Building Bridges to Middle School? Elementary School Departmentalization and Academic Achievement in the Upper Grades

Ben Backes
James Cowan
Dan Goldhaber

March 2025

WORKING PAPER No. 315-0325

CALDER
National Center for Analysis of Longitudinal Data in Education Research

AIR®

Building Bridges to Middle School? Elementary School Departmentalization and Academic Achievement in the Upper Grades

Ben Backes
American Institutes for Research / CALDER

James Cowan
American Institutes for Research / CALDER

Dan Goldhaber
American Institutes for Research / CALDER
University of Washington

Contents

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A210008 to American Institutes for Research (AIR). The opinions expressed are those of the authors and do not represent the views of the Institute, the U.S. Department of Education, or the Massachusetts Department of Elementary and Secondary Education. The authors thank Duncan Chaplin, Beth Gamse, Pierre Lucien, Roddy Theobald, Aubree Webb, and participants at the 2023 Association for Public Policy and Management and 2024 Association for Education Finance and Policy Annual Meetings for helpful comments. The authors also thank Daniel Moore for providing excellent research assistance. The research design for this paper was pre-registered at the Registry of Efficacy and Effectiveness Studies (#16020).

CALDER • American Institutes for Research
1400 Crystal Drive 10th Floor, Arlington, VA 22202
202-403-5796 • www.caldercenter.org

Building Bridges to Middle School? Elementary School Departmentalization and Academic Achievement in the Upper Grades
Ben Backes, James Cowan, & Dan Goldhaber
CALDER Working Paper No. 315-0325
March 2025

Abstract

In departmentalized elementary schools, teachers specialize in one or more core subjects and instruct different groups of students throughout the day. Proponents argue that this model leads to higher quality instruction because teachers can specialize in their strongest subjects and that it better prepares students for middle and high school, which are typically departmentalized. Using a difference-in-differences design with seven cohorts of elementary students in Massachusetts, we find that students in departmentalized elementary schools have better academic outcomes in middle school. The effects are concentrated in the first year of middle school, where each year of elementary school departmentalization improves test scores by about 0.03 standard deviations in math and ELA. The math and ELA effects mostly fade out by the end of 8th grade, although we do find persistent effects on science tests.

1. Introduction

Although most elementary school teachers are generalists and teach multiple subjects, about 15% specialize in fewer subjects and instruct multiple classes of students throughout the day (National Center for Education Statistics [NCES], 2023). Proponents of this departmentalized instruction (DI) model argue that it leads to higher quality instruction in the short run with teachers focusing on material in which they have a comparative instructional advantage (Backes et al., 2023a; Bastian & Fortner, 2020). But research has also found that individual teachers are less effective when they specialize and have less knowledge of their students' individual academic needs (Backes et al., 2023a; Bastian & Fortner, 2020; Hwang & Kisida, 2022; Fryer, 2018). The empirical literature on these departmentalized models has therefore reached mixed conclusions about their aggregate short-run effects (Backes et al., 2023a; Fryer, 2018).

Regardless of its effects in the short run, DI involves trade-offs between teacher content knowledge and student-teacher relationships that may have differing effects over longer time horizons. The switch to middle school coincides with a decline in engagement and achievement that may be partially attributable to changes in classroom organization and school structure (Marks, 2000; Rockoff & Lockwood, 2010). Proponents of DI argue that it prepares students for this transition by improving instruction and acclimating students to middle school routines (Chan et al., 2009). On the other hand, student-teacher relationships contribute to the development of independent study habits that have important long-run effects on students' academic success (Davis, 2003; Jackson, 2018). To the extent that it disrupts these relationships, elementary DI could exacerbate the difficulties students encounter in the transition to middle school.

In this paper, we provide some of the first evidence on the longer-run effects of DI on academic outcomes in middle and high school. We follow seven cohorts of elementary school students and assess outcomes between fifth and tenth grades. Using a difference-in-differences design that exploits variation in DI practices across cohorts from the same elementary school, we find that an additional year of DI increases student test scores in secondary school by about 0.01 standard deviations in math and ELA, and by about 0.02 standard deviations in science. These effects are largest in the transitional year between elementary and middle school and fade out in math and ELA by the end of eighth grade; the positive effects are more persistent in science. Overall, the findings suggest that elementary DI better prepares students for middle school but may not have lasting effects on academic outcomes.

2. Background and Literature Review

The existing literature on DI and teacher specialization is limited to short-run analyses and has produced mixed results. Fryer (2018) randomly assigned elementary schools in Houston to switch to DI and found that it reduced student achievement by about 0.05 standard deviations in both math and reading over a two-year period. Backes et al. (2023) study DI in Massachusetts using a difference-in-differences design. They found positive effects of DI on reading tests and null or small positive effects on math achievement. Bastian and Fortner (2020) and Hwang and Kisida (2022) examined teacher specialization in elementary schools and found little relationship between specialization rates and student achievement. There has not been any research to date on the longer-run effects of DI despite the fact that much of the debate on elementary DI focuses on the middle school transition and how it might affect achievement in later years.

Departmentalization is intended to provide deeper instruction by allowing teachers to specialize only in their stronger subjects. Backes et al. (2023a) and Bastian and Fortner (2020) find that this results in better matches between teacher and subject matter. Prior research has shown that teacher effects persist for several years (Chetty et al., 2014; Jacob et al., 2010; Kinsler, 2012). The persistent effects of teachers are only partly explainable based on their contemporary effects on test scores, which suggests that some teachers are better at imparting foundational skills (Gilraine & Pope, 2021). If specialization extends the research of these teachers, then learning gains from DI may persist into future years.

Elementary DI may also facilitate tracking or ability grouping because students attend multiple classes and class assignments are more flexible. For instance, DI allows schools to operate advanced classes only in certain subjects or to assign students to tracks separately by subject (Chan & Jarman, 2004). Students who have earlier access to advanced material may be better prepared for instruction in middle schools. This flexibility was one motivation for New York City's inclusion of DI in an initiative to expand access to algebra in middle schools (New York City Department of Education, 2015). To the extent that these reforms better prepare students for later courses, we might expect some students to enroll in more advanced coursework in later grades.

Finally, DI may smooth the transition to middle school, which entails significant changes in school organization and academic routines. Many students first encounter a departmentalized structure, with different teachers for each subject, when they enter middle school. This transition is associated with significant declines in student engagement, achievement, and the strength of student-teacher relationships (Anderman, 2003; Anderman & Maehr, 1994; Harter et al., 1992; Marks, 2000; Rockoff & Lockwood, 2010; Schwerdt & West, 2013). The structure and routines of departmentalized elementary schools mimic those found in middle schools and establishing them earlier in a closer-knit school setting may allow students to "rehearse" for the middle school transition (Chan et al., 2009; Haley, 2018). For instance, Chan et al. (2009) argue that "self-contained classroom organization may be a good way to ease students from home setting to school setting. However, when students come to fourth and fifth grades, a change in academic organization is needed to work toward a seamless transition as they move from elementary to middle schools."

These mechanisms suggesting potentially positive effects of DI must be considered against the potential drawbacks of the weaker student-teacher ties in departmentalized models. That is, DI may merely accelerate the drawbacks associated with standalone middle schools rather than prepare students for change. In surveys, teachers report that self-contained models better support students' emotional wellbeing (Brobst & Markworth, 2018). Empirically, Fryer (2008) finds that teachers report knowing their students less well under DI, while Backes et al. (2024) find that students report less engagement in DI schools. Part of the longer-run effects of teachers comes from imparting effective academic habits (Davis, 2003; Jackson, 2018), a process that may be interrupted in the departmentalized model. If students fail to develop as independent learners under these structures, their academic performance may also suffer in later grades.

Taken together, these research literatures suggest considerable ambiguity about the longer-run effects of DI. The empirical results in the short-run have so far been mixed. But, more importantly, the theoretical literatures on teacher specialization and teacher-student relationships point in opposite directions. In this paper, we provide direct empirical evidence based on changes over time in school instructional models.

3. Data

3.1 Categorizing Instructional Models

Our sample includes students who enroll in public elementary schools in Massachusetts as third graders between 2012 and 2018, excluding students enrolled in K-8 schools. We omit K-8 schools given the focus on middle school transitions and because part of the research design relies on comparing outcomes for students attending the same middle school but different elementary schools. To the extent possible given years of available data for each cohort, we follow students from the first grade higher than those offered by their elementary school through 10th grade. We use data on elementary school instruction prior to the pandemic: 2012-2018 third grade cohorts for K-4 schools, 2012-2017 cohorts for K-5 schools, and 2012-2016 cohorts for K-6 schools.[1] We then follow students through the 2023 school year.

We follow Backes et al. (2023a) and use student and teacher class assignment data to characterize instructional models. We first identify self-contained classrooms where students have only one teacher for the core academic courses or multiple co-teachers who all teach the same set of classes together. We define a school-grade-year cell as departmentalized if fewer than 50% of the students are in self-contained classrooms; most school-grade-year cells are either almost entirely self-contained or entirely departmentalized. Among departmentalized schools, 67% have no students in self-contained classrooms and only 12% have more than 20% of their students in self-contained classes.

After characterizing schools' instructional models, we identify the first school a student enrolls in for third grade. We then calculate the number of years of exposure students have to DI in elementary school assuming that students attended school continuously between third grade and the highest grade offered in the schools in which they are enrolled in 3rd grade. We use the anticipated DI exposure as an intent-to-treat measure to avoid the potential concern that school mobility is associated with instructional model. We explore the relationship between anticipated and observed exposure to DI in elementary school in Section 5.

[1] We use shorthand such as "K-5" for ease of notation, but note that we also include schools that do not offer kindergarten; i.e., grade 1-5 schools. We do not include schools whose grade configuration changed during the sample period (e.g., from K-5 to K-6).

3.2 Sample and Summary Statistics

We link the elementary school data to data on students' academic and behavioral outcomes in grades 5 through 10. For each student, we retain outcomes for grades above the highest grade offered in the elementary school in which they were enrolled in for third grade. The student achievement outcomes include tests offered through the state standardized testing system, including: math and ELA tests in grades 5 – 8 and 10; tests of science, technology, and engineering in 8th grade; and a biology test taken after the completion of a biology course in 9th or 10th grade. We also measure students' grade point average in core academic subjects (math, ELA, science, and social studies) in grades 6 through 10 to obtain an annual measure of core subject GPA.[2]

We next construct measures of course taking using student schedule data. Massachusetts uses a state course coding system based on the School Codes for the Exchange of Data (SCED) system (Bradby et al., 2007; National Forum on Education Statistics, 2011) that standardizes course identification across districts. We first identify students taking advanced math courses, which are more typically differentiated from standard courses in middle and early high school than in other subjects. We identify these as either above-grade-level math courses for students in grades 6 – 8 or courses at a level above the grade level standard in the state curricular frameworks.[3] We also identify students taking supplemental courses designed to help students reach grade level in a particular subject.[4] Finally, we use two behavioral measures included in

[2] Course grades are typically not available for students in fifth grade.
[3] In grades 6 through 8, most students are assigned a course titled "Grade X Mathematics" that covers instruction of topics in the state curricular frameworks. The above-level courses we consider are: Pre-Algebra or Algebra in 6th and 7th grade; Algebra or Geometry in 8th grade; Geometry, Algebra II, or Trigonometry/Pre-Calculus in 9th grade; and Algebra II or Trigonometry/Pre-Calculus in 10th grade. Because there is considerably less differentiation by course title in other subjects, we do not consider advanced course-taking in other subjects.
[4] We did not include supplemental courses in the original pre-registration (Backes et al., 2023b); however, in subsequent work, we have found these courses strongly predict baseline achievement in both middle and high

the administrative data: the total number of absences in each school year and the number of days a student is suspended in each school year.

Since 2018, Massachusetts has surveyed students in 5th, 8th, and 10th grade about the climate in their schools. The survey asks about experiences with school staff and other students. It consists of three dimensions (environment, engagement, and safety) each of which includes three distinct topics. The engagement dimension includes cultural competency (e.g., the extent to which students and faculty value diversity), participation in classroom and school life, and relationships with students and faculty. The safety dimension includes emotional safety (e.g., the extent to which students are comfortable approaching faculty with problems), physical safety, and bullying. The environment dimension includes topics on the instructional environment, mental health, and discipline. The surveys are included in the state standardized testing system and thus have high response rates. For the middle school grades we use in this study, response rates are generally around 80% (77-82%) except for the 2020-2021 school year (66% in 5th grade, 69% in 8th grade).

We summarize the data in Table 1. The main analysis sample includes students who are currently enrolled in Massachusetts public schools and enrolled in grades above the highest grade offered by their grade 3 elementary school. The first row displays the average number of years of exposure to elementary school DI. Because DI in K-4 schools is relatively rare, students in grade 5 who have advanced past elementary school typically have few years of DI exposure. Among students in higher grades, who mostly frequently attend K-5 schools, the average length of DI exposure is about 0.6 – 0.7 years.

school, which we interpret as evidence that they usefully proxy for students' academic track (Backes et al., 2023c). Nonetheless, we view their inclusion as exploratory given their omission from the pre-registration report.

In Figure 1, we plot average student achievement by grade and elementary school grade configuration (K-4, K-5, and K-6). The plots show the average differences from third grade test scores (the first grade of testing in Massachusetts). As in Rockoff and Lockwood (2010) and Schwerdt and West (2013), student achievement declines in the first year of middle school and gradually recovers thereafter. The declines are largest for K-4 transitions (about 0.12 standard deviations in math and ELA) and somewhat smaller for K-5 and K-6 (about 0.03 – 0.09 standard deviations). In the remaining sections, we investigate the extent to which elementary school structure influences these trajectories.

4. Research Design

We assess the long-run effects of DI using a difference-in-differences design that exploits variation in exposure to DI across cohorts in the same elementary school. We estimate

$$Y_{iecgs} = DI_{ec}\delta + \theta_{eg} + \lambda_{gc} + \epsilon_{iecgs} \qquad (1)$$

where *i* indexes student, *e* indexes elementary school, *c* indexes cohort (year of first third grade enrollment), *g* indexes current grade, and *s* indexes current school. DI_{ec} is a measure of DI exposure in elementary school; in our baseline specification, this is the number of years of DI instruction assuming a student continuously attends their elementary school from the third grade to the highest enrolled grade in that school. In the baseline specifications, we estimate intent-to-treat effects that use the instructional modality for the first school in which a student enrolled in third grade. In addition, we also assess treatment-on-the-treated effects using a student's observed school in each grade. We cluster standard errors at the elementary school level.

The research design in Eq. (1) compares changes in outcomes for students who attend elementary schools that switch instructional models to those in schools with a constant instructional model. The primary threat to this design is that changes in instructional models are correlated with other school or district initiatives that may affect student outcomes in later grades. For instance, the New York City algebra initiative discussed in previous sections suggests that districts might combine DI with policy changes affecting students in later grades. To test this possibility, we exploit the fact that multiple elementary schools typically feed into a single middle or high school. We can therefore compare changes in student outcomes across cohorts in elementary schools that switch models to other cohorts in the same middle or high school whose elementary schools did not switch models. In practice, we add a current-school-by-cohort fixed effect to Eq. (1):

$$Y_{iecgs} = DI_{ec}\delta + \theta_{eg} + \lambda_{gc} + \gamma_{sc} + \epsilon_{iecgs} \tag{2}$$

This specification allows the time trend to vary across middle or high schools, which allows for the introduction of other policies affecting instruction at higher grade levels that might coincide with changes in elementary DI. This would cover, for instance, the New York City algebra initiative discussed earlier. Because this specification also allows for changes in unobserved factors across cohorts for students enrolled in the same school, these results should not be sensitive to broader changes in district policy that occur simultaneously with changes in the instructional models used in elementary schools, including those that might affect student learning in earlier grades.

An additional concern in both designs is that schools adjust instructional models based on the current cohort of students. For instance, if schools departmentalize instruction for more advanced students, then we may find that DID designs falsely conflate compositional changes with the long-run effects of DI. In Table 2, we regress baseline student characteristics measured in 2nd grade on the cohort DI measures. The sample includes students in the analytic sample (grades 5 – 10), but we use grade 2 characteristics because very few schools departmentalize that early, so nearly all of these are measured prior to any potential treatment. The first characteristic is on-time grade progression, which is our sample selection indicator. We find no evidence of differential sample selection associated with DI in elementary school. We do, however, find some evidence that changes in DI are associated with compositional changes. DI is associated with more English learners and fewer special education students; the special education result is robust to the inclusion of school-by-cohort FE. Notably, the sorting patterns do not suggest any directionally consistent selection into instructional model. Nonetheless, adding controls for student background does not meaningfully change the results.

5. Results

5.1 School Assignments and Observed Treatment

We first demonstrate that DI assigned in a student's first third grade school predicts observed DI in Table 3. In the first column, we regress a student's number of observed years in elementary DI on the cumulative DI exposure and elementary school and cohort fixed effects (Panel A) and school-by-cohort fixed effects (Panel B). In the remaining columns, we repeat this exercise using DI status in each grade as the outcome. Overall, each year of assigned elementary DI raises actual DI by about 0.9 years. We also observe changes in treatment throughout the upper elementary grades: one year of DI is associated with an increase in the likelihood of departmentalization of 29 percentage points in 3rd grade, 35 percentage points in 4th grade, 23 percentage points in 5th grade, and 3 percentage points in 6th grade.

5.2 Dynamic Effects and Pre-Treatment Trends

Before turning to the estimated long-run effects of DI on student outcomes, we first plot the dynamic effects of DI in Figure 2. Because the treatment is continuous, we therefore follow

Freyaldenhoven et al. (2021) and estimate a version of Eqs. (1) and (2) that include transformations of the cohort leads and lags of DI:[5]

$$Y_{iecgs} = \sum_{k=-3,k\neq -1}^{2} \left(DI_{ec} - DI_{e,c-k}\right)\delta_k + \delta_3 DI_{e,c-3} + \delta_{-4}(1 - DI_{e,c+4}) + \theta_{eg} + \lambda_{gc} + \epsilon_{iecgs}$$

We bin the effects at t = -4 and $t = 3$. All effects are normalized relative to the next cohort's DI exposure.

There are two notable features of the results in Figure 2. First, we find little evidence of pre-existing trends in the academic outcome measures. Conditional on a cohort's DI treatment, leads of the treatment at the elementary school level are not predictive of changes in achievement outcomes. This is less true of grade point average, although the apparent pre-trend disappears in the model with school-by-cohort FE. Second, the estimated effects of DI increase over time, at least for the standardized achievement outcomes. Because the treatment variables are continuous measures of DI, this pattern suggests that the effect of a single year of DI in elementary school is more impactful the more years an elementary school has implemented DI.

5.3 Effects on Academic Outcomes

We present the main estimates of the effects of DI on academic outcomes in middle and high school in Table 4. The results in the first panel correspond to the baseline difference-in-differences design, while the second panel adds school-by-cohort FE to allow cohort trends to vary at the middle- or high-school level. We first examine the average of math and ELA test scores. The point estimates suggest that an additional year of DI raises achievement by 0.01 standard deviations on average, although the effect is not statistically significant. Results are

[5] The usual alternative to this specification is to include distributed lags of the treatment variable and then plot the cumulative effects of a one-unit change in treatment. Freyaldenhoven et al. (2021) note that their transformation produces the identical effects and obviates the need to compute the cumulative treatment effects for each event-time. We impute an unchanged treatment status for any period outside our observation window.

similar for math and ELA tests as the dependent variables. The math result is significant when we include school-by-cohort FE, and marginally significant in ELA without, but there are no clear trends. The effects are more positive for science, where we find that an additional year of DI raises achievement by about 0.02 standard deviations. This latter result accords with the evidence on the short-run effects of DI, which consistently show larger effects in science than in other subjects (Backes et al., 2024; Bastian & Fortner, 2020). The consistently larger effects in science may reflect the fact that it is the subject that typically receives the least instructional time in the mainstream, self-contained elementary classroom (Perie et al., 1997). Finally, we find little evidence of any effect on grades. Although the baseline estimate is marginally significant, recall that there was some evidence of a pre-treatment trend in this outcome. When we include school-by-cohort FE, the result is much closer to zero.

We next estimate the effects of elementary school DI by grade by interacting the DI measure with grade level. We plot the results in Figure 3. Recall that most elementary schools serve grades K through 5, so the first year of middle school is typically in 6th grade. For the math and ELA tests, we observe relatively large effects in 6th grade (0.02 in math; 0.03 in ELA) that fade out to close to zero by 8th grade. The science results are more imprecise, but we find positive effects on both the middle school (8th grade) and high school (biology) tests.

We investigate these patterns further in Table 5. We interact the DI dosage variable with an indicator for whether the student's current grade is the transitional grade between elementary and middle school. We define the transitional grade based on the grade range of a student's initial elementary school and their current grade level. For example, a student enrolled in a K-5 school would experience their transition grade in sixth grade. We omit science tests from these analyses because there is not a science test for the modal transitional grade in our sample (6th grade).

The first two panels replicate the analysis in Table 4. Consistent with the estimated effects by grade, the effects of elementary DI are most apparent in the transitional year. We estimate that an additional year of DI improves test scores by about 0.03 standard deviations in the transitional grade and estimates are similar in both math and ELA. In all other grades, the point estimates are close to zero and not statistically significant.[6]

There are two potential explanations for the effects of elementary DI on achievement during the transitional year. Backes et al. (2024) find that DI raises student achievement during elementary school. Given the well-documented tendency for the effects of educational interventions to fade out over time, the effects may be largest during the transitional year because it is most proximate to the elementary grades when students were treated. In this case, the effects of DI might reflect better academic preparation during elementary school, but not necessarily the kinds of middle school "rehearsal" effects suggested by Chan et al. (2009) and others. On the other hand, if these middle school rehearsal effects are an important explanation, we would expect to see that elementary DI leads to smaller declines in achievement – and hence higher test scores – as students switch from elementary to middle school.

We test these explanations in the remaining panel using the first difference in test scores for grades 5 – 8.[7] We use the more conservative specification with school-by-cohort FE, but results are similar in both specifications. The difference in test scores during the transitional year is the change in test scores from the last year of elementary school to the first year of middle school. This measure therefore captures changes in the negative middle school transition shock

[6] We omit K-8 schools given our focus on the middle school transition. Nonetheless, we can examine whether similar trends hold in these schools. In K-8 schools, the vast majority of grades 6-8 are departmentalized (about 93% for grade 6, and over 99% for grades 7 and 8). We thus examine the impact of an additional year of DI in grades 3-5 on student outcomes in grades 6-8. Although estimates are imprecise, we do not observe any evidence that DI in early grades in K-8 schools better prepares students for departmentalized upper grades within these same schools. Results are available upon request.

[7] We omit high schools from this analysis given that there is no 9th grade math or ELA test in Massachusetts.

associated with elementary DI. We estimate that elementary DI increases this difference by about 0.01-0.02 standard deviations per year, with somewhat larger effects in math than ELA. This suggests that part of the elementary DI effect comes from reducing the decline in test scores during the middle school transition, consistent with the rehearsal effects hypothesis described above. Comparing the estimates to those in the second panel, the findings suggest that about 50% of the elementary DI effect operates through reducing test score declines during the transitional year. That also suggests that about half of the effect can be explained by better preparation during elementary school and fade-out on the test score effects. Consistent with the fade-out hypothesis, test score growth is lower by about 0.01 standard deviations per year in the other middle school grades. Consequently, the achievement effects fully fade out by the end of 8th grade in math and ELA.

The estimated effects shown so far assume that each additional year of DI has a constant effect on student achievement. In Table 6, we relax this assumption. The effects of a single year of DI in elementary school are about one quarter to one half the magnitude of the continuous measure and not statistically significant. On the other hand, we estimate that two or three years of elementary DI increase achievement by about 0.02 – 0.03 standard deviations in math and ELA and about 0.05 standard deviations in science. We find little consistent evidence of any effect on grades.

Overall, the results indicate that student achievement would be about 0.06 standard deviations higher in 6th grade in both math and ELA for students who spent all of third through fifth grade in DI. It is useful to compare this to the effects of middle school transitions found in the previous literature. Both Rockoff and Lockwood (2010) and Schwerdt and West (2013) estimate that middle school transitions reduce achievement by about 0.10 – 0.13 standard deviations. Our results therefore suggest that early transitions to DI can ameliorate much of the negative effects of middle school transitions.

5.4 Accounting for Staggered Treatment Timing

Several recent studies have shown possible biases associated with traditional two-way fixed effects (TWFE) estimators for the difference-in-differences design (Goodman-Bacon, 2021; Roth et al., 2023). To ensure that our results are not sensitive to these issues, we implement an alternative estimator in the spirit of the stacked difference-in-differences procedure developed by Cengiz et al. (2019). We first divide treated units into mutually exclusive groups based on patterns of DI exposure across cohorts. For example, one such group is formed by elementary schools whose 2012 through 2014 third grade cohorts experienced no DI, whose 2015 cohort experienced one year of DI, and whose 2016 through 2018 cohorts experienced two years of DI. We then match each of these timing groups to the entire sample of never-DI schools. Let each such sample consisting of one timing group and the untreated schools be denoted by k. We estimate the stacked DID

$$Y_{iecgk} = DI_{ec}\delta + \theta_{egk} + \lambda_{gck} + \epsilon_{iecgk} \tag{3}$$

by interacting the elementary school and time indicators with the stacked sample indicators. Because there is no variation in treatment timing within subsample, the estimates from Eq. (3) are uncontaminated by differential treatment timing. We again cluster standard errors by school. Results for these specifications are shown in Table 7. For the academic outcomes examined in this section, results are quite similar for the overall estimates of DI and for the effect of DI in transitional grades. If anything, the results for math and ELA in transitional grades are somewhat larger, with a coefficient of 0.03 standard deviations for the combined result.

6. Mechanisms

In the prior section, we found that students who attend departmentalized elementary schools have higher achievement, particularly during the first year of middle school. In this section, we consider possible explanations for these findings.

6.1 Course-Taking, Attendance, and Engagement

In Table 7, we estimate the effects of DI on several intermediate outcomes that might proxy for student engagement in school. These include two measures for engagement in school (absences and suspensions), two measures for the academic level of courses taken (taking a supplemental class designed to help students reach grade level and taking an advanced math class), and an index of academic engagement from the student surveys. One limitation of the student surveys is that they are only available for some cohorts, and only in 5^{th} and 8^{th} grades; as a consequence, almost all of the observations are from 8^{th} grade.

Across all grades, we only find consistent effects of DI on absences. We estimate that a year of DI reduces absences by about 1 percent; this result is not sensitive to specification. In our baseline specifications, students attending departmentalized elementary schools are less likely to take advanced math courses and more likely to take supplemental courses, but these effects are not robust to controlling for school-specific trends. The point estimates are not significant and are close to zero. The survey results are less precise, but there is no clear evidence that students from departmentalized elementary schools self-report higher engagement, at least in 8^{th} grade.

In the remaining panel, we repeat these analyses by transitional grade given the pattern of results in Table 5 and Figure 3. Notably, we do not find a statistically significant effect on attendance in the transitional grade; rather, unlike the test score effects, the effect on attendance appears to persist into later years. In these specifications, we also find some evidence that students in departmentalized elementary schools are tracked into higher-level math classes, but only in the first year of middle school, when these courses are less frequently offered.

Overall, the evidence that DI affects engagement in middle school is mixed. We find positive effects on attendance that appear to persist throughout middle school. We also find some evidence that students take more advanced math classes, but only in the first year of middle school. This latter finding could reflect the fact the students from departmentalized elementary schools enter middle school with higher test scores, which determines course assignments in some places.

6.2 Persistence of Teacher Effects

Backes et al. (2024) find that elementary schools change teacher assignments when they departmentalize. Schools are more likely to assign teachers to the subjects in which they have a comparative advantage and increase the likelihood that teachers with subject-specific credentials teach the corresponding subjects. These changes in teacher staffing appear to offset the negative penalties to the departmentalized format for individual teachers. Prior research indicates that teacher contributions to student test scores persist to some extant across school years; estimates of the persistence in contemporaneous teacher effects vary between 20 and 50% (Chetty et al., 2014b; Jacob et al., 2010; Kinsler, 2012). In order to further test whether better academic preparation in elementary school could explain the benefits of DI in middle school, we follow Jacob et al. (2010) and estimate the persistence of prior teacher effects and how this varies by the instructional model used in elementary school.

We focus on fourth and fifth grade teachers (for whom we can estimate value added) and sixth grade test scores, where we observe the largest effects of DI. We first use data on teacher assignments in elementary schools between 2012 and 2019 to estimate teacher value added for fourth and fifth grade students in our elementary school sample. We estimate the VA model

$$Y_{ijgst} = X_{ijgst}\beta + \theta_{gst} + \epsilon_{ijgst}$$

$$\hat{\alpha}_{jt} = n_{jt}^{-1} \sum_{j(i,t)=j} (Y_{ijgst} - X_{ijgst}\hat{\beta} + \hat{\theta}_{gst}) \tag{3}$$

where $\hat{\alpha}_{jt}$ is the average of the residuals from the VA model for a teacher's students in each year. The control vector X_{ijgst} includes cubic polynomials in students' prior math and ELA test scores, gender, race/ethnicity, FRL eligibility, special education status, and English language learner status, as well as the means of these variables at the classroom level. We estimate Eq. (3) separately by subject for math and ELA teachers. We also include a school-grade-year fixed effect, θ_{gst}, to separate teaching effectiveness from the contributions of elementary schools' instructional models. Following Chetty et al. (2014a) and Stepner (2013), we then construct a forecast of teacher effectiveness in year *t* by aggregating the teacher effectiveness estimates, $\hat{\alpha}_{jt}$, from other years based on their ability to predict teacher effectiveness in year *t.* We omit data from year *t* in these forecasts so that the second-stage regressions of students' middle school test scores on teacher VA does not use a VA measure relying on their own elementary school test scores.

We then link the teacher VA estimates to linked student-teacher data in sixth grade. By construction, the VA estimates use data from students' elementary schools and the linked student-teacher dataset uses data from students' middle schools. We then estimate the following two-stage least squares regressions based on those in Jacob et al. (2010):

$$Y_{ict} = Y_{ict-k}\delta_k + X_{it}\beta + DI_{it-k}\lambda + \theta_{ct} + \epsilon_{ict}$$

$$Y_{ict-k} = \hat{\alpha}_{jt-k}^{-(t-k)}\rho_k + X_{it}\pi + DI_{it-k}\xi + \gamma_{ct} + \eta_{ict} \tag{4}$$

In Eq. (4), we instrument for prior achievement Y_{ict-k} in year *t-k* with VA for the teacher assigned to student *i* in year *t-k* (omitting data from the year in question) $\hat{\alpha}_{jt-k}^{-(t-k)}$. The control vector includes the same student variables included in the VA model, twice lagged achievement, and a classroom fixed effect. As discussed in Jacob et al. (2010), δ_k provides an estimate of the proportion of the variance in teacher effects on test scores in year *t-k* that persist to year *t*. To test whether this persistence is constant across elementary school instructional models, we construct two instruments by interacting estimated teacher VA $\hat{\alpha}_{jt-k}^{-(t-k)}$ with an indicator for each of the two possible instructional models for a student in year *t-k* and conduct the overidentification test that δ_k does not vary with the choice of instrument.

The results are shown in Table 9. In the first panel, we follow Jacob et al. (2010) and estimate the persistence of student learning by instrumenting prior test scores with twice-lagged test scores. The general persistence of student learning varies between 0.93 and 0.97 across tests but is similar in both DI and self-contained classrooms. In the second panel, we instrument prior achievement with the lagged teacher's VA estimate. In ELA, the persistence of teacher effects is about 0.2 in both DI and non-DI elementary schools. Hence, about 20% of teacher value-added persists into the next school year across the elementary-middle school transition. This is nearly identical to the rate of persistence found in Jacob et al. (2010). We do not reject the null hypothesis of the overidentification test that the two persistence rates are equal. The persistence of teacher effects does appear to differ in math, however. The rate of persistence in math VA is about 0.14 in non-DI models and 0.24 in DI models. In other words, the persistence rate is about 10 percentage points higher in DI than non-DI schools. In this case, the overidentification test does indicate that the difference in persistence between DI and non-DI elementary schools is statistically significant. In the final set of columns, we test differences in teacher persistence for both subjects simultaneously and do not find that the overall difference is jointly significant across subjects.

Overall, the results on teacher persistence provide some evidence that DI supports learning that persists, at least in math. A greater proportion of math teachers' contributions to student test scores persists into middle school when elementary schools departmentalize. However, we generally observe that these improvements in achievement are relatively short-lasting and mostly disappear by 8th grade. Although Gilraine and Pope (2021) suggest that teachers with more persistent effects also improve longer-run outcomes, it is unclear whether that holds in this setting.

7. Discussion

Proponents of transitioning upper elementary grades to DI focus on preparing students academically for middle school. Students gain experience taking multiple classes, transitioning between classrooms, and studying with more than one teacher. These experiences are intended to cultivate independent academic skills that will benefit them in the middle school environment. They are also matched with subject-matter specialists, who may better instill foundational concepts that will benefit them in more advanced courses. These arguments suggest that DI should improve student learning after they leave elementary school.

In this paper, we provide some of the first evidence on these hypotheses. We find clear evidence that elementary DI prepares students for the transition to middle school. We find that students have higher achievement in middle school after their elementary schools switch to departmentalized models. The effect is mostly concentrated in the transitional school year, where the estimated effect is about 0.03 standard deviations in both math and ELA tests per year of elementary DI. Our estimates suggest about half of this effect is driven by improvements in learning during elementary school, and the other half comes from a smaller decline in achievement during the elementary-to-middle school transition. The total effect of having attended a fully departmentalized elementary school is nearly enough to offset the middle school transition effect found in prior research (Rockoff & Lockwood, 2010; Schwerdt & West, 2013).

Nonetheless, we find mixed evidence on the persistence of elementary DI effects. Student achievement gains in math and ELA mostly fade out in later grades; by 8th grade, we estimate null effects of elementary DI on test scores in these subjects. The effects on science tests, by contrast, persist through the first year of high school. Students exposed to DI are no more likely to reach more advanced math courses or earn better course grades, and we find little evidence that students from departmentalized elementary schools report higher engagement in school by eighth grade.

One important caveat is that the lack of an observed effect on intermediate outcomes does not necessarily imply that DI has no long-run benefit to students. Educational interventions frequently produce short-term test score gains that fade out in subsequent grades, only to eventually reappear on longer-run outcomes. Examples include pre-kindergarten programs (Anderson, 2008; Deming, 2009), reductions in class size (Chetty et al., 2011), and exposure to effective teachers (Chetty et al., 2014). Although we observe rapid fade-out in math and ELA test score effects after the first year of middle school, effects on science tests appear more persistent and departmentalized teachers impart greater long-run learning than non-departmentalized teachers, at least in math.

References

Anderman, E. M., & Maehr, M. L. (1994). Motivation and Schooling in the Middle Grades. *Review of Educational Research*, *64*(2), 287–309. https://doi.org/10.3102/00346543064002287

Anderman, L. H. (2003). Academic and Social Perceptions as Predictors of Change in Middle School Students' Sense of School Belonging. *The Journal of Experimental Education*, *72*(1), 5–22. https://doi.org/10.1080/00220970309600877

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association*, *103*(484), 1481-1495.

Backes, B., & Cowan, J. (2019). Is the pen mightier than the keyboard? The effect of online testing on measured student achievement. *Economics of Education Review*, *68*, 89–103. https://doi.org/10.1016/j.econedurev.2018.12.007

Backes, B., Cowan, J., & Goldhaber, D. (2023). *Departmentalized instruction and elementary school effectiveness*. CALDER Working Paper.

Bastian, K. C., & Fortner, C. K. (2020). Is Less More? Subject-Area Specialization and Outcomes in Elementary Schools. *Education Finance and Policy*, *15*(2), 357–382. https://doi.org/10.1162/edfp_a_00278

Betts, J. R. (2011). The economics of tracking in education. In *Handbook of the Economics of Education* (Vol. 3, pp. 341-381). Elsevier.

Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3), 1405-1454.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly journal of economics*, *126*(4), 1593-1660.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American economic review*, *104*(9), 2633-2679.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014a). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, *104*(9), 2593–2632. https://doi.org/10.1257/aer.104.9.2593

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014b). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, *104*(9), 2633–2679. https://doi.org/10.1257/aer.104.9.2633

Condie, S., Lefgren, L., & Sims, D. (2014). Teacher heterogeneity, value-added and education policy. *Economics of Education Review*, *40*, 76–92. https://doi.org/10.1016/j.econedurev.2013.11.009

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), 111-134.

Freyaldenhoven, S., Hansen, C., Pérez, J. P., & Shapiro, J. (2021). *Visualization, Identification, and Estimation in the Linear Panel Event-Study Design* (No. 29170). National Bureau of Economic Research. https://doi.org/10.3386/w29170

Fryer, R. G. (2018). The "Pupil" Factory: Specialization and the Production of Human Capital in Schools. *American Economic Review*, *108*(3), 616–656. https://doi.org/10.1257/aer.20161495

Harter, S., Whitesell, N. R., & Kowalski, P. (1992). Individual Differences in the Effects of Educational Transitions on Young Adolescent's Perceptions of Competence and Motivational Orientation. *American Educational Research Journal*, *29*(4), 777–807. https://doi.org/10.3102/00028312029004777

Hong, K., Zimmer, R., & Engberg, J. (2018). How does grade configuration impact student achievement in elementary and middle school grades? *Journal of Urban Economics*, *105*, 1–19. https://doi.org/10.1016/j.jue.2018.02.002

Hwang, N., & Kisida, B. (2021). *Spread Too Thin: The Effects of Teacher Specialization on Student Achievement* (EdWorkingPaper No. 21–477). Annenberg Institute at Brown University. https://www.edworkingpapers.com/ai21-477

Jacob, B. A., Lefgren, L., & Sims, D. P. (2010). The Persistence of Teacher-Induced Learning. *Journal of Human Resources*, *45*(4), 915–943. https://doi.org/10.1353/jhr.2010.0029

Kinsler, J. (2012). Beyond Levels and Growth: Estimating Teacher Value-Added and its Persistence. *Journal of Human Resources*, *47*(3), 722–753. https://doi.org/10.1353/jhr.2012.0023

Marks, H. M. (2000). Student Engagement in Instructional Activity: Patterns in the Elementary, Middle, and High School Years. *American Educational Research Journal*, *37*(1), 153–184. https://doi.org/10.3102/00028312037001153

National Center for Education Statistics [NCES]. (2023). National Teacher and Principal Survey, 2020-2021. United States Department of Education, National Center for Education Statistics.

National Forum on Education Statistics. (2011). *Prior-to-secondary school course classification system: School codes for the exchange of data (SCED)* (NFES 2011–801). National Center for Education Statistics, U.S. Department of Education.

New York City Department of Education. (2015). Algebra for All Program memo. Accessed at https://admin.abcsignup.com/files/F474B456-0F3C-48EE-A066-0C58FCAAAC88_35/Algebra4All-ProgramOverview.pdf.

Pellegrini, A. D., & Bartini, M. (2000). A Longitudinal Study of Bullying, Victimization, and Peer Affiliation During the Transition From Primary School to Middle School. *American Educational Research Journal*, *37*(3), 699–725. https://doi.org/10.3102/00028312037003699

Perie, M., Baker, D.P., Bobbit, S. (1997). *Time Spent Teaching Core Academic Subjects in Elementary Schools: Comparisons Across Community, School, Teacher, and Student Characteristics* (NCES 97-293). Office of Educational Research and Improvement, U.S. Department of Education.

Rockoff, J. E., & Lockwood, B. B. (2010). Stuck in the middle: Impacts of grade configuration in public schools. *Journal of Public Economics*, *94*(11–12), 1051–1061. https://doi.org/10.1016/j.jpubeco.2010.06.017

Schwerdt, G., & West, M. R. (2013). The impact of alternative grade configurations on student outcomes through middle and high school. *Journal of Public Economics*, *97*, 308–326. https://doi.org/10.1016/j.jpubeco.2012.10.002

Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology*, *89*(3), 411–419. https://doi.org/10.1037/0022-0663.89.3.411

Tables and Figures

Table 1. Summary Statistics for Middle and High School Outcomes

	5	6	7	8	9	10
Years of Assigned DI	0.18	0.58	0.63	0.64	0.68	0.68
Years of Observed DI	0.24	0.59	0.63	0.63	0.68	0.68
On-time Progress	0.95	0.93	0.90	0.88	0.86	0.86
Enrolled in MA Schools	0.96	0.94	0.92	0.90	0.90	0.86
Math Test	-0.05	0.06	0.07	0.06		0.12
ELA Test	-0.03	0.06	0.06	0.08		0.13
Science Test				0.08	0.12	-0.02
Grade Point Average		3.26	3.16	3.11	2.85	2.96
Absences	7.24	7.63	8.59	10.34	11.74	12.44
Days Suspended	0.08	0.16	0.21	0.28	0.44	0.36
Advanced Math	0.00	0.02	0.20	0.28	0.28	0.26
Supplemental Course	0.09	0.11	0.11	0.11	0.04	0.04
Observations	69464	227580	246273	252010	239854	184413

Notes: Sample includes all students who enroll in an elementary school (K-4, K-5, or K-6) in Massachusetts for third grade.

Table 2. Changes in Instructional Models and Student Observables

	Baseline DID	School-Cohort FE
On-time Progress	0.002	0.001
	(0.001)	(0.001)
Limited English Proficient	0.007**	0.002
	(0.003)	(0.003)
Male	-0.001	0.002
	(0.002)	(0.003)
FRL	0.000	-0.005
	(0.003)	(0.003)
Special Education	-0.004*	-0.005**
	(0.002)	(0.002)
Asian	-0.001	0.000
	(0.001)	(0.001)
Black	0.002	0.002
	(0.001)	(0.002)
Hispanic	0.000	0.002
	(0.002)	(0.002)

Notes: Estimated effects of years of DI exposure in elementary school on student attrition and baseline characteristics using main regression sample. On-time progression indicates that a student is enrolled in a Massachusetts public school and is in the expected grade given the year of the student's first enrollment in third grade. Enrolled in MA school indicates that a student is enrolled in a Massachusetts public school. All other student characteristics are measured in the year prior to first third grade enrollment. All regression models include elementary school and cohort fixed effects. Standard errors clustered by elementary school in parentheses.
* p < 0.10 ** p < 0.05 *** p < 0.01

Table 3. Intent-to-Treat Effects on Observed Treatment

		Observed DI by Grade			
	Observed DI Years	Grade 3	Grade 4	Grade 5	Grade 6
Baseline DID					
DI Years	0.896***	0.289***	0.347***	0.229***	0.025***
	(0.008)	(0.017)	(0.013)	(0.017)	(0.006)
Num. Obs.	1219594	1219594	1219594	1219594	1219594
School-Cohort FE					
DI Years	0.850***	0.279***	0.335***	0.211***	0.019***
	(0.007)	(0.017)	(0.012)	(0.013)	(0.003)
Num. Obs.	1107601	1107601	1107601	1107601	1107601

Notes: Estimated effects of years of DI in elementary school cohort. Treatment is the number of years a student would have spent in DI assuming on-time grade progression in their assigned school in third grade. Years of DI (row 1) is the number of years a student is observed in DI during elementary school. Elementary school indicates the school a student first attends third grade. School indicates a student's current school. Standard errors clustered by elementary school in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 4. Effects of DI on Academic Outcomes in Middle and High School

	Math & ELA	Math	ELA	Science	Grades
Baseline DID					
DI Years	0.010	0.009	0.011*	0.015**	0.011*
	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
Num. Obs.	847398	841595	843200	337933	911277
School-Cohort FE					
DI Years	0.009*	0.012**	0.007	0.020***	0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Num. Obs.	846917	841129	842745	337661	910512

Notes: Estimated effects of years of DI in elementary school cohort. Treatment is the number of years a student would have spent in DI assuming on-time grade progression in their assigned school in third grade. Years of DI (row 1) is the number of years a student is observed in DI during elementary school. Elementary school indicates the school a student first attends third grade. School indicates a student's current school. Standard errors clustered by elementary school in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 5. Effects of DI in Transitional and Non-Transitional Grades

	Math & ELA	Math	ELA	Grades
Baseline DID				
Non-Transitional Grade	0.003	0.004	0.003	0.010
	(0.006)	(0.007)	(0.006)	(0.006)
Transitional Grade	0.029***	0.027***	0.033***	0.018**
	(0.008)	(0.009)	(0.008)	(0.008)
Num. Obs.	847398	841595	843200	911277
School-Cohort FE				
Non-Transitional Grade	0.004	0.007	0.001	0.001
	(0.006)	(0.006)	(0.006)	(0.005)
Transitional Grade	0.028***	0.028***	0.028***	0.006
	(0.008)	(0.008)	(0.009)	(0.007)
Num. Obs.	846917	841129	842745	910512
First Differences (School-Cohort FE)				
Non-Transitional Grade	-0.007*	-0.004	-0.011**	
	(0.003)	(0.004)	(0.004)	
Transitional Grade	0.015***	0.017**	0.011	
	(0.005)	(0.008)	(0.007)	
Num. Obs.	599363	593568	594946	

Notes: Estimated effects of years of DI in elementary school cohort. Treatment is the number of years a student would have spent in DI assuming on-time grade progression in their assigned school in third grade. Transitional grade indicates the grade above the highest grade offered in elementary school. Elementary school indicates the school a student first attends third grade. School indicates a student's current school. Standard errors clustered by elementary school in parentheses. * p < 0.10 ** p < 0.05 *** p < 0.01

Table 6. Effects of DI by Years of Assignment

	Math & ELA	Math	ELA	Science	Grades
Baseline DID					
1 Year DI	0.006	0.008	0.006	0.029***	0.018*
	(0.009)	(0.011)	(0.010)	(0.011)	(0.010)
2 Years DI	0.031**	0.025	0.039***	0.051***	0.026*
	(0.014)	(0.016)	(0.015)	(0.014)	(0.013)
3+ Years DI	0.019	0.023	0.019	0.027	0.028
	(0.019)	(0.021)	(0.018)	(0.019)	(0.019)
Num. Obs.	847398	841595	843200	337933	911277
School-Cohort FE					
1 Year DI	0.002	0.003	0.002	0.021*	-0.004
	(0.010)	(0.010)	(0.010)	(0.011)	(0.009)
2 Years DI	0.028**	0.031**	0.025*	0.052***	0.016
	(0.014)	(0.015)	(0.015)	(0.015)	(0.011)
3+ Years DI	0.024	0.032*	0.018	0.054***	-0.001
	(0.018)	(0.018)	(0.019)	(0.018)	(0.016)
Num. Obs.	846917	841129	842745	337661	910512

Notes: Estimated effects of years of DI in elementary school cohort. Treatment is the number of years a student would have spent in DI assuming on-time grade progression in their assigned school in third grade. Years of DI (row 1) is the number of years a student is observed in DI during elementary school. Elementary school indicates the school a student first attends third grade. School indicates a student's current school. Standard errors clustered by elementary school in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 7. Stacked Difference-in-Differences Estimates

	Math + ELA (1)	Math (2)	ELA (3)	Science (4)	GPA (5)
Overall	0.009	0.008	0.010*	0.010	0.013**
	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
Non-Transitional Grade	0.002	0.003	0.002		0.012*
	(0.006)	(0.007)	(0.006)		(0.006)
Transitional Grade	0.030***	0.024***	0.036***		0.018**
	(0.008)	(0.009)	(0.009)		(0.008)

Notes: Estimated effects of years of DI in elementary school cohort. Treatment is the number of years a student would have spent in DI assuming on-time grade progression in their assigned school in third grade. Transitional grade indicates the grade above the highest grade offered in elementary school. Elementary school indicates the school a student first attends third grade. Timing group represents the group of elementary schools whose cohorts experienced identical patterns of DI exposure. Data are stacked such that each timing group is matched to the entire sample of never-treated elementary schools. Standard errors clustered by elementary school in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 8. Effects of DI on Course-Taking and Behavioral Outcomes

			Courses		*Surveys*
	Absences	Suspensions	Advanced Math	Supplemental Class	Engagement
Baseline DID					
DI Years	-0.012**	0.000	-0.006**	0.008*	0.020**
	(0.005)	(0.002)	(0.002)	(0.004)	(0.010)
Num. Obs.	910164	929646	1107589	1107589	160320
School-Cohort FE					
DI Years	-0.011**	-0.002	0.002	-0.001	0.009
	(0.005)	(0.002)	(0.002)	(0.001)	(0.009)
Num. Obs.	909371	928670	1107589	1107589	160254
Effects of DI by Grade (School-Cohort FE)					
Non-Transitional Grade	-0.014***	-0.002	0.001	-0.001	
	(0.005)	(0.002)	(0.002)	(0.001)	
Transitional Grade	-0.002	-0.002	0.011***	0.000	
	(0.006)	(0.003)	(0.003)	(0.006)	
Num. Obs.	909371	928670	1107589	1107589	

Notes: Estimated effects of years of DI in elementary school cohort. Treatment is the number of years a student would have spent in DI assuming on-time grade progression in their assigned school in third grade. Elementary school indicates the school a student first attends third grade. School indicates a student's current school. Standard errors clustered by elementary school in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

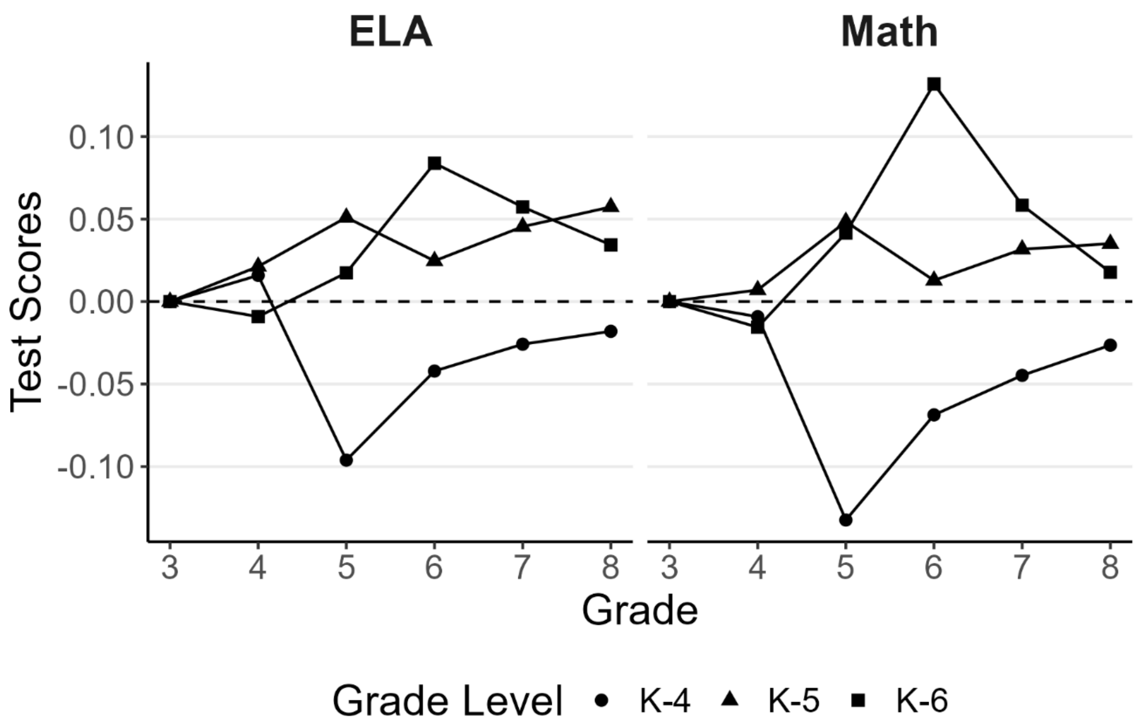
Table 9. Persistence of Elementary Teacher Effects by Instructional Model

	ELA		Math		Combined	
	(1)	(2)	(3)	(4)	(5)	(6)
Long Run	0.961***	0.967***	0.933***	0.948***	0.947***	0.957***
	(0.008)	(0.005)	(0.006)	(0.004)	(0.006)	(0.003)
N	213921	213921	202884	202884	416805	416805
Value Added	0.220*	0.186***	0.235***	0.139***	0.230***	0.158***
	(0.091)	(0.026)	(0.041)	(0.018)	(0.048)	(0.016)
N	171403	171403	173703	173703	345106	345106
Overidentification Test:						
F statistic	0.129		3.932		1.785	
p-value	0.719		0.047		0.182	

Notes: Estimated persistence of teacher effects. Sample includes students in transitional grades (grade higher than offered in elementary school). Estimates are derived from regressions of contemporaneous achievement in grades 5, 6, or 7 on lagged achievement, student race/ethnicity, gender, FRL status, participation in special education or English language learner programs, classroom fixed effects, and (in the second panel) twice-lagged achievement. Estimates in columns (1) and (3) use students whose classrooms were self-contained in the last grade of elementary schools; estimates in columns (2) and (4) use students whose classrooms were departmentalized in the last grade of elementary schools. Long-run persistence is estimated by instrumenting lagged achievement with twice-lagged achievement. The value added persistence is estimated by instrumenting lagged achievement with the lagged teacher's VA estimate. The overidentification test is constructed by fully interacting the model (including instruments) with DI status. Standard errors clustered by elementary school in parentheses. * p < 0.10 ** p < 0.05 *** p < 0.01

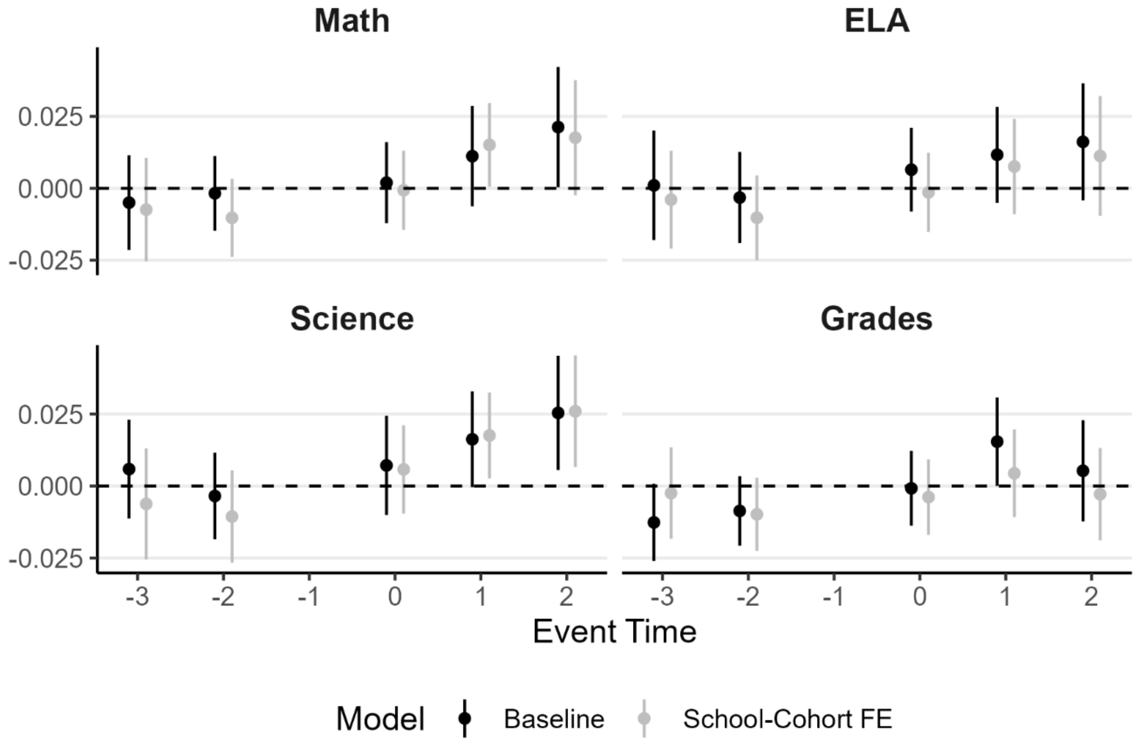
.

Figure 1. Student Achievement by Grade and Elementary School Configuration



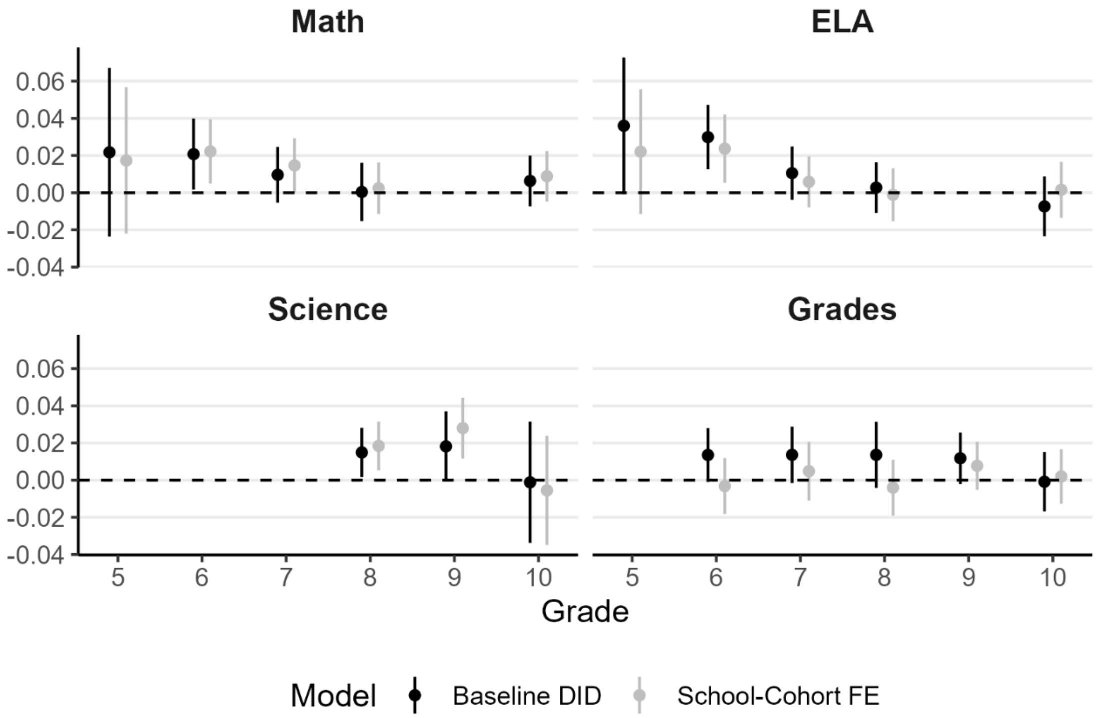
Notes: Standardized student achievement by elementary grade configuration and grade level. Student test scores centered around 3rd grade test scores. Elementary school grade level based on first school with 3rd grade enrollment. Sample includes all students in main analysis sample observed in each of grades 3 – 8 with math and reading scores.

Figure 2. Dynamic Effects of Elementary DI on Students' Long-Run Outcomes



Notes: Estimated effects of years of exposure to DI in elementary school by event time estimated as described in the text. Regression uses cohort leads and lags of DI transformed using method suggested by Freyaldenhoven et al. (2021) for event studies. Effects binned at $t = -4$ and t = 3. Regressions additionally include cohort-by-grade and elementary school-by-grade fixed effects. Confidence intervals constructed using standard errors clustered by elementary school.

Figure 3. Effects of Elementary DI on Test Scores by Grade



Notes: Estimated effects of years of DI in elementary school cohort by current grade. Treatment is the number of years a student would have spent in DI assuming on-time grade progression in their assigned school in third grade. All models include elementary school-grade and cohort-grade FE. Confidence intervals constructed using standard errors clustered by elementary school.