grade in 2007-08, we analyzed the total group of students, twenty-three high achievers, as a single entity (Figure A-1).

FIGURE A-1

Cohorts 3 and 4 by Grade and Year

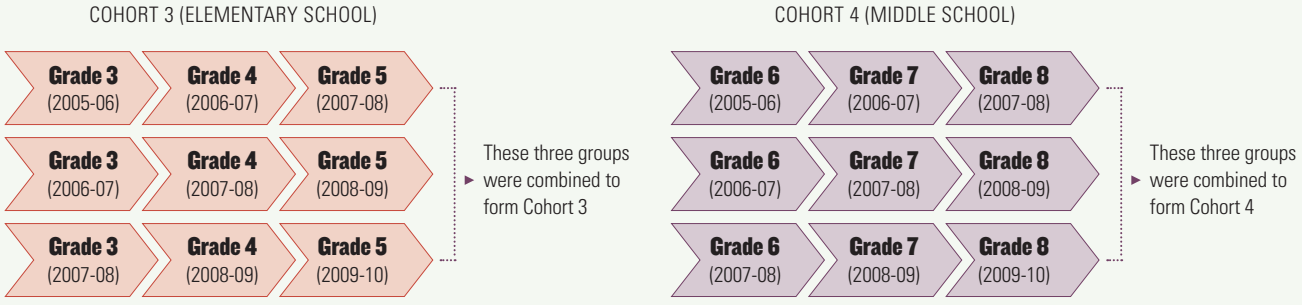

Table A-6 illustrates the gender and ethnic composition of the two cohorts. The proportions of males to females in Cohorts 3 and 4 were similar to the gender compositions of Cohorts 1 and 2, but the proportions of minority students in Cohorts 3 and 4 were higher than the proportions reflected in Cohorts 1 and 2. This is largely due to the two distinct definitions of "high achiever." Because high achievers in Cohorts 3 and 4 comprised those students who performed in the top 10 percent of their individual grades and schools, we expected a higher representation of minority students in those cohorts.

TABLE A-6

Demographics of High Achievers in Cohorts 3 and 4 (Initial Year of Study)

	Gender		Ethnicity		School Poverty	
	Female	Male	Minority	Non-minority	High Poverty	Low Poverty
Cohort 3 Math	39.5%	60.5%	16.2%	83.8%	27.8%	72.2%
Cohort 3 Reading	52.3%	47.7%	18.0%	82.0%	26.8%	73.2%
Cohort 4 Math	40.1%	59.9%	12.5%	87.5%	23.7%	76.3%
Cohort 4 Reading	51.7%	48.3%	13.8%	86.2%	22.9%	77.1%

Hierarchical Linear Modeling

This separate line of analysis employed hierarchical linear modeling (HLM) to evaluate the results for high achievers in high- and low-poverty schools.[15] HLM is an advanced form of linear regression. It is preferred over standard linear regression in circumstances in which data are nested, while linear regression might be preferred in circumstances where data are independent. For example, in a simple study in which one wanted to assess the relationship between the morning calorie intake and stamina of a group of recreational runners, simple regression would normally be sufficient, because the runners are not nested in a set of groupings (like running teams, for example).

These conditions are rarely present in educational data. In this particular study, there is not one single assessment but rather a series of assessments that are *nested*, if you will, within each student. In addition, each student nests within a grade level and school.[16] In these conditions, the individual test events of a student are likely to be highly correlated with one another, and the test events of all students within a school are also likely to be correlated. By accounting for nesting within this analysis, we can draw better inferences about the relationship between our high-achieving students' growth and factors such as gender, ethnicity, or school poverty rate that may influence performance and growth. Another reason for using HLM is that schools vary greatly in maintaining their high achievers' growth. Hence, by accounting for school-level random effects (variability across schools), the estimation of fixed effects (the effects of school poverty and location) are more precise.

The following three-level HLM was used to model the relationship between school poverty rate and school achievement and growth. It applied to Cohorts 3 and 4, separately examining performance and growth rates for elementary school mathematics, elementary school reading, middle school mathematics, and middle school reading. In the model, school poverty rate refers to the percentage of students who are eligible for free or reduced-price lunch at each school.

Level One: Test events (repeated measures)

Level one is an individual growth model of academic achievement at time *t* for student *i* in school *j*.

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}\,(ACADEMIC\ YEAR)_{tij} + e_{tij}$$

15 Additional models that examine characteristics other than school poverty will be described in a forthcoming report.

16 In truth, the student is nested inside a classroom, a grade level, *and* a school. One limitation of the data set used for this study is that we could not consider classroom effects because we did not have reliable data about the particular classrooms to which students were assigned.

Level Two: Students (individual growth trajectory)

We did not include any student-level variables to focus on the relationship between school poverty rate and school achievement and growth.

$$\pi_{0ij} = \beta_{00j} + r_{0ij},$$

$$\pi_{1ij} = \beta_{10j} + r_{1ij},$$

Level Three: Schools

$$\beta_{00j} = \gamma_{000} + \gamma_{001} X_{(FRL\%)j} + \mu_{00j}$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101} X_{(FRL\%)j} + \mu_{10j}$$