NYUWagner

Robert F. Wagner Graduate School of Public Service

Can Reorganizing K-8 Education Improve Academic Performance? The Impact of Grade Span on Student Achievement

NYU Wagner Research Paper No. 2010-06

Amy Ellen Schwartz
Leanna Stiefel
Ross Rubenstein
Jeffrey Zabel

Can Reorganizing K-8 Education Improve Academic Performance? The Impact of Grade Span on Student Achievement

Amy Ellen Schwartz
Steinhardt School and Wagner Graduate School
New York University

Leanna Stiefel
Wagner Graduate School
New York University

Ross Rubenstein
Maxwell School
Syracuse University

Jeffrey Zabel
Tufts University

Abstract

Although rearranging school organizational features is a perennially popular method of school reform, no consensus exists regarding the best organizational structure for educating students. Instead, there is wide variation in school organization across the United States and even within school districts. This paper examines how student performance in eighth grade is shaped by the grade span configuration. Exploiting the variation in New York City, we estimate the impact of several distinct grade span configurations, controlling for school and student characteristics including third grade performance. We find that students in schools with longer grade spans (K-8 schools particularly) outperform students in schools with shorter grade spans. Thus, creating more K-8 schools and minimizing structural articulations in the K-8 years may improve student performance.

We thank Luis Chalico for excellent research assistance. The Institute for Education Sciences, US Department of Education grant R305E040096 provided support for work leading to this paper. We are responsible for all analyses and conclusions.

Can Reorganizing K-8 Education Improve Academic Performance? The Impact of Grade Span on Student Achievement

Abstract

Although rearranging school organizational features is a perennially popular method of school reform, no consensus exists regarding the best organizational structure for educating students. Given the variation in school organization across and even within school districts, this paper examines how student performance in 8th grade is shaped by grade span configuration. Exploiting the variation in New York City, we estimate the impact of distinct grade span configurations on 8th grade math and reading performance, controlling for school and student characteristics including 3rd grade performance. We find that students in schools with longer grade spans (K-8 schools particularly) outperform students in schools with shorter grade spans. Thus, creating more K-8 schools, minimizing structural articulations, and reducing mobility may improve student performance.

Keywords: School reform; grade span; mobility; student achievement

Can Reorganizing K-8 Education Improve Academic Performance? The Impact of Grade Span on Student Achievement

INTRODUCTION

Rearranging school organizational features is a perennially popular method of school reform that promises improvements in student achievement at relatively low cost. Recently, for example, school districts around the country are re-organizing high schools into small schools and schools-within-schools in the hope of boosting graduation rates.[1] Changing the grade configurations of elementary and middle schools is another popular approach, with school systems including New York City, Boston and Philadelphia increasingly moving away from a combination of self-contained middle schools and self-contained elementary schools toward schools that serve kindergarten through eighth grade (K-8).[2] Reorganizing the grade spans of elementary and middle schools is a particularly attractive reform mechanism because it is relatively inexpensive compared to other organizational reforms such as reducing class size and, undoubtedly, because it is a "lever of change" within the control of school districts themselves.

It is, perhaps, unsurprising that there has been a significant transformation in the grade span structure of schools in the United States over the past 40 years. In this time, for example, self-contained middle schools housing combinations of grades 6-8 became increasingly popular around the country, spurred by concerns over meeting the unique needs of younger adolescents (see for example, Carnegie Task Force, 1989). Between 1970-71 and 2001-2002, the proportion

[1] It is noteworthy that large, comprehensive high schools were created in earlier reforms with similar intentions. See Hammack (2004) for more on this.

[2] The adoption of a K-8 model has become a popular education reform approach in a wide range of school districts, including New York City, Cleveland, Cincinnati, Philadelphia and Boston (DeJong and Craig, 2002; Gootman, 2007; Jonas 2007) Peoria, Illinois (Jellick, 2005) and Dayton, Ohio (Bowie, 2007).

of middle schools rose from 2.3 percent to almost 13 percent, while junior high schools fell from 8.7 percent to 3.9 percent of secondary schools (National Center for Education Statistics, 2004). [3]

No consensus exists, however, regarding the best organizational structure for educating students in grades one through eight. Instead, there is wide variation in school organization across the United States and even within school districts. In 2004-5, for example, elementary schools ending in the 5th grade were the most common (35.6 percent), followed by schools ending in the 6th grade (19.7 percent), while a variety of alternative forms comprised the rest (NCES, 2007).[4] Nationally, 19 percent of elementary schools serve grades 4, 5, or 6 to 6, 7 or 8, while the remainder combines the middle grades with lower grades or other configurations. Over 28 percent of 6th graders in 2001 were not in separate schools, but attended K-6 or K-8 schools (Bedard and Do, 2005).

Despite the growing popularity of grade reconfiguration among policymakers and its appeal to education reformers, researchers have noted that the evidence on grade span effects is "seriously wanting" (Howley, 2002) and suffers from a "dearth of empirical research" (Renchler, 2000). In this paper, we begin to fill the gap in knowledge using data on students and schools in New York City to examine the ways in which student performance in the 8th grade is shaped by the grade span configuration of the schools attended. In addition, we explore the effects of school transitions on student performance and the differences between mandatory and 'voluntary' school changes. The present study adds to our understanding of grade span configuration in several important respects. First, we exploit the variation in grade configuration found in the nation's largest school district (New York City) to identify grade span effects for a variety of configurations, controlling for a wide array of observable school and student

[3] Middle schools are defined by the NCES as schools beginning at grade 4, 5, or 6 and ending at grade 6, 7, or 8. Junior high schools are defined as schools serving grade 7 through grade 8 or 9.
[4] The lowest grade in these schools was prekindergarten, kindergarten or 1st grade.

characteristics. Second, we take advantage of unique longitudinal student data that allow us to examine the performance of cohorts of students through five years of schooling as they transition across grades and across schools, rather than at a single point in time. Third, we measure gains or "value-added" in student academic performance between the 3rd grade and the 8th grade.

National and international evidence has consistently shown that students tend to lose ground between the earlier grades and 8th grade, particularly in math. For example, on the 2005 National Assessment of Educational Progress (NAEP), 36 percent of 4th graders were rated at or above proficiency in math, as compared to 30 percent of 8th graders. While proficiency levels generally increased between 4th and 8th grade until 2000, this trend reversed beginning in 2003, pointing to the importance of improving the performance of students as they prepare to enter secondary school (NCES, 2007). Allensworth (2005) and Roderick (2006) provide evidence that improving academic achievement in 8th grade is critical to improving high school graduation rates. Thus, if re-configuring the grade spans of elementary and middle schools can improve 8th grade performance, it may also serve to improve high school graduation rates.

As the largest school system in the United States, New York City provides an excellent natural laboratory for this study. There are a large number of elementary and middle schools (over 915), with many different grade spans (approximately 28). In addition, all the schools are located in one district, so we do not have district-level variation within a given year in tax base or other unobserved district-level determinants of performance. More generally, the schools share a common set of institutions, policy changes and macro-economic and social environmental conditions. Finally, we are able to construct a unique longitudinal student-level database tracking cohorts of students in the New York City public schools in elementary and middle school years.

To preview the results, we find evidence that 8th graders in K-8 schools perform particularly well on math and reading tests relative to 8th graders in differently configured schools. Further, our results suggest that this higher performance may be driven partly by school stability - that is, by differences in school mobility. Importantly, we see little difference between the effects of moves made at institutional articulation grades (i.e., transitions following completion of the highest grade housed in a school and subsequent enrollment in the lowest grade in a new school) and moves made in non-standard years (i.e., leaving a school before completing its highest grade or enrolling in a new school at a grade other than its lowest). Thus, our results suggest that student performance in the 8th grade, an important indicator of high school readiness, may be improved by configuring elementary and middle school education to minimize school transitions. Policymakers should consider K-8 schools, in particular, as part of their school organizational plans.

The remainder of the paper proceeds as follows. The next section presents the underlying theory linking grade span configuration to student performance, describes recent grade span policy initiatives and reviews previous research on grade span effects. This is followed by an overview of the data used in the analyses and our model of the production of education outcomes. Discussion of the empirical results, including a series of sensitivity analyses, follows. A final section presents conclusions and policy implications.

GRADE SPAN AND STUDENT PERFORMANCE

Why might grade span organization affect student performance? As shown in Figure 1, school grade span can affect student performance through a variety of channels. First, the number and span of grades served in a school may have important impacts on school practices,

curriculum or other features of the school. For example, a longer grade span may allow greater coordination among teachers and administrators across grades, resulting in more cohesive and aligned curricular planning and more effective long-term strategies to address student needs. Anecdotal evidence often supports these hypotheses (see, for example, Gootman, 2007). Conversely, a shorter grade span may allow schools to better specialize in addressing the unique needs of students at specific ages (see for example, Carnegie Task Force, 1989). Differences in grade span may also imply differences in school size or cohort size, which may in turn influence student performance. Author (2008), for example, find that schools with longer grade spans have fewer students in each grade.

Second, school grade span determines the structural transitions or "articulation" points at which students move from one school to the next. The evidence on these transitions generally finds that students experience difficulties when moving to schools serving higher grade levels and when moving from the self-contained classrooms typically found in lower grades to schools in which they change classrooms during the day. Switching schools may serve to lower academic achievement and increase the risk of dropping out (Seidman, *et al*, 1994; Alspaugh and Harting, 1995; Alspaugh, 1998, Hanushek, Kain, and Rivkin, 1999). Both the number of transitions and the timing (the grade level at which the articulation occurs) could affect student performance. In addition, Wampler, Munsch and Adams (2002) suggest that this impact may differ across students from different racial and ethnic groups.

Third, grade span may shape performance because of its impact on the peers that a student encounters in school. Schools with longer grade spans will serve a greater diversity of ages, with a complex array of behavioral effects possible. As an example, if older students see younger students as "mentees," a longer grade span may improve older students' behavior and

produce valuable role models for younger students. On the other hand, if greater age diversity provides more opportunities for older students to bully or harass younger students, longer grade spans may produce negative effects. Note that these effects could differ by grade. While younger students may benefit from a schooling environment without older students, the older students themselves may benefit from greater age diversity, a lower concentration of students facing difficult adolescent transitions, and a more orderly school environment (MacIver and MacIver, 2006). Of course, these different impacts may magnify or offset each other over the course of the student's elementary and middle school career such that the cumulative effect in the eighth grade may be positive or negative (or null).

Despite the great diversity in school organization across the United States, there is little empirical research examining the effects of various school organizational models. Much of the extant research on grade span has focused on the middle school grades and suggests that student performance in the middle grades may suffer when these grades are housed in a separate school, or when they are included in secondary schools rather than elementary schools. For example, Bedard and Do (2005), using national data for 1987 to 1994, concluded that middle schools (grades 6-8) were associated with lower on-time graduation rates, perhaps by as much as three percent for an all-middle school structure. MacIver and MacIver (2006) studied Philadelphia's movement to K-8 schools, which occurred contemporaneously with a contracting out initiative, and found some evidence that students in older K-8 schools, but not in recently-converted schools, outperformed those in middle schools. Other studies of grade organization have examined a range of academic and non-academic outcomes, including student grades, attendance, suspensions, student perceptions of school safety and students' self-esteem. Weiss and Kipnes (2006) found higher self-esteem and perceived school safety among students

attending K-8 schools as compared to middle schools. Cook, *et al*. (2008) found that attending 6th grade in a middle school rather than an elementary school significantly increased the number of reported student infractions, and that this difference persisted into high school for black students.

In sum, much of the previous research on grade span effects has relied on cross-sectional comparisons across schools with varying organizational structures. In contrast, our analysis treats the effect of grade span as cumulative across the elementary and middle school grades. Our longitudinal data also allow us to disentangle the effects of grade span from the effects of changing schools, a distinction often overlooked in earlier research. Finally, by using detailed student-level data we are able to control for a wide array of other student and school characteristics that may affect student performance in eighth grade. Next we describe the data and our model of education production.

DATA AND MODEL

Data

This study makes use of longitudinal student level data on students registered in grades 4 through 8 in a New York City's public school between 1995-96 and 2001-2002. The data, provided by the New York City Department of Education (NYCDOE), include detailed information on student characteristics, including performance on citywide reading and math tests, free and reduced price lunch eligibility, race, age, gender, immigrant status, language spoken at home and eligibility for part-time special education services. We also include indicator variables to control for any changes in the configuration of the school each student attended over the period.

Since our interest is in understanding the way that student performance in 8th grade reflects the school "path" a student follows, we examine the grade spans of the schools each student has attended over the five-year period prior to 8th grade, focusing on a set of continuously registered students who moved sequentially through the appropriate grades. To be specific, our "Standard Academic Progress" (SAP) cohort includes students attending 4th grade in the 1997-98 school year, 5th grade the following year, 6th grade next, followed by 7th grade and, finally, 8th grade in the 2001-2002 school year. There are 40,286 students in the 98-02 SAP cohort and 41,120 students in a similarly constructed 1997-01 SAP cohort. In a later section of the paper, we perform sensitivity analyses in which we analyze possible biases that result from students who exit the SAP cohort.

We track each student's academic path, the grade configuration of the schools each student attends and the articulation points (if any) from one school to another and define a set of academic paths characterized by a sequence of schools of particular grade span configurations that a student might attend between 4th and 8th grades. The simplest academic path, for example, would entail continuous enrollment in a K-8 school between 4th grade and 8th grade. A slightly more complicated path would include enrollment in a K-5 school in 4th and 5th grades, followed by enrollment in a 6-8 school in 6th, 7th and 8th grades. Notice that a student following this path articulated "on time," meaning that he or she attended the same school through the final year offered (here, grade 5), "graduating" as expected, and then enrolling in a new school at the natural articulation or entry grade.

A more complex path would entail attendance in a K-6 school for 4th and 5th grades, followed by enrollment in 6th through 8th grade in a K-8 school. In this path, the student "exits early" and joins the school "midstream." In our analyses, we examine both the grade span path

and the articulation points (that is, the grade and time at which a student moves from one school to another) and also consider whether (and how) these differences in types of articulation matter.[5] Notice that, in some cases, schools change grade spans, such that students can experience a change in their school grade span without changing schools. As an example, students enrolled in a K-5 for 3rd and 4th grade may find themselves enrolled in a K-6 in 5th and 6th grade if their K-5 school expands to offer a 6th grade. Thus, we also construct variables that allow us to distinguish between changes that involve switching schools and changes that involve switches in the configuration of existing schools.

Model

To examine the effects of grade span paths on student performance in the 8th grade, we estimate a model of 8th grade achievement as a function of 3rd grade performance, student and school characteristics, and grade span paths, using a standard long-term education production function (EPF) model:

$$Y_{is8} = \beta_0 + \beta_1 Y_{is3} + \beta_2 Socio_{i8} + \beta_3 SC_{s8} + \beta_4 GS_i + \beta_5 Switch_i + e_{is8} \quad (1)$$

where i indexes students, s indexes schools, and a number such as "8" indexes grade; Y_{is8} indicates student i's reading or math scores in the 8th grade; Y_{is3} is student i's reading or math score five years previously, in 3rd grade; $Socio_{i8}$ is a vector of variables capturing the socioeconomic and educational characteristics of student i, including race, free lunch eligibility or immigrant status as well as other educational needs, such as limited English proficiency or special education participation; SC_{s8} represents resource characteristics for school s in 8th grade; GS_i is a vector of grade span path indicators that describe the grade spans of schools attended by student i from 4th to 8th grade; $Switch_i$ is a series of dichotomous variables indicating whether

[5] There is a fairly high level of within-school consistency in academic paths. As an example, in approximately 84 percent of schools, more than half of eighth graders followed the same path to eighth grade. In 13 percent of schools more than 92 percent of students followed identical paths.

student i's school changed its configuration during the student's attendance or whether the student switched to a school with a different configuration, again, based on the grade span path indicators; and e_{is8} is an error term with the usual properties. Note that because our interest is in the effect of alternative paths between grades 3 to 8, our specification is best viewed as a long-term production function, with a lagged performance score from 3rd grade, which serves as a baseline.

In addition to equation (1), we estimate a second model that combines the various grade span paths into a smaller number of options, focusing on the articulation timing and grade (for example, articulating early rather than on time, or articulating at specific grade levels).

To summarize, the centerpiece of our empirical work is a value-added model of long-term gains in academic performance – between 3rd and 8th grades – that provides estimates of the impacts of grade span paths on student performance. Thus, our estimates focus on the relationship between student achievement and grade span path, controlling for a wide variety of student characteristics, including prior student achievement. We estimate these models using pooled data for the cohorts of students attending 4th grade in 1997-98 or 1996-97 and 8th grade in 2001-2002 or 2000-01.[6]

EMPIRICAL RESULTS

Student Characteristics and the Path from 4th to 8th Grade

Table 1 shows, for each of two SAP cohorts, the frequency distribution of students by the academic path they have followed between 4th and 8th grade. As described earlier, we distinguish between standard paths that correspond to articulating "as expected," meaning that the student

[6] We exclude students retained in grade and those who exited or entered the system during the period, but we analyze the possible impact of these exclusions in the section on sensitivity analyses.

graduates from one school (on time with the rest of his grade mates) and joins the next school as an "incoming freshman" (again, on time with the rest of his grade mates). As shown, only 60 percent of all students in the cohort articulate on time and as expected and the majority of these students do so by attending a K-5 school and then moving to a 6-8 school between 5th and 6th grades. Approximately four percent of all students who articulate as expected begin in a K-6; three percent begin in a K-4 and then move to a 5-8 and, importantly, three to four percent spend all four years in a K-8 school.

Approximately 18 to 19 percent of the cohort enrolls in a middle school midstream, that is, after the beginning of the school's grade span. The largest group of these students attends a K-6 school though 6th grade, then moves into 7th grade in a 6-8 school. A smaller percentage of students (approximately six percent) leaves their schools early and enters a new school at its earliest grade, for example by leaving a K-6 in 5th grade and entering a 6-8 in 6th grade. Thus, nearly one-quarter of all students in the cohort switch schools at a non-standard grade – either exiting early from their elementary school or entering middle school late. Finally, somewhat over 15 percent follow one of the more than 500 "other" paths each of which includes a tiny percentage of students. Whether these non-standard moves are harmful to student performance – and how they compare to the standard moves – is an empirical question we consider.

Before turning to estimating impacts, we explore the possibility that different kinds of students follow different academic paths, which would complicate the interpretation of the differences in performance as an estimate of the impacts of the different spans. Thus, we examine the characteristics of the students by path, including their test scores in the 3rd grade, which serves as our baseline. As shown in Table 2, there are, indeed, some differences between the students who follow a standard articulation path and those that exit early or join midstream or

are in other paths, but these are, on the whole small.[7] For example, there is a significant difference in 3rd grade test scores for those who articulated as expected and students who follow other paths.[8] In addition, students who articulate as expected are more likely to be white, while students who articulate early are more likely to be black and those who articulate on time and join midstream are more likely to be Hispanic, as compared to the other groups. The other observable characteristics display very few differences across the groups. For example, the percentages of students eligible for free and reduced price lunch, in part-time special education, with limited English proficiency and who are foreign born are all within several percentage points across the three major groups, though those who articulate early have lower proportions of students for whom English is not the primary language spoken at home. The "other" group – students who do not follow one of these primary academic paths - have quite different characteristics, with lower percentages of students who are white or Asian, higher free and reduced price lunch eligibility and lower average test scores.

School Configuration and Student Characteristics

Turning next to the differences in the students attending different types of elementary schools, Table 3 displays the characteristics of students in the four major types of schools serving 4th graders, as well as an other category. Again, comparing column 1 with each of the others, the differences in mean characteristics across the groups are often relatively small though statistically significant. Students in K-4 and other upper elementary schools tend to have the lowest performance while students in K-5 – the most common elementary grade span – have the highest performance. Not surprisingly, given student performance, K-4 schools also have the

[7] Variable definitions are included in Appendix A.
[8] Test scores are presented as standardized z-scores with a mean of zero and standard deviation of one.

highest percentage of students eligible for free and reduced price lunches and K-5 has the lowest eligibility level, though all are quite high (between 69 and 91 percent). The K-5 schools also have the highest percentage of white students while K-8 have the highest proportion of black students and K-4 the highest percentage of Hispanic students. Other student characteristics, such as limited English proficiency, percentage foreign-born and percentage who are in part-time special education are quite similar across grade spans.

Finally, we examine the differences in the students attending different kinds of middle schools – that is, schools serving 8th graders. As shown in Table 4, the 6-9 and other elementary/middle configurations appear to be outliers, with much higher than average performance in 8th grade. The 6-9 schools also have well below-average free and reduced price lunch eligibility and much higher-than-average performance in 3rd grade. The other configurations exhibit more consistent 8th grade performance, although K-8 schools have the highest performance while scores in 7-8 schools are lower than average. Of particular note is that the K-8 schools display the largest gains between 3rd and 8th grades. Interestingly, other observable student characteristics are quite similar across the groups. While the 6-8 schools average 72 percent of student with free and reduced price lunch eligibility, the remaining three grade spans (K-8, 5-8 and 7-8) average between 83 and 89 percent eligibility. Schools serving K-8 and 7-8 have the highest percentages of black students while the 5-8 schools have the highest percentage of Hispanic students. Again, the other student characteristics are similar across the schools. The patterns in the schools serving the middle schools largely mirror those in the elementary grades, with higher proportions of white students attending K-5 and then 6-8 schools, while more Hispanic students attend K-4 and then 5-8 and more black students attend either K-8 or K-6 followed by 7-8.

Taken together, while schools of different spans have broadly similar student bodies, some differences are significant, pointing to the need to control for differences in students in estimating the grade span effects.

The Impact of Grade Span on Student Performance

As described earlier in equation (1), we examine the impact of grade span on students by estimating a series of models regressing 8th grade test scores in math and reading on academic paths controlling for a variety of student characteristics. Columns 1 and 5 in Table 5 provide estimates of equation (1) for math and reading. All models include the grade span path variables (that is, a set of dummy variables that correspond to different grade span paths that a student might follow), a set of variables that distinguish between paths induced by changes in schools and switches made by a student between schools, student characteristics, and 8th grade school resources as controls. Because the coefficients distinguishing paths induced by changes in schools rather than by students between schools are largely insignificant, the coefficients on this set of controls are not shown. The reference category is the dominant path taken -- attending a K-5 school followed by a 6-8 school.

Examining the results in columns 1 and 5, we see no significant differences between some of the most common grade span paths; K-6 to 7-8 is no different than K-5 to 6-8, for example.[9] Students attending K-8 schools, though, earned significantly higher scores than others, both among those who articulated as expected and among those who switched midstream. The results suggest that students who attend K-8 schools from 4th grade through 8th grade exhibit significantly higher performance, equivalent to roughly a quarter of a standard deviation in reading and slightly less in math, as compared to students in a more traditional K-5 to 6-8 path.

[9] The other columns are discussed in the next section of the paper.

Several other academic paths also display higher student performance gains compared to the most common path. Students who move out of a K-5 into a K-8 also show significantly higher 8th grade performance in both math and reading. This result has interesting implications. First it suggests that older students may benefit from a K-8 environment even (or perhaps, especially) if they attended a school with a shorter grade span for their earlier grades. Second, it suggests that the K-8 effect may not be simply an artifact of students remaining in the same school for their elementary and middle grades. Instead, it indicates that the K-8 structure itself appears to have a positive impact on student performance.

Students who move into 5-8 schools from K-5 after 4th grade rather than 5th also exhibit significantly higher 8th grade performance. No other paths, apart from the other category encompassing a variety of rarely observed paths, show consistently higher performance. Note that all of the significant coefficients are positive, indicating that academic paths that include non-traditional middle grade structures – K-8 or 5-8 – appear to have a positive effect on student performance relative to the typical K-5 to 6-8 path.

Taken together, these results point to significantly different student performance gains across some grade span paths, with strong positive effects in paths that include K-8 schools. In particular, students who remain in K-8 schools from 4th through 8th grades and students who switch from K-5 to K-8 schools after 4th grade experience test score gains between 3rd and 8th grade that are, on average, approximately 0.25 standard deviations higher than student in the traditional grade span path. This is a significant difference that policymakers would recognize as important. The impact is larger than the estimated impacts of many school reforms – often less than even one-tenth of a standard deviation – and large relative to, say, the black-white test score gap in New York City (see authors (2006) for more).

While the results suggest that even switching to a K-8 school midstream may produce positive effects, the models do not allow us to distinguish the effects of stability - attending a single school between 4th and 8th grade – from the grade span itself. To test these effects more directly, we construct a new series of variables. As shown in Table 6, rather than estimating separate coefficients for each common path as in Table 5, we examine only the timing of articulation from one school to the next.

The patterns are consistent with the previous results. Examining columns 1 and 5, articulating early appears to have little effect on student performance while articulating on-time and joining midstream has a small positive effect on reading. Attending one school from 4th through 8th grade, though, is associated with significantly higher performance in both math and reading, and the size of the coefficients is very close to the K-8 coefficient reported in Table 5. The similarity is not coincidental, of course, because the K-8 structure is the only one that permits students to attend only one school during these grades (4-8 schools are possible, but there are too few to examine separately). We also find a positive, though smaller, effect on math for students who articulate in 5th grade rather than 6th or 7th grade, but do not find the same effect in reading.

SENSITIVITY ANALYSES

We test the sensitivity of our results with a number of additional analyses. First, we examine whether non-SAP students differ from SAP students in ways that could bias the results of the SAP estimations. Second, by controlling for observable and unobservable school quality and resource measures, we analyze whether results are driven by selection of families into grade span paths. Third, we explore whether the short-term effects of grade span on non-SAP 8th

graders (who entered the district late or were retained in an earlier grade) reveal different patterns from the long-term effects on SAP students presented above. Finally, we estimate the differential effects of grade span paths on subgroups of students, specifically low-scoring and minority students.

Does Attrition from SAP Bias Long-term Effects?

Because we are primarily interested in the long-run impact of school organization on student performance, our previous analyses included only those 4th grade cohort students who were enrolled and who made standard academic progress through 8th grade. Students in the 4th grade cohort will not be included in the SAP sub-sample if they exit the school district before 8th grade or are retained in any of grades 4 through 7. The estimates of the impact of grade span on student performance in 8th grade will thus be biased if retention or exit rates are affected by grade span. The extent and sign of the potential bias is an empirical issue.[10]

To address this concern, we first calculate exit and retention rates in 4th grade for all students in that grade and then compare the mean and dispersion of test scores across grade spans to examine whether there are distinctly different patterns in the performance of these students across spans. Of the 71,152 students in 4th grade in 1998-99, for example, 2.28 percent were retained while 9.64 percent exited the district's schools by 1999-2000. Graphs of the test score distribution for retained and exiting students in each grade span (not shown) indicate that the upper bound for retention across grade spans is similar, though some spans have a higher concentration of the lowest performing students. The academic profile of exiting students is also

[10] The possibility of attrition bias is common in longitudinal analyses based upon administrative data obtained from a single city or state, however, it has received little attention. In much of the previous research, authors report that differences between the population of all students and those included in the analyses (that is, for whom sufficient data are available) are small.

similar across spans, with a normal distribution centered on a z-score of zero. Schools serving K-4 appear to have a slightly larger tail of higher performing exiters than other spans.

A more rigorous way to analyze the effect of non-SAP students on our previous estimates is to view their exit and retention as a specific example of the standard sample selection model, in particular, one with non-random attrition in panel data. Failing to account for this non-random attrition will mean that estimates using the panel data will be biased. To formalize this framework, we model the likelihood that a student is in the SAP sub-sample:

$$SAP_{is}^{*} = \gamma_0 + \gamma_1 Y_{is4} + \gamma_2 Y_{is3} + \gamma_3 Socio_{i4} + \gamma_4 SC_{s4} + e_{is4} \qquad (2)$$

where SAP_{is}^{*} is a continuous but unobserved measure of the likelihood or tendency of student i in school s in 4th grade to experience standard academic progress (SAP). This student is observed in the SAP cohort if $SAP_{is}^{*} \geq 0$. Hence we define the indicator variable SAP_{is} to take a value of one if the student is in the SAP sub-sample and zero otherwise; that is

SAP_{is} = 1 if student i in school s is in 4th to 8th grade SAP;
i.e. if $SAP_{is}^{*} \geq 0$
SAP_{is} = 0 if student i in school s is not in SAP; i.e. if $SAP_{is}^{*} < 0$

We assume that the error terms from equations (1) and (2), e_{is8}, and e_{is4}, are bivariate normal. Then

$$E[e_{is8} \mid SAP_{is} = 1] = E[e_{is8} \mid e_{is4} > -X_{is4}\gamma] = \beta_6 \frac{\phi(X_{is4}\gamma)}{\Phi(X_{is4}\gamma)} = \beta_6 \lambda(X_{is4}\gamma) \qquad (3)$$

Where $X_{is4}\gamma = \gamma_0 + \gamma_1 Y_{is4} + \gamma_2 Y_{is3} + \gamma_3 Socio_{i4} + \gamma_4 SC_{s4}$ and $\lambda(X_{is4}\gamma)$ is known as the inverse Mills ratio. It follows that:

$$\begin{aligned} &E[Y_{is8} \mid SAP_{is} = 1] = E[Y_{is8} \mid e_{is4} > -X_{is4}\gamma] \\ &= \beta_0 + \beta_1 Y_{is3} + \beta_2 Socio_{i8} + \beta_3 SC_{i8} + \beta_4 GS_i + \beta_5 Switch_i + \beta_6 \lambda(X_{is4}\gamma) \end{aligned} \qquad (4)$$

Hence, a consistent estimator of the EPF equation can be obtained by estimating the model with only the SAP subset and including the inverse Mills ratio on the right-hand-side of the equation.

Note that $\lambda(X_{is4}\gamma)$ must be estimated since it is not directly observed. It is therefore necessary to first estimate the SAP equation (2) and then use the results to calculate $\lambda(X_{is4}\hat{\gamma})$ as an estimate of the inverse Mills ratio. To do this, we begin with the population of 4th graders in 1998-99, identifying the ones that remain in the 2002 SAP cohort in 8th grade and those that do not. We include variables measured in 4th grade in the SAP equation since some members of the 4th grade cohort will exit by 5th grade and no further information will exist for them. The three exclusion restrictions in this equation are the (current) 4th grade test scores in math and reading and the (lagged) 3rd grade score of the other test (i.e. for math – reading). We estimate equation (2) using probit with 4th grade school resources and also estimate a separate equation for the population of 4th graders in 1997-98, identifying the ones that remain in the 2001 SAP cohort in 8th grade and those that do not (results available from the authors upon request).

Columns 2 and 6 of Table 5 display results for the effects of grade span paths on math and reading tests with the sample selection bias correction term included.[11] The sign and significance of the coefficients on the K-8, K-5 to K-8, K-5 to 5-8 and other path variables are virtually identical to those presented in columns 1 and 5. The only substantive difference from the results without the sample selection correction is that the K-4 to 5-8 path now has a larger and significant coefficient for both reading and math. In addition, including the sample selection bias correction in the models from Table 6 (columns 2 and 6 examining articulation timing), the results are again nearly identical, though – consistent with the K-4 to 5-8 results - articulating in

[11] Note that we use robust standard errors, in part, to adjust for the heteroskedasticity that arises from including the inverse Mills ratio in our model.

5th grade has a stronger positive effect. Thus, the results are robust to correction for possible non-random attrition from the sample.

Do Families Select Grade Span Paths?

A second threat to the validity of our results is that families (or students) may select certain grade span paths and these selections may be based on characteristics of the students, families or schools, some of which are unobserved. As an example, families may believe that there are significantly better resources in the upper grades of some grade spans than in earlier grades, inducing moves aimed at maximizing resources in all grades. Alternatively, parents may know something about their children (e.g. dislike of change, desire to remain or not with siblings in the same school, desire to meet new people etc.) that makes a particular set of grade spans unusually helpful to academic performance. This kind of selection may be most important in early grades, because of their importance in shaping grade span path to 8th grade.

To address these concerns, we estimate two additional models. First, we control for school resources in each school that a student attends from 4th to 8th grade, rather than just the resources of the 8th grade school. In this way, we estimate the effect of grade span holding constant the resources available throughout grade span. This model addresses the concern of different resources across grade span combinations. Thus, for example, if parents choose a school because more experienced teachers are available in the upper grades, our resource variables will control for this difference. Second, we estimate a model with 4th grade school fixed effects to control for the possible selection at the beginning of the grade span paths. This model addresses the concern that parents choose an elementary school with the grade span path to 8th grade in mind, and that school factors unobserved in our data (for example, school culture or curriculum) may affect these choices.

Results of the estimation of these models are displayed in Tables 5 and 6, columns 3 and 4 for math and columns 7 and 8 for reading. The estimates that include resources and fixed effects are largely unchanged from the earlier results. In particular, the K-8, K-5 to K-8, and the K-5 to 5-8 early articulation paths continue to have the largest positive effects compared to the K-5 to 6-8 path. In Table 6, attending one school again has the largest positive effect even when the additional controls are added. For math, in fact, the model with the full set of controls in column 4 shows the largest coefficient for attending one school, as compared to the other models. The only substantive difference from the earlier results is that articulating in 7th grade has a significant negative effect on reading when school resources are included, though the coefficient becomes insignificant again when the 4th grade fixed effects are added. Thus, selection into paths seems not to substantively affect the results presented above.

Do 8th Graders with Non-Standard Academic Progress (Non-SAP) Respond Differently to Grade Spans?

Another group of students not included in the SAP cohort is comprised of those who *enter* the New York public schools after 4th grade or those who followed a non-standard academic path to 8th grade. These students have either attended school elsewhere -- in the case of immigrants, in another country altogether -- or they repeated grades. Unfortunately, we do not have data on their full academic histories. Instead, we analyze the short-term gains (7th to 8th grades and 6th to 8th grades) by grade span paths, for the 8th graders of 2001 and 2002 (the last years of our SAP cohort) who attended NYC schools for two (7th and 8th) or three (6th, 7th, and 8th) grades, respectively.

For the pooled 2001 and 2002 data, 110,952 8th graders with valid math scores were present in New York City in 7th grade in the previous year and have all the data needed to

estimate the EPF model of short-term gains. Similarly, 100,098 students with 8th grade math scores attended both 6th and 7th grades in New York City.[12] Of all 8th graders in 2001 or 2002 with math scores, 67.5 percent are in our SAP cohort.

Panel A of Table 7 presents results for reading and math for 8th graders in New York City public schools for at least two years and Panel B for students who are in these schools for at least three years. Although we estimate the full model from Table 5, with all variables interacted for non-SAP students, we present only the results for the path variables (full results available from authors). First, note that for SAP students in columns 1 and 3 of both panels in Table 7, the K-8 grade span path has a positive impact on 8th grade reading and math, controlling for 7th or 6th grade test scores and other covariates. The one year positive effect holds in math for students who change into the K-8 span from K-7.[13] Columns 2 and 4 in Panel A show that the K-8 span is associated with significantly larger math gains between 7th and 8th grade for non-SAP students as compared to SAP students in some cases, such as one year math gains for those who stay in K-8 and for those who switch from either K-7 or 6-8 to K-8. We find no differential K-8 effects for non-SAP student gains between 6th and 8th grade, however (Panel B). The 5-8 span is also associated with larger math and reading gains between 6th and 8th grade for SAP students, though non-SAP student have smaller gains than non-SAP students in reading. Several other paths have significant coefficients as well, though the ones with positive signs are few. We conclude that the short-term effects do not contradict the long-term effects of grade span paths and, as important, students who are not part of the SAP cohort because they enter late or are retained also perform well in the K-8 span.

Do Low Achieving or Minority Students Respond Differently?

[12] The numbers for reading are 107,253 and 97,843 respectively.
[13] Note that these are largely cases of schools switching from K-7 to K-8 rather than students moving to a different school.

We next turn to consider whether the impact differs for minority or low performing students. To examine these sub-groups, we re-estimate the full models presented in columns 4 and 8 of Table 5 and add interactions between grade span paths and 3rd grade test scores (to test for differential effects among high and low achieving students), and between grade span paths and indicators for student race (black and Hispanic). Results for the coefficients on the path variables are shown in Table 8 (other coefficients not shown). In math, a variety of grade span paths positively impact low achieving students. For example, as seen in column 1, a student who scores one standard deviation below the mean in 3rd grade math and follows a K-8 path from 4th to 8th grade has an average 0.09 standard deviation larger gain than a similar student in a K-5/6-8 path.[14] Interestingly, students with lower 3rd grade scores also appear to benefit in math from non-standard moves into 6-8 schools (either early or midstream) and on-time transitions from K-4 to 5-8. These results indicate that the K-8 grade span has particularly positive effects on math scores for lower-achieving students, as do several paths associated with the wider middle school span (5-8). We also find strongly positive results for K-8 schools for black and Hispanic students, including both those in K-8 schools between 4th and 8th grades and those who move from K-5 to a K-8. It also appears that moving from either K-8 to 6-8 or from K-6 to 6-8 early are particularly *ineffective* paths for minority students' math achievement.

In several cases, the results for reading differ from results for math for low achieving students. In particular, K-8 is associated with lower 8th grade reading scores for low achieving 3rd graders. No paths appear more effective for reading than K-5/6-8 for lower achieving 3rd graders. While K-8 schools do not affect black and Hispanic students differently, several other paths (K-5 to 5-8 or 6-9, K-6 to 7-9) have positive reading effects for minority students. As in

[14] Note that the negative sign on the coefficient indicates that lower performing 3rd graders have higher 8th grade scores in this path.

the math results, moving out of a K-8 into a 6-8 appears particularly ineffective for black students. In sum, the math results for these subgroups are similar to those for the full sample, with positive effects of K-8 schools for lower achieving and minority students. The reading results do not yield the same consistent patterns for these sub-groups though and suggest that black and Hispanic students' reading scores do not respond very differently to grade span paths than do those for other students.

DISCUSSION AND POLICY IMPLICATIONS

Our analysis of academic cohorts over a five-year span in New York City public schools finds strong evidence of grade span effects on student performance. Specifically, we find consistent evidence that students who attend schools serving longer grade spans – in particular K-8 schools -- have significantly larger reading and math test score gains between third and eighth grades. This result is particularly noteworthy because these schools do not appear to serve more privileged or higher-achieving students. On the contrary, K-8 schools have, on average, lower 3rd grade test scores and somewhat higher proportions of students from poverty and students who are black as compared to schools serving grades K-5 or 6-8. The results are robust to tests for specification and sample selection bias, and we find consistent positive K-8 effects for both students making standard academic progress and those who do not. Only on reading scores for low performing 3rd graders do we find negative effects of the K-8 span.

These results demonstrate the potential for improving student performance from policy initiatives to create more schools with a K-8 grade span and to minimize structural articulations as students move through the elementary and middle grades. Nevertheless, caution is warranted before school districts rush to wholesale implementation of a K-8 model. First, it is not clear that

simply shifting all schools to a K-8 model will necessarily produce the effects found here. In particular, conversions of existing schools to a K-8 span may take time to show positive effects (MacIver and MacIver, 2006). Second, additional research could provide important guidance for policy design. For example, while the present research builds on theories regarding the underlying reasons for grade span effects, it does not reveal the specific features of different grade span models that promote stronger achievement. Our findings suggest that grade span may matter, in part, because of its impact on school mobility, but other mechanisms and factors may also be critical. Third, the natural variation in New York City provides a unique opportunity to examine grade span effects in a "real-world" setting, thereby increasing the generalizability of the findings. While we have been able to utilize econometric methods to minimize threats to validity and draw causal inferences, causality based on natural variation is difficult to prove. Grade span research could also benefit from controlled experimental designs, with random assignment of students and teachers to schools of different grade spans. Evidence from such randomized trials could supplement the natural experiment presented here, providing stronger evidence on causality, and also help to identify specific features of K-8 spans that might increase student performance.

Despite the popularity of grade span reconfiguration to improve student performance and the growing popularity of K-8, in particular, the research base supporting these changes has been thin. Our research helps to fill this gap and, perhaps surprisingly, finds that these policies do appear to lead to better student performance among diverse groups of students, particularly as compared to traditional self-contained middle schools. As the movement to K-8 grade configuration picks up additional steam around the country, researchers and policy makers will

have numerous opportunities to examine whether these improvements can be implemented at large scale, and whether the gains can be sustained over the long-term.

REFERENCES

Allensworth, E. (2005). Graduation and Dropout Trends in Chicago: A look at cohorts of students from 1991 through 2004. Chicago: Consortium on Chicago School Research. Retrieved October 15, 2007 from <ccsr.uchicago.edu/publications/p75.pdf>.

Alspaugh, J.W. (1998). Achievement Loss Associated with the Transition to Middle School and High School. The Journal of Educational Research, 92(1), 20-25.

Alspaugh, J.W. and R.D. Harting. (1995). Transition Effects of School Grade-Level Organization on Student Achievement. Journal of Research and Development in Education 28(3), 145-149.

Bedard, K. and C. Do. (2005). Are Middle Schools More Effective? The Impact of School Structure on Student Outcomes. Journal of Human Resources 40(3), 660-682.

Bowie, L. (2007). Study raises doubts on K-8: JHU research finds no academic benefit over middle schools, undercuts city's push. Baltimore Sun. January 16.

Carnegie Task Force on Education of Young Adolescents. (1989). Turning Points: Preparing American Youth for the 21st Century. New York: Carnegie Council on Adolescent Development of the Carnegie Corporation.

Cook, P.J, R. MacCoun, C. Muschkin, and J. Vigdor. (2008). "The Negative Impacts of Starting Middle School in Sixth Grade." Journal of Policy Analysis and Management, 27 (1), p104-121

DeJong, W.S. and J. Craig. (2002). How Should Schools Be Organized? School Planning and Management 41(6), 26-32.

Gootman, E. (2007). Taking the Middle Schoolers Out of the Middle. New York Times, Section A, page 1, January 22.

Hammack, F.M. (Ed.) (2004). The Comprehensive High School Today. New York: Teachers College Press.

Hanushek, E.A., J.F. Kain and S.G. Rivkin. (1999). The Cost of Switching Schools. Texas Schools Project, Mimeo

Howley, C.B. (2002). Grade-Span Configurations. School Administrator 59(3), 24-29.

Jellick, C. (2005). District Plans School Setup Review; Peoria Will Consider a K-8 Configuration. Peoria Journal Star, May 27.

Jonas, M. (2007). Junior High. Boston Globe, Section E, page 1, January 28.

Mac Iver, M.A. and D.J. Mac Iver. (2006). Which Bets Paid Off? Early Findings on the Impact of Private Management and K-8 Conversion Reforms on the Achievement of Philadelphia Students. Review of Policy Research 23(5), 1077-1093.

National Center for Education Statistics. (2007). Digest of Education Statistics. Washington, DC: Government Printing Office, table 95.

Renchler, R. (2000). Grade Span. Research Roundup 16(3).

Roderick, M. (2006). Closing the Aspirations-Attainment Gap: Implications for High School Reform. New York: Manpower Demonstration Research Corporation.

Author (2008).

Seidman, E., L. Allen, J. L. Aber, C. Mitchell and J. Feinman. (1994). The Impact of School Transitions in Early Adolescence on the Self-System and Perceived Social Context of Poor Urban Youth. Child Development, 65, 507-525.

Author (2006).

Wampler, R.S., J. Musch, and M. Adams. (2002). Ethnic Differences in Grade Trajectories During the Transition to Junior High. Journal of School Psychology 40(3), 213-37.

Weiss, C.C. and L. Kipnes. (2006). Reexamining Middle School Effects: A Comparison of Middle Grades Students in Middle Schools and K-8 Schools. American Journal of Education, 112(2), 239-272.

Figure 1. Theoretical Model of the Relationship between School Grade Span and Student Performance

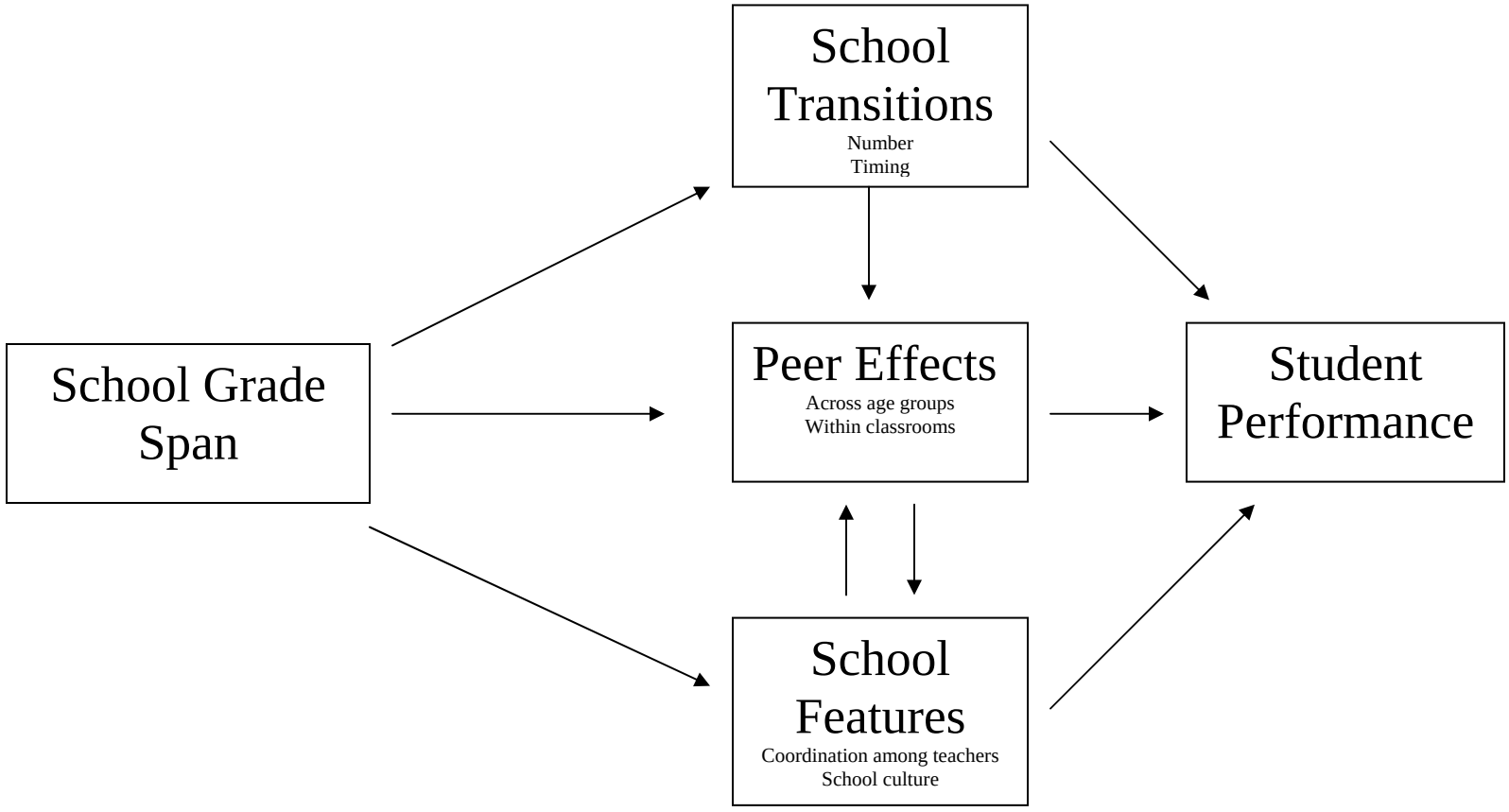

Table 1: Frequency Distribution of Students by Academic Grade-Span Paths, 98-02 and 97-01 SAP cohorts

Path	Freq. SAP 98-02	Percent	Freq. SAP 97-01	Percent
Articulate as expected				
K-5 to 6-8	19,072	47.34	18,668	45.40
K-8	1,559	3.87	1,374	3.34
K-6 to 7-8	1,515	3.76	2,003	4.87
K-4 to 5-8	1,180	2.93	1,155	2.81
K-5 to 6-9	1,036	2.57	1,195	2.91
Total	24,362	60.47	24,395	59.33
Articulate on time and joined midstream				
K-6 to 6-8	3,428	8.51	3,572	8.69
K-6 to 6-9	2,191	5.44	2,519	6.13
K-5 to 5-8	933	2.32	1,140	2.77
K-6 to 5-8	358	0.89	445	1.08
K-5 to K-8	221	0.55	246	0.60
Total	7,131	17.71	7,922	19.27
Articulate early to be on time at next level				
K-6 to 6-8 early	2,184	5.42	1,761	4.28
K-8 to 6-8	314	0.78	328	0.80
K-5 to 5-8 early	198	0.49	192	0.47
Total	2,696	6.69	2,281	5.55
Others				
Other Paths	6,097	15.31	6,522	15.86
Total Number of Students	40,286	100	41,120	100

Note: i) Total number of Paths: 540 for the 98-02 cohort and 520 for the 97-01 cohort.
ii) The student (or the school) might have switched to 7-9 post articulation. The regressions control for these cases.

Table 2: Student Mean Characteristics by Path Category, 8th graders, 98-02 and 97-01 SAP cohorts

	Articulate as expected	Articulate on time and joined midstream	Articulate early to be on time at next level	Others
3rd grade test score	0.25	0.18***	0.30 ***	0.11***
White	0.23	0.13 ***	0.19 ***	0.12 ***
Black	0.32	0.35 ***	0.45 ***	0.40 ***
Asian Plus	0.12	0.15 ***	0.08 ***	0.08 ***
Hispanic	0.32	0.37 ***	0.28 ***	0.41 ***
Free Lunch Eligible	0.61	0.68 ***	0.62	0.73 ***
Red. Price Lunch Eligible	0.09	0.10 ***	0.09 **	0.08 ***
Age	13.43	13.44 ***	13.44	13.45 ***
Female	0.53	0.53	0.52	0.53
Part-Time Special Ed	0.07	0.06 ***	0.05 ***	0.06
Foreign Born	0.14	0.14	0.12 ***	0.12 ***
Non-English at Home	0.42	0.43 **	0.31 ***	0.41
LAB Percentile	0.51	0.58 *	0.40	0.78 ***
Took the LAB	0.02	0.03 ***	0.02 ***	0.03 ***
LEP	0.02	0.02 ***	0.01 ***	0.03 ***
Number of Students	48,757	15,053	4,977	12,619

Total number of students: 40,286

Note: Asterisks show the significance of t-tests for equivalence of means across groups (all tests compare the mean value of the first column –articulate as expected– with each of the remaining ones).

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Student Mean Characteristics by Grade Span, 4th graders, 1996-97 and 1997-98

	K-4	K-5	K-6	K-8	Other Upper Elementary
4th grade test score	0.09	0.27***	0.18***	0.14***	0.06
3rd grade test score	0.08	0.27***	0.17***	0.16***	0.03**
White	0.04	0.24***	0.14***	0.19***	0.07***
Black	0.33	0.31***	0.39***	0.41***	0.44***
Asian Plus	0.05	0.12***	0.14***	0.07***	0.07***
Hispanic	0.58	0.33***	0.33***	0.32***	0.43***
Free Lunch Eligible	0.82	0.60***	0.66***	0.72***	0.76***
Red. Price Lunch Eligible	0.07	0.09***	0.10***	0.09**	0.07
Age	13.48	13.43***	13.45***	13.43***	13.47
Female	0.53	0.53	0.53	0.53	0.54
Part-Time Special Ed	0.05	0.06**	0.07***	0.07**	0.07
Foreign Born	0.12	0.14***	0.13**	0.13**	0.09**
Non-English at Home	0.51	0.43***	0.37***	0.39***	0.39***
LAB Percentile	1.08	0.50***	0.52***	0.76**	1.13
Took the LAB	0.05	0.02***	0.02***	0.03***	0.04
LEP	0.04	0.02***	0.02***	0.02***	0.03**
Number of Students	3,480	47,234	24,915	4,494	1,283

Total number of students: 81,406

Note: Asterisks show the significance of t-tests for equivalence of means across groups (all tests compare the mean value of the first column –"K-4"– with each of the remaining ones).

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Student Mean Characteristics by Grade Span, 8th graders, 2000-01 and 2001-02

	K-8	5-8	6-8	7-8	6-9	Other elementary/ middle
8th grade test score	0.20	0.09***	0.14***	-0.08***	0.57***	0.31***
3rd grade test score	0.11	0.14**	0.22***	0.02***	0.52***	0.31***
White	0.15	0.06***	0.21***	0.05***	0.31***	0.10*
Black	0.40	0.33***	0.37***	0.46***	0.12***	0.40
Asian Plus	0.06	0.07	0.11***	0.11***	0.30***	0.02***
Hispanic	0.39	0.54***	0.31***	0.38	0.27***	0.48***
Free Lunch Eligible	0.76	0.79***	0.63***	0.76	0.42***	0.79
Red. Price Lunch Eligible	0.07	0.08**	0.09***	0.10***	0.13***	0.09
Age	13.45	13.47***	13.44*	13.46*	13.40***	13.44
Female	0.54	0.54	0.53**	0.54	0.50***	0.55
Part-Time Special Ed	0.07	0.05***	0.06	0.07	0.06	0.02**
Foreign Born	0.11	0.14***	0.13***	0.12	0.16***	0.10
Non-English at Home	0.40	0.50***	0.39**	0.33***	0.55***	0.45
LAB Percentile	0.93	0.97	0.49***	0.50***	0.35***	0.04**
Took the LAB	0.04	0.05***	0.02***	0.02***	0.01***	0.00**
LEP	0.03	0.04***	0.02***	0.01***	0.01***	0.00**
Number of Students	6,111	8,184	54,569	4,165	8,167	210

Total number of students: 81,406

Note: Asterisks show the significance of t-tests for equivalence of means across groups (all tests compare the mean value of the first column –"K-8"– with each of the remaining ones).

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Academic Paths: 8th Grade Math and Reading Test Scores

	Math				Reading			
Path	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Articulate as Expected				
K-8	0.218**	0.218**	0.214**	0.224**	0.256**	0.255**	0.241**	0.252**
	(0.030)	(0.030)	(0.031)	(0.030)	(0.034)	(0.034)	(0.035)	(0.035)
K-6 to 7-8	-0.007	-0.004	-0.035	-0.009	-0.037	-0.031	-0.049	-0.027
	(0.039)	(0.039)	(0.036)	(0.036)	(0.052)	(0.051)	(0.049)	(0.050)
K-4 to 5-8	0.066	0.090*	0.068	0.073*	0.028	0.064*	0.038	0.046
	(0.036)	(0.036)	(0.035)	(0.033)	(0.030)	(0.030)	(0.031)	(0.032)
K-5 to 6-9	0.087	0.089	0.085	0.067	0.114*	0.117*	0.116*	0.102*
	(0.056)	(0.058)	(0.060)	(0.059)	(0.050)	(0.051)	(0.053)	(0.052)
				Articulate on Time and Joined in Midstream				
K-6 to 6-8	0.013	0.019	0.009	0.029	0.004	0.013	-0.000	0.016
	(0.023)	(0.023)	(0.023)	(0.023)	(0.024)	(0.023)	(0.023)	(0.023)
K-6 to 7-9	0.030	0.029	0.034	0.045	0.027	0.026	0.034	0.043
	(0.034)	(0.034)	(0.035)	(0.034)	(0.039)	(0.039)	(0.037)	(0.035)
K-5 to 5-8	0.016	0.022	0.046	0.052	0.069	0.077	0.077	0.088
	(0.046)	(0.046)	(0.040)	(0.043)	(0.054)	(0.053)	(0.052)	(0.051)
K-6 to 5-8	0.037	0.044	0.042	0.041	0.083	0.092	0.088	0.090
	(0.044)	(0.044)	(0.046)	(0.052)	(0.050)	(0.048)	(0.047)	(0.048)
K-5 to K-8	0.264**	0.264**	0.272**	0.269**	0.285**	0.284**	0.281**	0.282**
	(0.062)	(0.061)	(0.060)	(0.061)	(0.068)	(0.067)	(0.067)	(0.068)
				Articulate Early to be on Time at Next Level				
K-6 to 6-8 early	0.015	0.021	0.018	0.031	0.024	0.034	0.033	0.042
	(0.030)	(0.029)	(0.030)	(0.029)	(0.023)	(0.022)	(0.023)	(0.022)
K-8 to 6-8	0.104*	0.094	0.100*	0.103*	0.076	0.059	0.060	0.063
	(0.052)	(0.052)	(0.049)	(0.049)	(0.063)	(0.063)	(0.061)	(0.062)
K-5 to 5-8 early	0.235**	0.236**	0.248**	0.227**	0.201**	0.202**	0.193*	0.185*
	(0.057)	(0.057)	(0.051)	(0.051)	(0.073)	(0.074)	(0.078)	(0.080)
				Other Paths				
Other	0.045*	0.051**	0.041*	0.045*	0.076**	0.085**	0.074**	0.079**
	(0.019)	(0.019)	(0.020)	(0.019)	(0.022)	(0.022)	(0.023)	(0.023)
Lambda	No	Yes	Yes	Yes	No	Yes	Yes	Yes
4th Grade FE	No	No	No	Yes	No	No	No	Yes
4th -7th Grade Resources	No	No	Yes	Yes	No	No	Yes	Yes
Observations	81,406	81,406	81,406	81,406	80,344	80,344	80,344	80,344
R-squared	0.53	0.53	0.54	0.54	0.49	0.50	0.50	0.50

Robust standard errors in parentheses
** p<0.01, * p<0.05

Notes:

i) The omitted path is K-5 to 6-8.

ii) Missing values of independent variables have been recoded to zero. Models include dummy variables to indicate missing values and the following dummy variables: School Switched to a Different Configuration Prior to Articulation, Student Switched to a School with a Different Configuration Prior to Articulation, School Switched to a Different Configuration Post Articulation, and Student Switched to a School with a Different Configuration Post Articulation. Models include the following controls: 3rd grade test score, student race, free lunch eligibility, reduced price lunch eligibility, age, gender, part-time special education status, foreign born, non-English speaking at home, LAB Percentile, Took the LAB, LEP, student enrollment, teacher-pupil ratio, percent teachers with: 2-year degree, MA degree, at least 5 years experience, with license.

iii) 4th grade resources are excluded in regressions in columns (4) and (8) that include 4th grade school fixed effects

Table 6: Timing of Articulation: 8th Grade Math and Reading Test Scores

	Math				Reading			
Path	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Articulate on Time and Joined Midstream	0.043	0.046	0.071**	0.068*	0.079*	0.082*	0.089**	0.089**
	(0.029)	(0.029)	(0.026)	(0.027)	(0.035)	(0.035)	(0.033)	(0.033)
Articulate Early to be on Time at Next Level	0.034	0.036	0.038	0.047	0.036	0.039	0.041	0.048*
	(0.026)	(0.026)	(0.026)	(0.025)	(0.022)	(0.022)	(0.021)	(0.021)
Attended One School	0.214**	0.214**	0.211**	0.221**	0.251**	0.249**	0.237**	0.248**
	(0.030)	(0.030)	(0.030)	(0.030)	(0.033)	(0.034)	(0.035)	(0.035)
Articulate in 5th Grade	0.083*	0.103**	0.086**	0.086**	0.043	0.073*	0.051	0.055
	(0.034)	(0.033)	(0.033)	(0.031)	(0.030)	(0.030)	(0.030)	(0.032)
Articulate in 7th Grade	-0.023	-0.022	-0.052	-0.033	-0.062	-0.060	-0.073*	-0.060
	(0.029)	(0.029)	(0.027)	(0.028)	(0.036)	(0.036)	(0.033)	(0.034)
Other	0.042*	0.048**	0.038*	0.042*	0.071**	0.080**	0.069**	0.073**
	(0.018)	(0.018)	(0.019)	(0.019)	(0.021)	(0.021)	(0.021)	(0.021)
Lambda	No	Yes	Yes	Yes	No	Yes	Yes	Yes
4th Grade FE	No	No	No	Yes	No	No	No	Yes
4th -7th Grade Resources	No	No	Yes	Yes	No	No	Yes	Yes
Observations	81,406	81,406	81,406	81,406	80,344	80,344	80,344	80,344
R-squared	0.53	0.53	0.54	0.54	0.49	0.49	0.50	0.50

Robust standard errors in parentheses
** p<0.01, * p<0.05

Notes:

i) The omitted path is K-5 to 6-8.

ii) Missing values of independent variables have been recoded to zero. Models include dummy variables to indicate missing values and the following dummy variables: School Switched to a Different Configuration Prior to Articulation, Student Switched to a School with a Different Configuration Prior to Articulation, School Switched to a Different Configuration Post Articulation, and Student Switched to a School with a Different Configuration Post Articulation. Models include the following controls: 3rd grade test score, student race, free lunch eligibility, reduced price lunch eligibility, age, gender, part-time special education status, foreign born, non-English speaking at home, LAB Percentile, Took the LAB, LEP, student enrollment, teacher-pupil ratio, percent teachers with: 2-year degree, MA degree, at least 5 years experience, with license.

iii) 4th grade resources are excluded in regressions in columns (4) and (8) that include 4th grade school fixed effects

Table 7: 8th Grade Math and Reading Test Scores, Short-term EPFs, SAP and Non-SAP Regression Results

Panel A: 7th-8th Grade

	Math		Reading	
	SAP	Diff Non-SAP	SAP	Diff Non-SAP
	(1)	(2)	(3)	(4)
Path		Attend One School		
K-8	0.099***	0.052**	0.143***	0.002
	(0.024)	(0.022)	(0.030)	(0.024)
5-8	0.055**	0.034*	0.064	-0.030
	(0.022)	(0.018)	(0.043)	(0.019)
6-9	0.014	0.013	0.060	-0.002
	(0.035)	(0.027)	(0.041)	(0.052)
7-8	-0.001	0.000	-0.032	0.050*
	(0.040)	(0.030)	(0.041)	(0.026)
7-9	-0.013	-0.023	-0.010	0.015
	(0.028)	(0.028)	(0.050)	(0.022)
		Changed School or School Changed		
K-7/K-8	0.249**	0.067*	0.245	0.002
	(0.113)	(0.040)	(0.164)	(0.066)
6-8/K-8	0.012	0.163**	0.032	0.101
	(0.046)	(0.074)	(0.060)	(0.067)
6-8/5-8	-0.038	0.067*	-0.044	0.022
	(0.029)	(0.038)	(0.046)	(0.031)
5-8/6-8	0.055*	-0.016	0.037	-0.012
	(0.029)	(0.032)	(0.052)	(0.050)
6-7/6-8	0.010	0.056	0.121	-0.064
	(0.056)	(0.060)	(0.076)	(0.071)
6-9/6-8	-0.149***	0.104*	-0.066	0.104**
	(0.029)	(0.056)	(0.069)	(0.044)
7-8/6-8	-0.081	0.100***	-0.087	0.064***
	(0.061)	(0.038)	(0.082)	(0.021)
6-8/7-8	-0.055*	0.020	-0.030	-0.004
	(0.030)	(0.042)	(0.048)	(0.031)
Other	0.030	0.026	0.020	-0.028
	(0.035)	(0.027)	(0.033)	(0.029)
Observations		110,952		107,253
R-squared		0.66		0.60

Panel B: 6th-8th Grade

	Math		Reading	
			Attend One School	
K-8	0.130***	-0.002	0.152***	0.014
	(0.033)	(0.029)	(0.035)	(0.027)
5-8	0.064*	0.032	0.075*	-0.040*
	(0.036)	(0.026)	(0.044)	(0.024)
6-9	0.060	-0.001	0.110**	-0.043
	(0.049)	(0.030)	(0.044)	(0.059)
			Attend Two Schools or School Changed	
K-6/5-8	-0.034	-0.031	-0.028	0.010
	(0.041)	(0.045)	(0.041)	(0.059)
K-6/6-8	-0.080***	-0.018	-0.103***	-0.042**
	(0.024)	(0.023)	(0.022)	(0.021)
K-6/6-9	-0.167***	-0.062*	-0.177***	0.003
	(0.037)	(0.036)	(0.051)	(0.096)
K-6/7-8	-0.114**	0.006	-0.152***	0.073**
	(0.053)	(0.035)	(0.043)	(0.029)
K-6/7-9	-0.493***	0.499**	-0.340***	0.454
	(0.177)	(0.243)	(0.100)	(0.290)
K-6/7-8/6-8	-0.148***	0.045	-0.158**	-0.012
	(0.046)	(0.031)	(0.080)	(0.040)
			Attend Three Schools or Schools Changed	
K-6/6-8/7-8	-0.148***	-0.116	-0.198***	-0.031
	(0.046)	(0.091)	(0.048)	(0.049)
5-8/5-8/6-8	0.028	0.048	-0.099***	0.131**
	(0.061)	(0.069)	(0.032)	(0.053)
6-8/5-8/5-8	0.040	0.090	0.006	-0.000
	(0.046)	(0.058)	(0.053)	(0.068)
6-8/5-8/6-8	0.070	-0.098	0.153***	0.008
	(0.098)	(0.133)	(0.059)	(0.086)
6-8/6-8/5-8	-0.001	-0.001	-0.045	0.012
	(0.030)	(0.041)	(0.036)	(0.046)
6-8/6-9/6-8	-0.121***	0.161**	0.063	0.105**
	(0.046)	(0.067)	(0.040)	(0.045)
6-8/6-8/7-8	-0.047	-0.098*	-0.011	-0.028
	(0.053)	(0.057)	(0.069)	(0.043)
6-9/6-9/6-8	-0.089***	0.043	-0.226***	0.153
	(0.033)	(0.057)	(0.031)	(0.116)
Other	-0.041	0.009	-0.021	-0.033*
	(0.025)	(0.021)	(0.029)	(0.020)
Observations		10,0098		97,843
R-squared		0.63		0.57

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes:

i) The omitted path is 6-8.

ii) Missing values of independent variables have been recoded to zero. Models include dummy variables to indicate missing values the following controls: 3rd grade test score, student race, free lunch eligibility, reduced price lunch eligibility, age, gender, part-time special education status, foreign born, non-English speaking at home, LAB Percentile, Took the LAB, LEP, plus the following school resources for 8th grade for Panel A and for 7th and 8th grader for Panel B: student enrollment, teacher-pupil ratio, percent teachers with: 2-year degree, MA degree, at least 5 years experience, with license

Table 8: Subgroup Analysis: 8th Grade Math and Reading Scores in 2001-02 and 2000-01, paths interacted with third grade scores and race

	Math			Reading		
Path	Third Grade Test Score	Black	Hispanic	Third Grade Test Score	Black	Hispanic
	(1)	(2)	(3)	(4)	(5)	(6)
K-8	-0.088***	0.202***	0.103*	0.179***	0.088	0.113
	(0.030)	(0.055)	(0.053)	(0.025)	(0.069)	(0.069)
K-6 to 7-8	-0.008	-0.144**	-0.027	0.204***	0.026	0.143*
	(0.030)	(0.069)	(0.067)	(0.025)	(0.098)	(0.078)
K-4 to 5-8	-0.073**	0.017	0.009	0.189***	0.058	0.096
	(0.036)	(0.095)	(0.094)	(0.027)	(0.074)	(0.068)
K-5 to 6-9	0.047**	0.011	0.010	0.339***	0.095*	0.166***
	(0.020)	(0.042)	(0.051)	(0.030)	(0.057)	(0.062)
K-6 to 6-8	-0.065**	0.080	0.101**	0.174***	0.030	0.072
	(0.026)	(0.050)	(0.044)	(0.019)	(0.057)	(0.056)
K-6 to 7-9	0.038	-0.016	0.080	0.306***	0.081*	0.111**
	(0.025)	(0.051)	(0.051)	(0.024)	(0.048)	(0.046)
K-5 to 5-8	-0.042	0.105	-0.041	0.195***	0.237***	0.129**
	(0.033)	(0.067)	(0.076)	(0.035)	(0.065)	(0.058)
K-6 to 5-8	-0.035	0.161	0.046	0.207***	0.260	0.170
	(0.037)	(0.114)	(0.112)	(0.028)	(0.227)	(0.209)
K-5 to K-8	-0.009	0.382***	0.174	0.242***	0.335*	0.149
	(0.046)	(0.123)	(0.118)	(0.058)	(0.180)	(0.165)
K-6 to 6-8 early	-0.062***	-0.109**	-0.182***	0.179***	-0.055	-0.079*
	(0.019)	(0.051)	(0.049)	(0.016)	(0.047)	(0.047)
K-8 to 6-8	-0.053	-0.288***	-0.196*	0.161***	-0.353***	-0.072
	(0.050)	(0.109)	(0.101)	(0.050)	(0.106)	(0.091)
K-5 to 5-8 early	-0.019	0.041	0.104	0.207***	0.162	0.204*
	(0.056)	(0.117)	(0.147)	(0.066)	(0.138)	(0.116)
Other	-0.062***	-0.051	-0.046	0.197***	0.014	0.068
	(0.018)	(0.044)	(0.041)	(0.014)	(0.049)	(0.043)
Observations	81,406			80,344		
R-squared	0.54			0.51		

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes:

i) The omitted path is K-5 to 6-8.

ii) Missing values of independent variables have been recoded to zero. Models include dummy variables to indicate missing values and the following dummy variables: School Switched to a Different Configuration Prior to Articulation, Student Switched to a School with a Different Configuration Prior to Articulation, School Switched to a Different Configuration Post Articulation, and Student Switched to a School with a Different Configuration Post Articulation. Models include the following controls: 3rd grade test score, student race, free lunch eligibility, reduced price lunch eligibility, age, gender, part-time special education status, foreign born, non-English speaking at home, LAB Percentile, Took the LAB, LEP, lambda, 4th grade fixed effect, 5th -8th grade resources: student enrollment, teacher-pupil ratio, percent teachers with: 2-year degree, MA degree, at least 5 years experience, with license.

Appendix A: List of Variables Used in the Regressions

Variable	Description
3rd grade test score	3rd grade z-score (math or reading, depending the model)
4th grade test score	4th grade z-score (math or reading, depending the model)
8th grade test score	4th grade z-score (math or reading, depending the model)
White	Dummy variable equal to 1 if student is white, equal to 0 otherwise
Asian Plus	Dummy variable equal to 1 if student is Asian (or native American or unknown ethnicity), equal to 0 otherwise
Black	Dummy variable equal to 1 if student is black, equal to 0 otherwise
Hispanic	Dummy variable equal to 1 if student is black, equal to 0 otherwise
Free Lunch Eligible	Dummy variable equal to 1 if student is eligible for free lunch, equal to 0 otherwise
Red. Price Lunch Eligible	Dummy variable equal to 1 if student is eligible for reduced price lunch, equal to 0 otherwise
Age	Age of the student
Female	Dummy variable equal to 1 if student is female, equal to 0 otherwise
Part-Time Special Ed	Dummy variable equal to 1 if student is in part time special education, equal to 0 otherwise
Foreign Born	Dummy variable equal to 1 if student is foreign-born, equal to 0 otherwise
Non-English at Home	Dummy variable equal to 1 if the language at home is not English, equal to 0 otherwise
LAB Percentile	LAB spring percentile (test of English language knowledge)
Took the LAB	Took the LAB test
LEP	Dummy variable equal to 1 if the LAB percentile is less or equal than 40, equal to 0 otherwise (less than 40 is cutoff for eligibility for English language services)