WORKING PAPER

Beyond Tracking and Detracking

The Dimensions of Organizational Differentiation in Schools

Thurston Domina, Andrew McEachin, Paul Hanselman, Priyanka Agarwal, NaYoung Hwang, and Ryan Lewis

RAND Labor & Population

WR-1155

July 2016

This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the RAND Labor and Population Unit.

RAND CORPORATION

For more information on this publication, visit www.rand.org/pubs/working_papers/WR1155.html

Published by the RAND Corporation, Santa Monica, Calif.

Beyond tracking and detracking:

The dimensions of organizational differentiation in schools

Thurston Domina,[1] Andrew McEachin,[2] Paul Hanselman,[3] Priyanka Agarwal,[3] NaYoung Hwang,[3] and Ryan Lewis[3]

Abstract: Schools utilize an array of strategies to match curricula and instruction to students' heterogeneous skills. While generations of scholars have debated "tracking" and its consequences, the literature fails to account for diversity of school-level sorting practices. In this paper we draw upon the work of Sørenson (1970) to articulate and develop empirical measures of five distinct dimensions of school cross-classroom tracking systems: (1) the degree of course differentiation, (2) the extent to which sorting practices generate skills-homogeneous classrooms, (3) the rate at which students enroll in advanced courses, (4) the extent to which students move between tracks over time, and (5) the relation between track assignments across subject areas. Analyses of longitudinal administrative data following 24,000 8th graders enrolled in 23 middle schools through the 10th grade indicate that these dimensions of tracking are empirically separable and have divergent effects on student achievement and the production of inequality.

June, 2016

Acknowledgements: Research reported in this article was supported by the Spencer Foundation (award no. 201400180), by the W. T. Grant Foundation (award no. 183913), and by the Institute for Education Sciences (IES; award R305B120013). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Spencer Foundation, the W. T. Grant Foundation, or IES. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Spencer or W.T. Grant Foundations or IES. The authors thank the following for helpful comments and suggestions: Steven Hemelt, Douglas Lauen, Mark Long, Andrew Penner, Emily Penner, Karolyn Tyson and commenters at the Education and Inequality Seminar at the University of North Carolina, Chapel Hill and the 2016 American Educational Finance and Policy and American Educational Research Association meetings.

[1] Corresponding author. University of North Carolina, Chapel Hill, tdomina@email.unc.edu
[2] RAND Corporation
[3] University of California, Irvine

Schooling may be a “great equalizer” (Mann 1848, Downey et al. 2004; Raudenbush & Eschmann 2015). But at the organizational level, schools are deeply implicated in the production, maintenance, and legitimation of educational inequality. Schools repeatedly sort students, conferring opportunities, resources, and status distinctions unequally in the process (Barr & Dreeban 1983; Kerckhoff 1995). As such, many scholars argue that organizational differentiation practices within schools serve to generate and perpetuate social inequalities. Much of this research focuses on “tracking” – an umbrella term that refers to a broad array of practices associated with the grouping of students into distinct courses of study. The practices that American secondary schools utilize to sort students for instruction have evolved considerably from tracking’s origins in early-Twentieth Century social Darwinist and social efficiency movements (Cremin 1964; Kleibard 1995; Lucas 1999; Oakes 1985; Domina & Saldana 2012). However, controversies surrounding tracking persist and most American public secondary schools continue to sort students into different learning environments and curricula in an attempt to match instruction to diverse learning styles, skills, and instructional needs (Loveless 2013).

The research literature on tracking is diverse. Several studies explore the ideological, political, and technical pressures that lead educators to group students for instruction (Hallinan, 1992; Oakes & Guiton, 1995; Rickles 2011). Others document the relationship between tracking and the distribution of educational opportunities (Oakes, 1985; Argys, Rees, & Brewer, 1996; Kerckoff 1986). Still others consider strategies to detrack schools and improve equity (Burris, & Garrity, 2008; Wells, & Oakes, 1996). Despite its strengths, we argue that this empirical literature on tracking is hindered by an overly simplistic conceptualization and operationalization of tracking itself. Tracking is typically measured as a unidimensional, if not binary, construct.

Some studies compare tracked and untracked schools; others compare students placed in different track locations or courses of study. As a result, the existing empirical literature has strikingly little to say about the specific organizational practices that educators engage in, how these vary between schools and over time, and which matter for student outcomes and inequality. This omission is problematic because it potentially obscures and/or misses important ways that schools contribute to the production of inequality.

In this paper, we draw upon seminal research by Sørenson (1970) and others (Gamoran, 1992; Kelly, 2007; Lucas, 1999; Lucas, & Berends, 2002) to articulate several dimensions of school-level academic tracking systems. Focusing on middle school mathematics and English courses, we hypothesize that school-level tracking systems differ in at least five important ways: (1) the extent to which schools use distinct courses to differentiate curricula, (2) the degree of within-classroom skills homogeneity school tracking practices create, (3) the proportion of students who enroll in high-track courses, (4) the amount of between-track mobility that occurs as students move from middle to high school, and (5) the extent to which course placements are related across subjects. Our project thus contributes to the research literature on tracking and its consequences, which we review below. More generally, our approach suggests new ways to understand and study how institutions structure social inequality. Much of the relevant work on institutional structure in education has taken a macro-level approach, exploring variation in national educational systems (c.f. Buchmann & Park 2009; Hanushek & Woessman 2006; Shavit & Blossfeld 1993). Consistent with recent work on between-firm variation in workplace inequality (Stainback, Tomaskovic-Devey, & Skaggs 2014), we take a meso-level approach, exploring school practices related to instructional organization and their consequences.

We use a unique set of administrative data from 24,000 8th graders in 23 ethnically- and economically-diverse California public middle schools to measure the dimensions of school tracking systems and study their relation to student academic skills development. In contrast to the national probability sample data that are widely used elsewhere in the tracking literature, our data provide detailed longitudinal achievement, demographic, and transcript information for all students enrolled in sample schools. As such, they make it possible to move beyond the prior literature's relatively simple descriptions of school tracking systems to generate time-varying measures these dimensions in each sample school. Our analyses indicate that the dimensions of tracking vary relatively independently between schools and within schools over time. We use this between school and temporal variation in our five measures of tracking to test their effects on student outcomes, allowing us to account for persistent unobserved differences between schools. Our findings indicate that tracking strategies do little to improve average levels of student achievement within schools. However, we find that different dimensions of school tracking systems have independent (and occasionally counter-acting) consequences for student achievement and student achievement inequality. Further, we find some evidence to suggest that school-level tracking systems may exacerbate achievement inequalities within schools by providing a boost for high-achievers relative to their lower achieving peers.

Organizational differentiation and its implications

Our work builds on the theory that schools' organizational differentiation practices have fundamental consequences for student achievement and educational inequality. Sorensen (1970:355) defined organizational differentiation as "the division of a school's student body into subgroups of a permanent character." Organizational differentiation in some form is a practical imperative, and it is difficult to imagine an educational system operating at scale in which all students receive identical instruction at all times. Indeed, the sorting of students into age-based grades is arguably the most fundamental component of the "grammar" of contemporary schools (Tyack & Cuban 1995) and a clear example of organizational differentiation. The question facing educators, then, is generally not *whether* to differentiate instruction, but *how* to differentiate instruction. These decisions likely have important consequences (Sorensen 1989), since the nature of school-level organizational differentiation structures likely shape the style and rigor of the instruction to which students' are exposed (Gamoran & Nystrand 1984), the ability and behavior of their classroom peers (Zimmer 2003), and students' identities as learners (Domina, Penner, & Penner 2016).

Understanding the effects of school tracking systems

The study of tracking and its consequences is central to understanding the role that education plays in the construction of social inequality. Several studies suggest that students in tracked schools demonstrate no greater academic achievement, on average, than students in untracked schools (Hoffer, 1992; Kerckoff, 1986; Slavin 1988). However, there is considerable evidence to suggest that students in high-track classes enjoy a wide range of educational advantages relative to their peers in low-track classes including access to high-achieving peers, high educator expectations, and relatively rigorous instruction (Carbonaro, & Gamoran, 2002; Gamoran & Nystrand 1994; Kelly & Carbonaro, 2012, Van Houte 2004). These educational advantages translate to higher levels of educational achievement, greater access to post-secondary education, and higher levels of ultimate educational attainment (Attewell & Domina, 2008; Long, Conger, & Iatorola, 2012). Further, poor students, students whose parents have relatively low levels of educational attainment, and students of color are all less likely to enroll in high-track classes. Accordingly, much of the research literature suggests that school tracking practices have negligible average effects on student achievement, but that these practices contribute to achievement inequalities by providing relative educational advantages to students in high-track classes.

However, the research literature is by no means unanimous on tracking's impact on achievement and achievement inequality. From a teacher's point of view, tracking is a technical response to pedagogical challenges that almost inevitably arise in educational systems that provide schooling to large and heterogeneous student populations (Hallinan, 1994; Rosenbaum, 1999). One might expect some forms of tracking to help teachers target instruction to their students' needs, yielding positive effects for a broad range of students. Consistent with this hypothesis, a handful of studies using experimental and quasi-experimental methods indicate that sorting students into skills-homogeneous classes has positive achievement effects for students across the skills distribution (Betts, & Shkolnick, 2000; Figlio, & Page, 2002). Further, large-scale policy efforts to create more skills-heterogeneous classroom assignments often have unintended negative consequences for high- and low-achieving students alike (Allensworth, Nomi, Montgomery, & Lee, 2008; Penner, Domina, Penner, & Conley, 2015). Perhaps most notably, Duflo, Dupas, and Kremer (2008) present evidence from an experiment in which students in 61 Kenyan schools were randomly assigned to first-grade classes and students in 60 other Kenyan schools were grouped into classes based on their prior achievement. Their analyses indicate that enrolling in a tracked school has large and lasting positive effects on the achievement of high- and low-achieving students alike. While the extent to which these findings generalize is unclear, the Duflo et al. study provides internally valid evidence regarding the effects of one tracking strategy in one educational setting.

Selection bias and tracking effects

We propose two potential explanations for contradictory evidence regarding the consequences of differentiation on achievement and achievement inequality: First, selection processes – including the nonrandom distribution of tracking practices across schools and nonrandom student-level selection into tracked classes – likely bias estimates of school-level effects of tracked curricula and the student-level effects of tracked course assignment based on observational data (Argys, Rees, & Brewer, 1996; Betts, & Shkolnick, 2000). Researchers interested in estimating the effects of tracking often use regression and propensity-score matching approaches to control for potentially spurious correlations between school tracking systems and student track location and student outcomes. However, these approaches do not account for potentially confounding unmeasured (or imperfectly measured) covariates. The fact that recent studies using random assignment and other quasi-experimental designs find positive effects (Figlio, & Page, 2002; Duflo et al., 2008; Slavin 1990) suggests that unmeasured characteristics may introduce a downward bias on the average effect of attending a school that offers differentiated curricula and an upward bias on the effect of high-track course attendance in other studies (Argys, Rees, & Brewer, 1996).

Conceptualizing "tracking"

Second, we argue that the simplistic conceptualization and measurement of "tracking" in the empirical studies referenced above may contribute to this literature's mixed and ambiguous findings. Scholars utilize a variety of measures to operationalize tracking; including principal reports of school differentiation practices and written school policies related to course assignments (Betts & Shkolnick 2000; Hoffer 1992; Kelly, 2007; Kelly, & Price, 2011), student reports of track assignment (Gamoran & Mare, 1989), teacher reports of classroom composition (Argys, Rees, & Brewer 1996) and transcript-verified measures of student course assignments

(Lucas, 1999). In many cases, these measures impose simplistic categorizations on school tracking systems, classifying schools as “tracked” or “untracked” or dividing students between “vocational” and “academic” tracks. While this literature demonstrates the importance of tracking for educational achievement and inequality, it largely fails to address the ways in which tracking systems likely differ and the consequences of these differences for student outcomes. As an example, the Duflo et al. (2008) study estimates the effects of an isolated change in one dimension of a school tracking system – the degree to which students are grouped by ability into separate classrooms for instruction – but provides little evidence regarding the relations among this change and other dimensions of school tracking systems. Understanding these relations is essential to understanding the social organization of schooling and designing effective and equitable instructional practices.

A handful of studies attempt to operationalize a more nuanced view of school tracking systems. Using school course catalogues and assignment policies to measure the several dimensions of tracking systems, Kelly and Price (2011) find that schools with high levels of variation in student skills are most likely to develop highly differentiated academic tracking systems. Lucas (1999) uses student-level data from the nationally representative High School & Beyond (HSB) to measure the flexibility of secondary school tracking systems, demonstrating that despite the dissolution of an over-arching track system, the curricular experiences of students U.S. high schools remain highly stratified by race and class. Using the same data Gamoran (1992), provides evidence to suggest that different dimensions of school tracking systems have different consequences for students, demonstrating that the achievement effects of enrolling in high-track courses varies across schools. In particular, Gamoran demonstrates that relatively flexible school tracking systems are associated with high levels of mean student achievement and low levels of cross-track achievement inequality.

These studies point to the potential for a more nuanced view of tracking practices for understanding the role that schools play in the production and reproduction of social inequality. However, each faces substantial data limitations. Lacking access to student-level data, Kelly & Price (1999) are unable to test the relationship between tracking systems and student outcomes. Meanwhile, scholars who have studied tracking using NCES cohort-based studies (including the HSB, NELS, ELS, and HSLS) are limited by the paucity of available contextual data (Argys et al. Betts; Figlio & Page; Lucas 1999; Gamoran 1992). These panel studies generally provide detailed data on 20-50 students sampled from each of approximately 500 secondary schools. While this stratified sampling scheme provides data on a nationally representative sample, it situates the student as the unit of analysis and provides limited direct data on the emergent institutional structures in which students are situated. In particular, these panel studies provide limited data about the range of courses schools offer and the ways in which schools sort students across those courses. As a result, several highly salient dimensions of school tracking systems are unobservable in these widely utilized nationally representative panel datasets.

The dimensions of tracking

In this paper, we adopt a framework for conceptualizing tracking and its consequences. We understand school tracking systems as the culmination of an array of school-level processes related to the provision of differentiated academic coursework and the allocation of students among the available courses (Hanselman, Domina, and Hwang 2016). We thus measure track structures as school-level variables. Building upon Sørenson's theoretical work (1970) as well as prior efforts to measure the dimensions of tracking, we develop a framework for thinking about and measuring school tracking systems. We articulate five conceptually distinct dimensions of school tracking systems, and measure these dimensions using administrative data gathered from 23 middle schools in 3 large southern California public school districts. We then link these school-level data to repeated measures of student academic achievement to generate multilevel models of the mean effects of school tracking systems on student achievement. Since students' experiences in school tracking structures likely vary considerably with their own location in these track structures, these mean effects estimates may conceal important inequality-producing consequences of school-level track systems. Accordingly, we investigate the extent to which the effects of school tracking systems vary with students' prior achievement.

Central to this undertaking is the supposition that tracking systems vary on multiple dimensions both across schools and over time. In particular, we identify and measure the following five dimensions of school tracking systems:

1. *Degree of Course Differentiation.* Sørenson (1970, p. 355) defines organizational differentiation as "the division of a school's student body into subgroups of a permanent character." Some form of organizational differentiation is nearly ubiquitous in our setting. The U.S. public education system sorts children into schools by neighborhood and parental preferences. These schools then sort children by age into grades. However, beyond these basic forms of differentiation, schools vary considerably in the degree to which they differentiate curriculum and instruction. Schools may differentiate curriculum and instruction *horizontally*, by providing students with various learning environments in which they can be exposed to different bodies of knowledge, as when a university offers a wide range of graduate seminars focusing on distinct topics. In addition they may differentiate curriculum and instruction *vertically*, by creating different learning environments that expose students to similar bodies of knowledge but at different paces, levels of rigor, and/or with differing degrees of social status. In our conceptualization, schools that offer students a broad range of classes – whether vertically or horizontally differentiated – display a high degree of differentiation (as measured by the number of course offerings), while schools that offer few classes display a low degree of differentiation. All else equal, one might expect course differentiation to have positive consequences for student achievement, since it allows both educators to develop subject-matter and skill-level specializations and students to find classes that match their academic interests and instructional needs.

2. *Cross-classroom ability grouping.* By sorting students across learning environments according to their measured skills, many tracking strategies attempt to simplify the task of instruction. While teachers in skills-heterogeneous (or ungrouped) classrooms may struggle to deliver instruction that is at the appropriate level for a wide range of students (Rosenbaum 1999), skills-homogenous grouped classrooms may allow teachers to provide instruction that is more appropriately tailored to their students (Eccles, & Roeser, 2011). Schools vary in the extent to which their assignment processes generate skills-homogeneous classrooms. Some schools attempt to assign students to courses exclusively on the basis of their prior test scores (Dougherty et al., 2015; Kelly, 2009). However, scheduling constraints and limited resources often restrict educators' discretion over students' classroom assignments. Further, many schools allow teacher recommendations as well as parent and student preferences to influence classroom assignments (Oakes & Guiton 1995; Rickles 2011). As a result, even in otherwise "tracked" schools, students with very different skills levels may sit in the same academic classrooms (Mickelson 2001). Conversely, even in explicitly "untracked" schools, informal pathways may develop that lead students to be grouped based on skills levels across

classrooms (Horvath 2015; Agarwal 2016). Building upon Sørenson's notion of "selectivity," we conceptualize the degree to which schools assign students to skills-homogeneous classrooms as a distinct dimension of tracking systems.[4]

This dimension of tracking systems likely has mixed consequences for students. While skills-homogeneous classroom assignments may allow teachers to target their instruction to student skills; such grouping strategies may broaden skills gaps by exposing high-achieving students to positive peer effects and low-achieving students to negative peer effects (Becker 1987; Epple, Newland & Romano, 2002; but see also Zimmer 2003). Further, skills-homogeneous classroom assignments may create status hierarchies in schools, creating inequalities in learning opportunities and academic expectations across high- and low-achieving classrooms (Kelly & Carbonaro 2012; Metz 1978; Nystrand & Gamoran 1997; Oakes 1985; Page 1991).

3. *Track inclusiveness.* Over the last several decades, policy-makers and educators have undertaken a concerted effort to intensify academic curricula in American schools. This change is particularly noticeable in middle and high school mathematics, where policy-makers have attempted to enroll students in Algebra early in an effort to insure that all students graduate from high school college-ready (Domina, & Saldana, 2012; Domina, McEachin, Penner, & Penner, 2015; Stein et al., 2011). Nonetheless, schools likely continue to vary in the extent to which they expose students to high-level academic content. Some schools enroll all students in courses previously reserved for relatively high-achieving students; others allocate relatively advanced or academically rigorous instruction to some students, and less advanced and rigorous instruction to others (Hanselman et al. 2016). Following Sørenson, we label this dimension of

[4] Indeed our conception of skills-homogeneity is nearly identical to Sorenson's notion of selectivity, which he defines (1970, p. 363) as "the amount of homogeneity that educational authorities intend to produce by the assignment, in terms of the index of learning used, shall be denoted the selectivity of the assignment."

school differentiation systems “track inclusiveness.” Sørenson defines track inclusiveness as “the number of opportunities assumed to be available at different educational levels” (p. 360). Our conceptualization of inclusiveness is arguably a simplification of this conceptualization, since we focus on the relative size of the upper track as a proxy for the more difficult to define and measure distribution of “opportunities.”

If enrolling a student in a more advanced course increases the rigor of the instruction to which they are exposed, one might expect track inclusiveness to boost student achievement. However, there are scenarios in which increases in track inclusiveness might have negative effects. If, for example, many students in a highly inclusive system are exposed to instructional materials for which they are unprepared, inclusivity could have negative effects on student learning (cf Domina, McEachin, Penner, & Penner, 2014). Further, increases in track inclusivity might depress achievement for students left in low-track classes by creating new stigmas associated with this classes (Gamoran 1992).

4. *Track mobility.* School tracking systems likely also vary in the extent to which they create opportunities for students to move between tracks over time. We describe this dimension of school tracking systems as “track mobility,” and seek to distinguish between schools in which track placements are fairly permanent and students have few opportunities to move up or down in a track system from schools in which track placements are relatively fluid over time. Rosenbaum’s classic portrayal of tracking at “Grayton High” (1976) provides an example of a “tournament-style” track mobility system, in which few students move from low-track courses to high-track courses and upward track mobility is thus exceedingly rare. Less rare, however, is downward mobility, or the phenomenon of students moving from high-track courses to low-track courses. Subsequent analyses suggest that this description may not always hold,

indicating that some schools provide opportunities for both upward and downward track mobility (Hallinan, 1996; Lucas, 1999; Lucas, & Good, 2002; McFarland, 2006).

Systems that allow for high degrees of track mobility may be particularly effective at matching students with instruction. If so, exposure to relatively mobile track system may boost student achievement. However, these positive effects may be less common in "tournament" track systems, where upward mobility is rare and downward mobility is common. It is possible that tournament mobility systems may also boost achievement by facilitating an appropriate match between students and instructional offerings and motivating students. Alternatively, one might expect a high degree of tournament mobility to depress student achievement and broaden inequalities by stigmatizing track mobility and associating it with failure.

5. *Track scope.* The tracking system that was common in American secondary schools throughout the first half of the 20th Century sorted students to vocational, general, college preparatory tracks, which typically defined students' secondary school curricula. One distinguishing characteristic of this system, as well the between-school tracking systems that are common in secondary education in much of Europe and Asia, is that it places students into overarching tracks such that students who are exposed to high-level instruction in one subject tend to be subject to high-level instruction in all areas (Hanushek & Woessman, 2006; Lucas, 1999). As such, this system can be said to have a high degree of "scope." As Lucas (1999) documents, American schools dismantled this overarching track system during the 1960s and 1970s, creating a system that theoretically allows students to take high-track classes in some subjects and low-track classes in others. Although Lucas's analyses suggest that track scope remained high in American high schools through the 1980s, he shows that track scope varies

considerably across schools. We consider “scope” as a fifth dimension of contemporary tracking systems.

One might expect scope to relate negatively with student achievement, if schools with high degrees of track scope find it difficult to match students with instruction appropriate for their course-specific skills (Sørenson 1970; Hallinan 1994). High-scope tracking systems may also intensify a tendency toward social closure – or cliquishness – in student peer networks, since it limits the extent to which students have the chance to socialize in class with peers outside of their academic track (McFarland et al. 2014). The resulting social processes may increase the extent to which students identify with their academic track position, exacerbating the association between track assignments and achievement inequality.

Data

In this paper, we operationalize the above five dimensions of tracking using administrative panel data consisting of approximately 24,000 students enrolled as 8th graders during the 2010-11, 2011-12, and 2012-13 school years in 23 Southern California middle schools. Our analyses draw upon student-level administrative data, which districts collect annually from nearly all enrolled students. These data include: student demographics (gender, race/ethnicity, language status, free/reduced lunch eligibility); 7th and 8th grade annual California Standards Test (CST) mathematics and English Language Arts (ELA) scores; transcript data on student middle and high school math and ELA course assignment and performance; course title, teacher ID, and course period data for these middle and high school courses; and California High School Exit Exam (CAHSEE) scores, which provide a standardized measure of student math and ELA achievement in the spring of 10th grade. We supplement these data with qualitative data gathered in interviews with administrators from each district and approximately 25 teachers who teach 8th grade mathematics courses in the three districts.[5]

Table 1 provides a summary of the longitudinal student-level administrative data that we have available from our sample schools during the 2009-10, 2010-11, and 2011-12 school years. Our sample is by no means nationally representative, and in particular our sample schools enroll a disproportionately large number of Latino and Asian-American students and a correspondingly small number of white and African-American students. However, the sample is ethnically and economically diverse. Districts A and B, both of which are among the 10 largest public school districts in California, are situated in inner-ring suburban communities that include both middle class and relatively poor neighborhoods. District C spans an affluent beach community as well as a considerably poorer inland city. The share of students in our sample eligible for the federal Free and Reduced Lunch Program, based on their family incomes, roughly matches the state average (55% in 2010-11) [6].

TABLE 1 ABOUT HERE

Methods

These data provide a unique opportunity to develop nuanced measures of school tracking systems. Since we have a census of transcript, achievement, and demographic data for three cohorts of students enrolled as 8th graders in our 23 sample schools, including teacher and period identifiers, we can identify the classrooms in which students took core academic courses and

[5] We conducted fifteen formal/informal interviews and two focus groups with district administrators, math coaches, and teachers during 2014-2015 school year in the three districts that are part of this study. During the summer of 2014, the team observed six professional development sessions for the three districts and a pilot testing effort at one of the school district with the director of assessments. Interviews consisted of open-ended questions on the district policies and challenges pertaining to student testing, student course placement, curriculum changes and the implementation of new state standards. Observations were done naturalistically and recorded using field-notes. Some focus groups and interviews were audio recorded and transcribed, while others were recorded manually.

[6] http://www.cde.ca.gov/

each of their peers in these classrooms. In addition, we draw upon school and district course listings and academic policy documents as well as interviews with educators at the school and district levels to contextualize these transcript and administrative data. In the analyses that follow, we draw upon these data to measure (1) the degree to which schools offer differentiated curricula in math and ELA, (2) the degree to which schools group students in math and ELA classrooms based on their measured ability, (3) the inclusiveness of high-track math and ELA course placements in schools, (4) the extent to which students experience track mobility in math and ELA between the 8th and 9th grade years, and (5) track scope, or the extent to which students' 8th grade math and ELA course placements correlate.

We first analyze these measures at the school/year level (N=69). To explore the extent to which "tracking" as implemented in contemporary schools is a single practice or a collection of relatively independent practices, we estimate a correlation matrix for our measures of the dimensions of tracking. If tracking is best conceptualized as a single institutional practice, one might expect the dimensions of tracking to correlate highly across schools and over time. Alternatively, weak correlations among the dimensions of tracking suggest that tracking may be better conceptualized as a diverse set of structural elements and practices that are realized in different ways across schools and over time.

In this multidimensional conception of tracking, the school-level practices that define the social organization of instruction likely result from time-variant contextually-specific technical, political, and cultural factors. As such, it seems likely that different school-level factors predict different dimensions of tracking. To test this notion, we estimate a series of mixed models of the following form:

(1) $Y_{sdt} = \beta_0 + \beta_1 X_{sdt} + \alpha_t + \alpha_d + u_s + e_{sdt}$

where Y_{sdt} measures the dimensions of organizational differentiation in 8th grade math and ELA for school *s* in district *d* at year *t*; X_{sdt} is a set of time-varying school-level covariates describing observable characteristics of *s* at time *t* including: school enrollment, an index of school disadvantage calculated as the mean of the standardized proportion of black and Hispanic students in the school, the standardized proportion of students who qualify for free and reduced lunch, the standardized proportion of students who are English-Language learners, students' mean prior achievement levels,[7] and dispersion in students' prior achievement; α_t is a vector of year fixed effects; α_d is a district-level fixed effect; u_s represents school-level random effects; and e_{sdt} is the time-varying school-level error term.

A multidimensional conception of tracking suggests a more nuanced set of answers to historically contentious questions regarding the effects of tracking for student achievement and inequality. If tracking is actually a collection of conceptually and empirically separable practices, it may be possible to develop school structures that realize the potential benefits associated with instructional differentiation while avoiding the costs that are commonly associated with tracking. To address these questions, we use student-level data to investigate the effect of exposure to the dimensions of tracking in 8th grade on students' 10th grade achievement scores. These models take the following general form:

$$(2)\ Y_{ics} = \gamma_{00} + \gamma_{01}(Tracking_{st}) + \gamma_{02}(X_i) + \alpha_t + \alpha_d + u_c + u_s + r_{ics}$$

In these analyses Y_{ics} is students' 10th grade math and ELA test scores as measured on the California High School Exit Exam (CAHSEE). This exam is administered to all students in the spring of their 10th grade year. At the time of its administration to the students in our sample, the

[7] Standardized percent black or Hispanic correlates with standardized percent free or reduced lunch at 0.86 and standardized percent English Language Learner at 0.74. Standardized percent free or reduced lunch correlates with standardized percent English Language Learner at 0.86. School mean test score measures correlate at 0.96. Since school-level standard deviations in math and ELA test scores correlate less closely (0.61), we enter these variables separately into the models.

CAHSEE was a requirement for high school graduation.[8] While the test measures relatively simple skills and is aligned to 6th-8th grade level standards, it is useful for our analyses since it was administered in a consistent form throughout the study period to virtually all students regardless of their skill level, postsecondary plans, and course enrollments. X_{isdt} is a set of student-level characteristics including: demographics and prior achievement as measured by students' 7th grade test scores and grade; α_t is a vector of year fixed effects;α_t and α_d are cohort and district fixed effects; u_c is a class-level random effect; $u_s$is a school-level random effect; and e_{sdt} is the time-varying student-level error term. [9]

The coefficients of interest in this model, γ_{01}, represent the relationship between school-by-year measures of the dimensions of tracking and students' achievement, independent of the other relevant measures of the dimensions of tracking as well as school, district, and year fixed effects and student-level controls. Assuming that student demographics and lagged achievement measures capture the selection of students into schools with different organizational differentiation structure, these models generate unbiased estimates of the independent effects of these dimensions of school tracking systems. Since that assumption is restrictive, however, we fit additional models in which we center each of the tracking measures on their school-level mean. These models thus estimate the effect of tracking exclusively from the within-school variation in tracking systems. Assuming that students do not select into school on the basis of cross-year variation in their 8th grade math and ELA tracking systems, these models generate unbiased estimates of the effects of these school tracking systems on students' achievement.

---

[8] California has since reversed course on the requirement that students pass the CAHSEE.

[9] In both equations (1) and (2), the fixed effects terms represented by α_t and α_d absorb all cross-cohort and district-level variation in the outcomes. By contrast, the school-level random effects term u_s (as well as the course-level random effects term u_c in equation 2) simply account for the non-independence of repeated observations of schools across time.

Finally, to understand the extent to which tracking practices work to exacerbate achievement inequalities within schools, we add an interaction between students' 7th grade test scores and the school-mean centered version of the school dimension of tracking. Positive values on these interaction terms suggest that tracking practices magnify the association between 7th grade test scores and 10th grade test scores, as one would expect if tracking increases achievement inequality. For the purposes of simplicity, we interpret results in terms of the predicted associations between 7th and 10th grade achievement under different tracking regimes and report these interactions graphically.

Measuring the dimensions of tracking

Based on a review of school course catalogues as well as conversations with educators at sample schools and districts, we categorize 8th grade math and ELA courses into three levels: advanced, college prep, and remedial. We refer to the middle track as "college prep" since it is designed to prepare students to complete the high school course sequence required for admission to the four-year colleges in the University of California and California State University systems. As Figure 1 indicates, schools tend to place relatively high-achieving students in advanced and honors courses, students at the middle of the test score distribution in college prep courses, and low-achieving students in remedial courses. However, consistent with Mickelson (2003) we also find evidence of considerable skills-heterogeneity among students in each of these tracks.[10]

[10] To quantify the degree to which the observed track placements deviate from a model in which students are placed in 8th grade courses strictly on the basis of their content-relevant 7th grade test scores, we conducted a simple simulation in which we rank-ordered students relative to their 8th grade school peers based on their 7th grade student scores. Then, taking the relative magnitude of schools' 8th grade advanced/honors, college preparatory, and remedial math and ELA tracks as a given, we constructed simulated counterfactual course assignments for all students in our sample based on a simple model in which students queue for high level courses based on their 7th grade test score rank. (Such that in a school in which x students enroll in advanced mathematics courses, y students enroll in college prep track mathematics courses, and z students enroll in remedial mathematics courses; students in the top x of the 7th grade test score distribution are assumed to be placed in advanced courses, students in the top x+y in college prep courses, and the remaining students place in remedial courses.) The exercise reveals an approximately 70 percent correlation between students' actual 8th grade math and ELA track assignment and their simulated assignments.

FIGURE 1 HERE

In this paper, we move beyond the broad representation of tracking systems represented in Figure 1 and empirically measure each of the five dimensions of school tracking systems. Since we have access to testing and transcript data in mathematics and ELA for every student in our sample schools, we can identify the title and level of all courses that sample schools offer to 8th graders in these key academic areas. In addition, by identifying students who take the same class with the same teacher during the same school period, we can identify every peer in 8th graders' math and ELA classrooms. These data allow description of schools' tracking systems and students' places in these systems. Table 2 provides a descriptive overview of our measures of the five dimensions of tracking.

TABLE 2 HERE

Course differentiation is the range of different topics and activities that a school makes available to students. We measure the degree to which mathematics and English instruction is differentiated in our sample schools as the number of different course titles schools make available to 8th graders in any given year. As Table 2 reveals, the schools in our sample offer an average of four mathematics classes during the study period. However, schools vary appreciably on this measure. We observe schools that offer as few as two distinct 8th grade mathematics courses (Algebra and Pre-Algebra) and schools that offer as many as seven (including a remedial General Mathematics Skills course, Pre-Algebra courses in English and Spanish, Algebra courses in English and Spanish, an Honors Algebra course, and a doubly-advanced Honors Geometry course.) While our sample schools offer slightly fewer ELA courses to 8th graders during the study period, we observe no less cross-school variation in 8th grade ELA course offerings.

We measure the degree of **skills-homogeneous classroom assignments** in schools' 8th grade math and English classes by using students' 8th grade classroom assignments to predict their 7th grade standardized test scores within each school and year for which we have data. The intraclass correlation (ICC) from this multi-level model captures the amount of between class variation that exists within a given school-by-year based on students' prior achievement. We interpret this ICC as the degree of skills homogeneity in 8th grade mathematics and English classrooms in a school in a given year on a zero-to-one scale. This measure has a mean of 0.52 in our sample schools and a standard deviation of 0.17 for mathematics and a mean of 0.50 and a standard deviation of 0.18 in ELA.

Conversations with school and district leaders reveal substantial variation in course assignment policies, both across schools and within schools over time. Throughout the study period, District B encouraged schools to enroll students in 8th grade math and ELA exclusively on the basis of the prior test scores. While teachers report that they occasionally overruled the district's placement formulae, our analyses indicate that classroom assignments are relatively skills-homogeneous in District B over time. By contrast, Districts A and C gave schools relatively little guidance regarding course placements. In District A, schools typically used a fairly informal approach to course assignments, allowing teachers, parents, and teachers to place students independently of their prior test scores. Finally, schools in District C experimented with an array of course assignment practices over time, ranging from explicitly skills-heterogeneous course assignments to rigid test-score based assignments.

Figure 2 provides an illustration of our measure of homogeneous classroom assignments, plotting the distribution of 7th grade mathematics test scores by 8th grade mathematics classroom for 8th graders in one District C school in 2010 and 2012. During this period, this school moved from an informal course placement system to a system that explicitly attempts to create skills-heterogeneous classrooms in middle-track mathematics. In the process, the schools' skills-homogeneity measure decreased from 0.51 to 0.24, a change equivalent to approximately 1.5 standard deviations in the sample-wide distribution. There is considerable overlap across classrooms in the distribution of student achievement in both years. However, in 2010, the bulk of students scored within 25-30 points of their classroom mean. The distribution of scores within classrooms is considerably broader in 2012, especially in the 9 middle-track mathematics classrooms where a large proportion of students score more than 50 points higher or lower than their classroom mean (roughly a standard deviation in 7th grade CST scores among the 24,000 students for which we have data).

FIGURE 2 HERE

Track inclusiveness refers to the extent to which schools assign students to high-track courses. We measure inclusiveness as the proportion of 8th graders enrolled in accelerated or honors-level courses in our sample schools. As Table 2 indicates, we observe a higher degree of track inclusiveness in mathematics in our sample schools than in ELA. This is likely largely due to a policy effort to use state educational accountability policies to encourage schools to boost 8th grade Algebra enrollments. While the state began to moved away from this effort as it transitioned to the Common Core State Standards in both math and ELA, California schools continued to enroll students in 8th grade Algebra – a course we consider accelerated since it puts students on a track to complete Calculus by the end of 12th grade – at a considerably higher rate than their peers across the U.S. (Domina, McEachin, Penner, & Penner, 2015). The state's Algebra-for-all effort limits the degree of variation in math track inclusiveness in our sample schools. However, we observe a large degree of both between-school variation as well as within-school temporal variation in ELA track inclusiveness, where the mean is 0.63 and the standard deviation is 0.27.

Track mobility refers to the extent to which students' move across track levels as they progress through school. While our sample schools enrolled a large proportion of students in advanced courses in both math and ELA during their 8th grade year, these middle school placements do not ensure that students will remain on an advanced track through high school. Consistent with Rosenbaum's observations in "Grayton High" (1976), we find that virtually no students in our sample schools move from 8th grade remedial classes to 9th grade college prep classes or 8th grade college prep classes to 9th grade advanced classes. However, 41 percent of the students in our sample schools experienced downward mobility in mathematics between 8th and 9th grade and 34 percent experienced downward mobility in ELA.[11]

We use the proportion of a school's 8th graders in advanced or college prep courses who repeated the same course in 8th and 9th grade or took a lower-level course in 9th grade than 8th to measure the degree of downward mobility in school tracking systems.[12] In interviews, teachers and district leaders report that they prefer to place students in relatively high-level middle school courses, so as not to foreclose students' opportunities to take advanced courses later in their educational careers. There is some evidence to suggest that state policy around 8th grade Algebra reinforced this tendency (see Domina et al. 2015 for more detail), leading schools to create nominally "accelerated" 8th grade Algebra classrooms in which the vast majority of students retook Algebra as 9th graders. As Table 2 indicates, this arrangement, which is measured as the

---

[11] Many of educators we interviewed expressed frustration at the lack of upward track mobility in their schools. Curricular planners in Districts A and C have dedicated particular attention to attempting to facilitate upward mobility by creating multiple "course acceleration" opportunities, including double-dose and summer courses. However, these efforts to create upward mobility paths were not in place in sample schools during the study period.

[12] Most downward mobility in mathematics occurred when students took Algebra in the 8th grade and retook it in the 9th grade. In ELA, a the most common example of downward mobility is from 8th grade Honors course to a 9th grade College Prep course.

proportion of students who experience downward track mobility in 9th grade course placement varies appreciably within and between schools in both mathematics and ELA.

Finally, **Track Scope** refers to the relation between students' classroom assignments during one part of the school day and their assignments during the rest of the day. Following Lucas (1999), we measure scope as the correlation between 8th grade mathematics course placements and 8th grade ELA courses placements. In schools that approach 1 on this measure, students who are assigned to high-track mathematics courses are typically also assigned to high-track ELA courses. In schools that approach 0 on this measure, mathematics and ELA courses placements are largely unrelated. On average, this measure of scope is fairly high in our sample schools, and students' math course assignments correlate with their ELA course assignments at 0.67. This correlation corresponds closely with Lucas's (1999) findings regarding track scope in a nationally representative sample of U.S. high schools. Underlying this measure, however, we find considerable variation in track scope between schools as well as temporally within schools. In some schools, students' math track placements rarely diverge from their ELA course placements while in others it is not uncommon for students to enroll in advanced math and college preparatory ELA courses (or vice-versa.)

Testing a multi-dimensional conception of tracking

In the popular conception, a highly "tracked" school is one in which curricula are highly differentiated, students are grouped into very skills-homogeneous classrooms, access to high-track classes is constrained to a relatively small proportion of high-achieving students, track scope is high and track mobility is low. In this conception, it seems reasonable to categorize schools as "tracked" or "untracked" and to expect little movement among these categories within a school over time. However, as the discussion above indicates, "tracking" is a multidimensional construct and the dimensions of tracking need not closely covary. At least in principle, schools can offer a highly differentiated curriculum composed of a wide array of distinct courses even as they place students into highly skills-heterogeneous classrooms. Likewise, schools can in theory maximize track scope by having students spend the entire school day with the same set of peers even as they maximize track mobility by changing students' location in the track system year after year. Furthermore, the descriptive statistics in Table 2 indicate that these dimensions of school tracking systems vary in a continuous fashion both across schools and within schools over time.

The correlation matrix reported in Table 3 investigates the extent to which the theoretically separable dimensions of school tracking systems are separable in practice among our 69 school observations. We observe close associations between our measures of track inclusiveness and track mobility. Schools that enroll large proportions of students in advanced courses in 8^{th} grade tend to have more students who make downward moves in the track system in 9^{th} grade. This correlation is particularly pronounced in ELA, at 0.95.

More generally, however, Table 3 indicates that the correlations among the dimensions of tracking are low. For example, while schools that sort students into relatively skills-homogeneous math classes tend to have lower levels of enrollment in advanced math classes and lower levels of track mobility, these associations are fairly small at -0.18 and -0.16 respectively. In ELA, the correlations between skills-homogeneous course assignment and track inclusiveness and mobility are even smaller. Further, the associations between skills-homogeneous classroom assignments in both mathematics and ELA and track scope are also quite small. We observe positive associations between the degree of curricular differentiation in schools and the degree of within-classroom ability grouping, consistent with the idea that curricular differentiation facilitates the sorting of students into skills-homogeneous classrooms. In both mathematics and ELA, we find that as the number of courses schools increases so to does its practice of skills homogeneous classroom assignments. However, these associations are quite modest, at 0.36 and 0.44 respectively.

TABLE 3 HERE

Consistent with Table 3, the multilevel models reported in Table 4 indicate that associations between school characteristics and school tracking practices vary across the dimensions of tracking. In these models both the dependent variables and the independent variables are standardized, so that the coefficients can be interpreted as the expected increase in the dimensions of tracking (expressed in standard deviation terms) associated with a one standard-deviation increase in each of the independent variables, conditional on all other controls.

While we find evidence to suggest that the degree of mathematics curricular differentiation significantly varies across districts and over time, none of our measured school characteristics significantly predict the number of different mathematics courses offered by schools in our sample. Similarly, we find no significant association between school characteristics and ability grouping in mathematics. Indeed, the only relatively consistently significant school-level predictor of school mathematics tracking systems is schools' total enrollment. In particular, these analyses indicate that relatively large schools tend to enroll a large proportion of students in advanced 8^{th} grade math courses, but that students in these large schools tend to experience relatively high rates of downward track mobility in mathematics between 8^{th} and 9^{th} grade. Since the relatively small school-level sample size limits the power in these analyses, the nonsignificant negative conditional associations between school socioeconomic disadvantage and all four mathematics tracking dimensions is worth noting. These non-significant associations indicate that schools that educate relatively large proportions of poor, minority, and EL students may tend to offer fewer mathematics courses and place students in relatively heterogeneous mathematics classes. School mean prior year achievement is negatively related to three of the four mathematics-specific tracking dimensions. Notably, school achievement is a significant negative predictor of school level downward track mobility rates.

[TABLE 4 HERE]

The pattern of school-level predictors of the dimensions of tracking in ELA is somewhat different. We find that schools with relatively disadvantaged student populations tend to offer significantly *more* 8th grade ELA courses than more advantaged schools, net of controls. However, school socioeconomic disadvantage is a significant negative predictor of track inclusiveness and downward track mobility. Consistent with mathematics, we find that school mean prior achievement relates negatively with curricular differentiation, ability grouping, and downward mobility in ELA. Finally higher school enrollment is positively associated with all four dimensions of ELA tracking, although this association is only significant for downward track mobility. Finally, we find that school disadvantage negatively predicts track scope while higher school enrollment positively predicts it.

The effects of the dimensions of tracking

In light of the above evidence suggesting that the dimensions of tracking are empirically separable, the remaining analyses examine the links between these dimensions and student achievement. Table 5 reports the results of a series of multilevel models regressing the dimensions of mathematics tracking systems in students' 8th grade middle schools on students' 10th grade math achievement; Table 6 reports the results of parallel analyses in ELA. All dependent and independent variables are standardized in each of the models reported in both tables, such that each has a mean of zero and a standard deviation of one in the student population under analysis. The first model in Table 5 provides an unconditional look at these relationships. The second model adds student-level demographic and prior achievement controls as well as indicator variables that account for commonalities among students enrolled in the same school district (district fixed effects) and students in the same grade cohorts (cohort fixed effects). Finally, in the third model, we mean-center the time-varying school-level measures of the dimensions of tracking around schools' 3-year mean scores on these measures. Doing so controls for time-varying school characteristics that may confound the link between the dimensions of school tracking regimes and student achievement.[13]

TABLE 5 AROUND HERE

The first model of Table 5 indicates that there is no average association between the number of courses that schools offer in mathematics and students' mathematics achievement. This null relationship continues to hold as we add background controls in Model 2 and condition on time-invariant school characteristics in Model 3. Similarly, the relation between track scope and math achievement is nonsignificant and substantively small in all three models in Table 5, suggesting that track scope is unrelated to student achievement.

By contrast, we find that students in schools that have relatively skills-homogeneous 8th grade math classroom assignment practices score significantly less well on 10th grade courses than their peers in schools where math courses are less rigidly grouped by student achievement. That association continues to hold after adding student-level controls in Model 2 and after controlling for time-invariant school characteristics in the school mean-centered Model 3.

---

[13] All models include school and classroom level random effects terms to adjust standard error estimates for the clustering of students in schools and classrooms.

Accordingly, the analyses presented in Table 5 suggest that that homogeneous math course assignment practices may have small negative effects on students' mathematics achievement.

Models 1 and 2 of Table 5 suggest that there is no average association between 8th grade school math track inclusiveness and students' 10th grade mathematics achievement, even after controlling for student background characteristics. However, after controlling for time-invariant school characteristics in Model 3, we find evidence to suggest that attending middle schools with highly inclusive 8th grade math tracking systems significantly *depresses* student achievement by the 10th grade. This estimate suggests that, all else equal, a student who moves from a school that enrolls 73 percent of students in advanced math classes to a school that enrolls 85 percent in advanced math classes will experience a test score decline of approximately 1/12th of a standard deviation. While somewhat counter-intuitive, these findings are consistent with evidence elsewhere in the research literature suggesting that efforts to intensify middle school mathematics curricula may have unintended negative consequences for students' achievement (Clotfelter, Ladd & Vigdor 2014; Domina, McEachin, Penner & Penner 2015).

Interestingly, we find the reverse association between track mobility and student achievement. Our reduced-form model indicates that students in schools where downward track mobility is relatively common score less well in 10th grade than their peers in schools with less downward mobility. Since downward track mobility in this context typically takes the form of failing 8th grade Algebra and repeating it in the 9th grade, it is not surprising that the association is negative. Notably, however, this association seems to be entirely driven by between-school variation. After controlling for time-invariant school characteristics in Model 3, we find a positive conditional link between downward track mobility and 10th grade achievement. Given this relationship, it appears that conditional on prior achievement, students in schools that provide them a math class closer to their ability level do better than students who continue in a track that may be too challenging for them. Taken together, our findings for track inclusiveness and track mobility suggest that students may not benefit on average when they and/or a large proportion of their peers are placed in advanced courses. While we are unable to explore the mechanisms through which these negative average effect operate, they are consistent with the idea that students are especially likely to succeed when instruction is matched to their skill level and cognitive needs (Domina 2014).

Table 6 reports parallel models exploring the relation between the dimensions of school ELA tracking systems and students' 10^{th} grade ELA achievement. In general the results reported here indicate that ELA test scores are less sensitive to the dimensions of tracking than math scores. While we find that 8^{th} grade ELA track differentiation, inclusiveness, and mobility are all associated with 10^{th} grade ELA scores, none of these associations are significant after controlling for student characteristics and time-invariant school characteristics. Perhaps most notably, Models 2 of Table 6 indicates that students who attend schools with high levels of downward track mobility score lower on average on 10^{th} grade ELA tests than similar peers in schools with less downward mobility. However, Model 3 indicates that this association is largely a function of unmeasured school effects. We find no evidence to suggest that school level *changes* in ELA track mobility rates are associated with students' 10^{th} grade ELA test scores.

TABLE 6 AROUND HERE

Taken together, the results reported in Tables 5 and 6 suggest that the constellation of practices researchers often refer to as "tracking" have mixed and modest average effects on student achievement. We find that placing students into ability grouped 8^{th} grade mathematics classrooms has a small negative effect on students' mean 10^{th} grade mathematics achievement.

However, our findings regarding the average effects of track inclusiveness and mobility suggest that efforts to detrack mathematics instruction by enrolling all students in accelerated courses may have unintended negative consequences. Meanwhile, we find no evidence to suggest that any of the dimensions of 8th grade ELA tracking systems influence student achievement in ELA.

However, since the analyses reported in Tables 5 and 6 focus on the mean effects of school-level tracking systems, they neglect crucial questions regarding to the effects of tracking systems on achievement inequality. Figures 3 and 4 address the equity effects of tracking by taking a closer look at one key dimension of school tracking systems – the degree to which schools group students into classrooms based on their prior test scores. Building on the third models in Table 5 and 6, these figures illustrate the results of models in which we investigate the extent to which the effects of school-level ability grouping vary with students' 7th grade test scores.

FIGURES 3 AND 4 AROUND HERE

The y-axis in this graph represents students' z-scored predicted 10th grade mathematics achievement scores, while the x-axis represents students' z-scored 7th grade mathematics scores. The dashed line represents the predicted relation between 7th grade achievement and 10th grade achievement in mathematics for students in schools that have implemented ability grouping to an above-average degree in 8th grade mathematics classrooms. The solid line, meanwhile, represents that same relation in schools that have implemented a below-average degree of ability grouping in 8th grade mathematics classrooms. The shaded areas around both lines represent 95% confidence intervals. Consistent with the results indicating a negative average effect of ability grouping in 8th grade mathematics reported in Table 5, the dashed line is lower than the solid line across the 7th grade math test score distribution in Figure 3. Notably, however, the disadvantage associated with attending a school in which students attend largely skills-homogeneous 8th grade mathematics courses is particularly pronounced for students at the bottom of the 7th grade mathematics test score distribution. Put differently, this figure suggests that low-achieving students disproportionately bear the achievement costs associated with ability grouping in middle school mathematics. The full model, reproduced in Appendix Table 3, indicates that this interaction term is highly statistically significant, if small in magnitude. All else equal, this model suggests that enrolling in a school with a high degree of ability grouping will increase the gap between students who come into the 8th grade 1 standard deviation above and below the math test score average by approximately 0.08 standard deviations.

Figure 4, and the corresponding model reported in Appendix Table 3, suggests that the null average effect of homogeneous ELA classroom assignments reported in Model 3 of Table 6 conceals importantly variable effects across the skills distribution. While low-achieving students experience negative achievement effects when then enroll in a middle school with a high degree of skills-based sorting across 8th grade ELA classrooms, high achieving students experience positive effects. As in the case of mathematics grouping, the interaction with prior skills is highly statistically significant. While these interactions are arguably small, they are notable since they suggest that ability grouping – a strategy that is ostensibly designed to improve instruction for all students – broadens within-school achievement inequalities.

Discussion

This study is the first to rigorously measure multiple dimensions of tracking and identify their effects on student achievement. Building on the work of Sørenson (1970) and others (Gamoran 1992; Lucas 1999; Kelly, 2007; Becker, 1987), we identify five theoretically distinct dimensions of school math and ELA tracking systems: (1) curricular differentiation, (2) classroom ability grouping, (3) track inclusiveness, (4) track mobility, and (5) track scope. We take advantage of a unique set of student-level administrative data gathered from 3 medium- to large-enrollment public school districts, as well as qualitative data gathered from administrators and educators in these three districts, to measure the dimensions of tracking systems in 23 elementary schools and the ways in which these tracking systems changed over the course of three years.

Our findings indicate that "tracking" is a multidimensional phenomenon in contemporary secondary schools. We observe considerable variation on each dimension both between our sample of 23 middle schools and within these schools over time. Further, we find that the dimensions of school tracking systems do not highly correlate with one another. Our findings thus suggest that widely used methods in the tracking literature obscure important organizational variation in track practices and their consequences. Our findings indicate that these dimensions of school tracking practices are fairly independent of one another (with observed school-level correlations in the 0.2-0.4 range). Further, consistent with a multidimensional conception of school tracking practices, our analyses indicate that the predictors of school tracking systems vary across the dimensions of tracking.

Our investigation of the dimensions of tracking reveals new insights into the ways in which school tracking systems influence student achievement. In the area of ELA, our findings are largely consistent with earlier sociological research on the effects of school tracking. While we find little evidence to suggest that the five dimensions of school tracking systems have an effect on student achievement, this null effect conceals important inequality-producing consequences of school tracking systems. In particular, we find that when schools group students into ELA classes based on their prior achievement, high-achieving students tend to experience rapid test score growth in ELA while low-achieving students fall behind.

Furthermore, the dimensions of tracking have potentially cross-cutting effects on students' mathematics achievement growth. We find evidence to suggest that ability grouping has a weak negative effect on achievement growth in mathematics and that this negative effect is particularly pronounced for low-achieving students. This finding suggests that students may benefit from placement in relatively skills-heterogeneous classrooms for secondary mathematics instruction. Based on this finding, it is tempting to recommend that schools eliminate low-track classes and attempt to enroll all students in high-level courses. Many recent policy efforts narrow inequalities in opportunities to learn in U.S. secondary schools have taken exactly this approach, attempting to expose all students to high-quality instruction and high-achieving peers by universalizing accelerated course placements.

However, our analyses also reveal negative effects of math track inclusiveness and mobility on student mathematics achievement. Consistent with several recent policy analyses (Clotfelter, Ladd, & Vigdor 2015; Domina, McEachin, Penner & Penner, 2015; Stein et al., 2011), these findings suggest that efforts to detrack instruction by enrolling more students in accelerated courses can have unintended negative effects if they lead students to courses for which they are academically unprepared. From a practitioner's perspective, therefore, our findings point to a tension between the benefits of skills-heterogeneous learning environments and the shortcomings of instruction that is insensitive to student skills. Curricular reform efforts that simultaneously provide disadvantaged students with access to higher achieving peers *and* sufficient skill-building opportunities provide one promising strategy for resolving this tension (Nomi & Allensworth 2012, Nomi & Raudenbush 2016).

Much research on tracking – and indeed, much research in the sociology of education and inequality – takes an individualistic approach, focusing in the case of the tracking literature on the consequences of students' track locations. By contrast, our approach is more explicitly organizational. Since we posit that organizational context matters both in the ways in which "tracking" is realized in practice and in the consequences of these practices, we measure school contexts using quantitative data on all students in sample schools as well as qualitative data on schools and districts. Our estimates thus speak to the school-wide effects of tracking systems and the ways in which these effects vary with student prior achievement.

This organizational approach has multiple advantages. First, it seems unlikely that the processes that determine which students enroll in which schools are likely to change appreciably year-to-year in ways that would confound estimates of the effects of the dimensions of school tracking systems. Accordingly, our estimates of the effects of school tracking systems are likely less subject to selection biases than estimates of the effects of track location. Second, the effects of school-level tracking systems are substantively important. Educational researchers from Coleman (1966) to contemporary scholars interested in teacher effects (c.f. Chetty, Friedman, & Rockoff, 2013) emphasize within-school variation in student achievement, raising questions about the extent to which variation across schools in quality matters in the production of educational inequality (Jennings et al. 2015). Our analyses shed light on school-level practices that produce (or ameliorate) within-school inequalities. As such, they indicate that organizational differentiation may account for some of the within-school achievement variation that scholars often attribute to teacher or student background factors. Our findings thus point to a frequently overlooked way in which schools and their organizational processes shape student achievement and achievement inequality.

More generally, this meso-level examination of the curricular differentiation patterns within schools contributes to sociological understanding of the organizational bases of social inequality. As Stainbeck, Tomaskovich-Devey, and Skaggs point out (2010, p. 226), "organizations are the primary site of the production and allocation of inequality in modern societies." This insight has stimulated considerable research in the sociology of work, where scholars have demonstrated considerable variation across firms in the degree of wage inequality (Avent-Holt Tomaskovich-Devey 2012), attributable at least in part to firm-level organizational characteristics and practices (Kalev, Dobbin, & Kelly 2006.) This body of research suggests that even the most durable workplace inequalities are contingent on local circumstances. Further, this approach points to important opportunities for organizational actors to narrow inequalities, even as it acknowledges the ways in which macro-level forces structure social inequalities (Tomaskovich-Devey 2014).

Given the increasing availability of administrative data from U.S. public schools, the time is ripe to extend this insight to the study of education and social inequality. Schools operate in diverse funding and policy climates (Jackson, Johnson, & Persico 2014; Meyer, Rowan & Meyer 1978; Reed 2014) and serve widely varying student populations (Reardon & Owens 2014). Despite the presence of cultural and institutional pressures to conform to a broadly accepted "grammar of schooling" (Tyack & Cuban 1995, Weick 1976), schools vary considerably in many respects that are relevant to the production of educational inequality (c.f. Fiel 2015; Hanselman, Domina, Hwang 2016; Legewie & DiPrete 2012; Raudenbush & Bryk 1986). As this paper demonstrates, careful study of this organizational variation can shed light on the complex and interacting mechanisms through which schools produce, reproduce, and even ameliorate social inequality. Ultimately, such an approach may point to promising strategies for building more effective and equitable organizations.

References

Agarwal, Priyanka, and Thurston Domina. 2016. "Evidence of “concealed” within-course sorting in middle school algebra: A case study of three school districts." Manuscript in preparation.

Alexander, Karl L., and Edward L. McDill. 1976. "Selection and allocation within schools: Some causes and consequences of curriculum placement." *American Sociological Review*: 963-980.

Allensworth, Elaine, Takako Nomi, Nicholas Montgomery, and Valerie E. Lee. 2009. “College preparatory curriculum for all: Academic consequences of requiring algebra and English I for ninth graders in Chicago” *Educational Evaluation and Policy Analysis*, *31*(4): 367-391.

Attewell, Paul, Thurston Domina, David Lavin, and Tania Levey. 2004. "The black middle class: progress, prospects, and puzzles." *Journal of African American Studies 8*(1-2): 6-19.

Brewer, Dominic J., Daniel I. Rees, and Laura M. Argys. 1996. “Detracking America's Schools: Equity at zero cost?.” *Journal of Policy analysis and Management*, *15*(4): 623-645.

Avent-Holt, Dustin, and Donald Tomaskovic-Devey. 2012. “Relational Inequality: Gender Earnings Inequality in US and Japanese Manufacturing Plants in the Early 1980s.” *Social Forces*, 91:157-80.

Becker, Henry Jay. 1987. “Addressing the Needs of Different Groups of Early Adolescents: Effects of Varying School and Classroom Organizational Practices on Students from Different Social Backgrounds and Abilities. Report No. 16.” Baltimore: Johns Hopkins University Center for Research on Elementary and Middle Schools.

Betts, Julian R., and Jamie L. Shkolnik. 2000. “The effects of ability grouping on student achievement and resource allocation in secondary schools.” *Economics of Education Review*, *19*(1): 1-15.

Buchmann, Claudia, and Hyunjoon Park. 2009. "Stratification and the formation of expectations in highly differentiated educational systems." *Research in Social Stratification and Mobility* 27(4): 245-267.

Burris, Carol Corbett, and Delia T. Garrity. 2008. *Detracking for excellence and equity*. ASCD.

Carbonaro, William. 2009. "Tracking, students' effort, and academic achievement." *Sociology of Education* 78(1): 27-49.

Carbonaro, William J., and Adam Gamoran. 2002. "The production of achievement inequality in high school English." *American Educational Research Journal* 39(4): 801-827.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2013. "*Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood* (No. w19424)." National Bureau of Economic Research.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2015. "The Aftermath of Accelerating Algebra Evidence from District Policy Initiatives." *Journal of Human Resources*, *50*(1): 159-188.

Coleman, James S. 1966. *Equality of educational opportunity.* Washington, DC: U.S. Government Printing Office.

Cremin, Lawrence A. 1965. *The genius of American education*. Pittsburgh: University of Pittsburgh Press.

Domina, Thurston. 2014. "The link between middle school mathematics course placement and achievement." *Child development 85*(5): 1948-1964.

Domina, Thurston, Andrew M. Penner, and Emily K. Penner. 2016. "Membership Has Its Privileges': Status Incentives and Categorical Inequality in Education." *Sociological science* 3 (2016): 264-295.

Domina, Thurston, Andrew McEachin, Andrew Penner, and Emily Penner. 2015. "Aiming High and Falling Short California's Eighth-Grade Algebra-for-All Effort." *Educational Evaluation and Policy Analysis 37*(3): 275-295.

Domina, Thurston and Joshua Saldana. 2012. "Does raising the bar level the playing field? Mathematics curricular intensification and inequality in American high schools, 1982–2004". *American Educational Research Journal 49*(4): 685-708.

Downey, Douglas B., Paul T. Von Hippel, and Beckett A. Broh. 2004. "Are schools the great equalizer? Cognitive inequality during the summer months and the school year." *American Sociological Review* 69(5): 613-635.

Dougherty, Shaun M., Joshua S. Goodman, Darryl V. Hill, Erica G. Litke, and Lindsay C. Page. 2015. "Middle school math acceleration and equitable access to eighth-grade algebra evidence from the wake county public school system." *Educational Evaluation and Policy Analysis* 37(1 suppl): 80S-101S.

Dreeben, Robert and Rebecca Barr, 1998. "Classroom Composition and the Design of Instruction." *Sociology of Education* 61(3): 129-42.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2008. *Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya* (No. w14475). National Bureau of Economic Research.

Eccles, Jacquelynne and Robert W. Roeser. 2011. "Schools as developmental contexts during adolescence." *Journal of research on adolescence* 21(1): 225-241.

Epple, Dennis, Elizabeth Newlon, and Richard Romano. 2002. "Ability tracking, school competition, and the distribution of educational benefits." *Journal of Public Economics* 83 (January): 1–48.

Fiel, Jeremy. 2015. "Closing Ranks: Closure, Status Competition, and School Segregation." *American Journal of Sociology* 121(1): 126-170.

Figlio, David N., and Marianne E. Page. 2002. "School choice and the distributional effects of ability tracking: does separation increase inequality?." *Journal of Urban Economics* 51(3): 497-514.

Gamoran, Adam. 1992. "The variable effects of high school tracking." *American Sociological Review*: 812-828.

Gamoran, Adam, & Martin Nystrand. 1994. "Tracking, instruction and achievement." *International Journal of Educational Research* 21(2): 217-231.

Gamoran, Adam, and Robert D. Mare. 1989. "Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality?." *American journal of Sociology* 1146-1183.

Hallinan, Maureen T. 1994. "Tracking: From theory to practice." *Sociology of Education* 67(2): 79-84.

Hallinan, Maureen T. 1996. Track mobility in secondary school. *Social Forces*, 74(3): 983-1002.

Hanselman, Paul, Thurston Domina, and NaYoung Hwang. 2016. "Organizational Inequality Regimes in School: Provision, Allocation, and Production of Educational Opportunities." Manuscript submitted for publication.

Hanushek, Eric A., and Ludger Woessmann. 2006 "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries." *Economic Journal* 116(510): C63-C76.

Hoffer, Tomas B. 1992. “Middle School Ability Grouping and Student Achievement in Science and Mathematics. “*Educational Evaluation and Policy Analysis* 14(3): 205-27

Horvath, Hedvig. 2015. “Classroom Assignment Policies and Implications for Teacher Value-Added Estimation” Unpublished manuscript.

Jackson, C. Kirabo, Rucker Johnson, and Claudia Persico. 2014. *The effect of school finance reforms on the distribution of spending, academic achievement, and adult outcomes* (No. w20118). National Bureau of Economic Research.

Jennings, Jennifer L., David Deming, Christopher Jencks, Maya Lopuch, and Beth E. Schueler. 2015. "Do Differences in School Quality Matter More Than We Thought? New Evidence on Educational Opportunity in the Twenty-first Century." *Sociology of Education* 88(1): 56-82.

Kalev, Alexandra, Frank Dobbin, and Erin Kelly. 2006. “Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies.” *American sociological review*, *71*(4): 589-617.

Kelly, Sean. 2007. “The contours of tracking in North Carolina.” *The High School Journal*, *90*(4): 15-31.

Kelly, Sean. 2009. “The black-white gap in mathematics course taking.” *Sociology of Education*, *82*(1): 47-69.

Kelly, Sean, and Heather Price. 2011. "The Correlates of Tracking Policy Opportunity Hoarding, Status Competition, or a Technical-Functional Explanation?." *American Educational Research Journal* 48(3):560-585.

Kelly, Sean, and William Carbonaro. 2012. “Curriculum tracking and teacher expectations: Evidence from discrepant course taking models.” *Social Psychology of Education*, *15*(3): 271-294.

Kerckhoff, Alan C. "Effects of ability grouping in British secondary schools. 1986." *American Sociological Review* :842-858.

Kerckhoff, Alan C. 1995. "Institutional arrangements and stratification processes in industrial societies." *Annual review of sociology* 21: 323-347.

Kliebard, Herbert M. "Why history of education?. 1995." *The Journal of Educational Research* 88(4): 194-199.

Legewie, Joscha, and Thomas A. DiPrete. 2012. “School context and the gender gap in educational achievement.” *American Sociological Review 77*(3): 463-485.

Long, Mark C., Dylan Conger, and Patrice Iatarola. 2012. "Effects of high school course-taking on secondary and postsecondary success." *American Educational Research Journal* 49(2): 285-322.

Loveless, Tom. 2013. The Algebra Imperative: Assessing Algebra in a National and International Context. Washington, DC: Brookings. Retrieved September 1, 2013 from http://www.brookings.edu/research/ papers/2013/09/04-algebra-imperative-education-loveless.

Lucas, Samuel R. 1999. *Tracking Inequality: Stratification and Mobility in American High Schools. Sociology of Education Series*. New York: Teachers College Press.

Lucas, Samuel R., and Mark Berends. 2002. "Sociodemographic diversity, correlated achievement, and de facto tracking." *Sociology of Education*, 328-348.

Lucas, Samuel. R., Aaron D. Good. 2001. "Race, class, and tournament track mobility." *Sociology of Education*, 139-156.

McFarland, Daniel. A. 2006. "Curricular flows: Trajectories, turning points, and assignment criteria in high school math careers." *Sociology of Education*,*79*(3): 177-205.

Metz, Mary Haywood .1978. *Classrooms and Corridors: The Crisis of Authority in Desegregated Secondary Schools.* Berkeley: University of California Press.

Meyer, John W., Brain Rowan, and M.W. Meyer. 1978. "The Structure of Educational Organization."

Mickelson, Roslyn Arlin. 2001. "Subverting Swann: First-and second-generation segregation in the Charlotte-Mecklenburg schools." *American Educational Research Journal*, *38*(2): 215-252.

Mickelson, Roslyn Arlin. 2003. "When are racial disparities in education the result of racial discrimination? A social science perspective." *The Teachers College Record*, *105*(6): 1052-1086.

Nomi, T., & Allensworth, E. M. 2012. Sorting and Supporting Why Double-Dose Algebra Led to Better Test Scores but More Course Failures. *American Educational Research Journal* 50(4): 756-788.

Nomi, T., & Raudenbush, S. W. (2016). Making a Success of "Algebra for All" The Impact of Extended Instructional Time and Classroom Peer Skill in Chicago. *Educational Evaluation and Policy Analysis* 38(2), 431-451.

Nystrand, Martin and Adam Gamoran. 1997. "The big picture: Language and learning in hundreds of English lessons." *Opening Dialogue*. New York: Teachers College, Columbia University 30-74.

Oakes, J. 1985. *Keeping track: How Schools Structure Inequalities*. Yale University Press.

Oakes, Jeannie, and Gretchen Guiton. 1995. "Matchmaking: The dynamics of high school tracking decisions." *American Educational Research Journal 32*(1): 3-33.

Page, Reba Neukom. 1991. *Lower-Track Classrooms: A Curricular and Cultural Perspective*. New York: Teachers College Press.

Penner, Andrew M., Thurston Domina, Emily K. Penner, and AnneMarie Conley. 2015. "Curricular policy as a collective effects problem: A distributional approach." *Social science research* 52: 627-641.

Raudenbush, Stephen, and Anthony S. Bryk. 1986. "A hierarchical model for studying school effects." *Sociology of education*: 1-17.

Raudenbush, Stephen W., and Robert D. Eschmann. 2015. "Does Schooling Increase or Reduce Social Inequality?." *Annual Review of Sociology* 41: 443-470.

Reardon, Sean F., and Ann Owens. 2014. 60 Years after Brown: Trends and consequences of school segregation. *Annual Review of Sociology* 40: 199-218.

Reed, D.S. 2014. *Building the Federal Schoolhouse: Localism and the American Education State.* Oxford: Oxford University Press.

Rickles, Jordan H. 2011. "Using interviews to understand the assignment mechanism in a nonexperimental study: The case of eighth grade Algebra." *Evaluation review* 35(5); 490-522.

Rosenbaum, James E. 1976. *Making Inequality; the Hidden Curriculum of High School Tracking.* New York: John Wiley & Sons.

Rosenbaum, James E. 1999. "If Tracking Is Bad, Is Detracking Better?" *American Educator 23*(4): 24-29.

Shavit, Yossi, and Hans-Peter Blossfeld. 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries. Social Inequality Series*. Westview Press, 5500 Central Avenue, Boulder, CO 80301-2847

Slavin, Robert E. 1988. "Synthesis of research on grouping in elementary and secondary schools." *Educational Leadership* 46(1): 67-77.

Slavin, Robert E. 1990. “Cooperative Learning: Theory, research, and practice.” Englewood Cliffs, NJ: Prentice Hall.

Sørensen, Aage Bøttger. 1970. Organizational differentiation of students and educational opportunity. *Sociology of Education*: 355-376.

Stainback, Kevin, Donald Tomaskovic-Devey, and Sheryl Skaggs. 2010. “Organizational approaches to inequality: Inertia, relative power, and environments.” *Annual Review of Sociology 36*(1): 225-247.

Stein, Mary Kay, Julia Heath Kaufman, Milan Sherman, and Amy F. Hillen. 2011. “Algebra: A challenge at the crossroads of policy and practice.” *Review of Educational Research* 81: 453–492.

Tomaskovic-Devey, Donald. 2014. “The relational generation of workplace inequalities.” *Social Currents 1*(1): 51-73.

Tyack, David B., and Larry Cuban. 1995. *Tinkering toward utopia*. Harvard University Press.

Van Houtte, Mieke. 2004. "Tracking effects on school achievement: A quantitative explanation in terms of the academic culture of school staff." *American Journal of Education* 110(4): 354-388.

Weick, Karl E. 1976. "Educational organizations as loosely coupled systems." *Administrative science quarterly*: 1-19.

Wells, Amy Stuart, and Jeannie Oakes. 1996. "Potential pitfalls of systemic reform: Early lessons from research on detracking." *Sociology of Education*: 135-143.

Zimmer, Ron. 2003. A new twist in the educational tracking debate. *Economics of Education Review 22*(3): 307-315.

Tables and Figures

Table 1: Descriptive statistics, 8th grade students in 3 Southern California public school districts, 2009-10—2011-12 school years

	District A	District B	District C
District administrative information			
Total 8th grade student enrollment, 2010-2012	12,212	7,913	3,714
N traditional schools enrolling 8th graders	9	10	4
N 8th grade mathematics classrooms[14]	116	103	41
N 8th grade ELA classrooms	165	80	35
Student demographics (averaged over available cohorts)			
% Female	50.6	50.7	47.1
% African American	2.5	0.5	0.9
% Asian	18.1	37.0	6.5
% Hispanic or Latino	67.1	51.4	44.9
% White	12.3	11.1	47.6
% Free- and Reduced-Price Lunch	70.7	69.9	50.4
% English Language Learners	20.3	28.0	16.9
% Reclassified Fluent English Speakers	43.8	47.1	22.1
% Special Education	6.2	2.4	12.5
7th grade ELA CST	-0.13 (0.97)	0.18 (0.93)	0.06 (1.15)
7th grade Mathematics CST	-0.16 (0.98)	0.22 (0.91)	0.08 (1.13)

[14] Classroom counts average over the 3 study years.

Figure 1: Distribution of standardized 7th grade achievement scores by eighth grade course track, math and ELA in 3 California school districts 2010-2012.

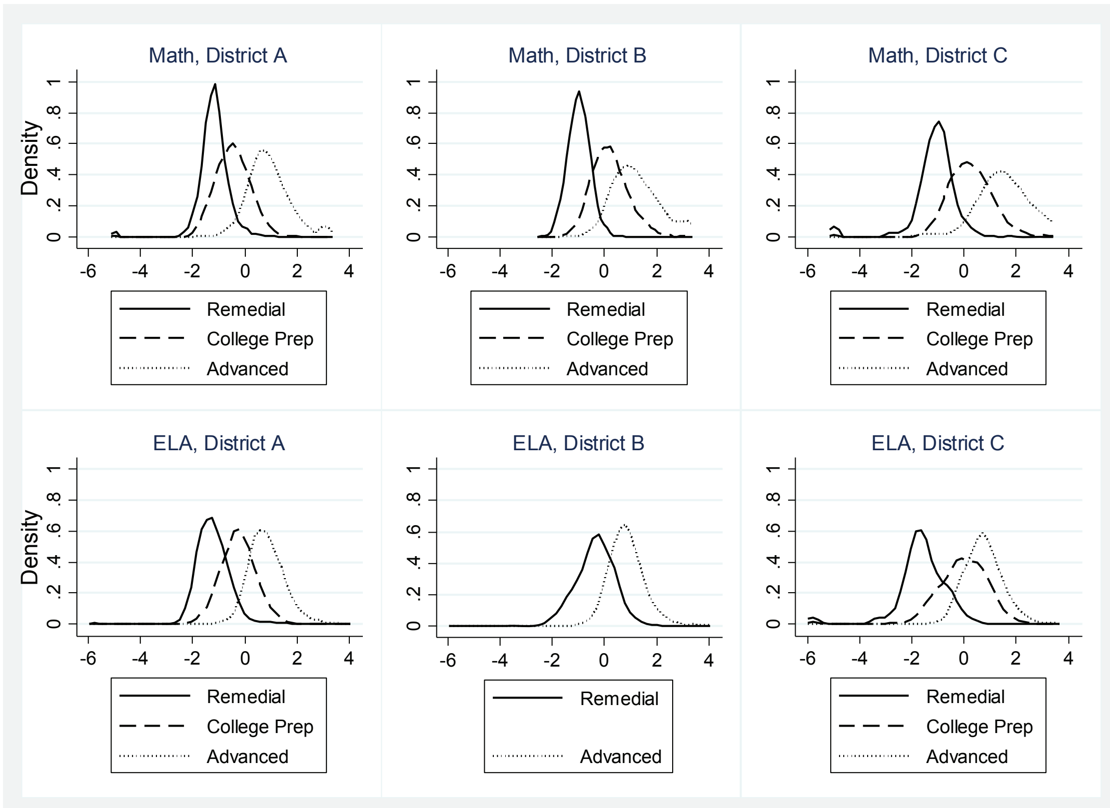


Table 2: Descriptive statistics for measures of dimensions of organizational differentiation in 3 Southern California public school districts, 2009-10—2011-12 school years

		Mean	SD	% variance between schools	% variance within schools (over time)
Differentiation					
# distinct courses					
	Math	4.06	1.06	67.7	32.3
	ELA	3.26	1.31	56.6	43.4
Homogeneous classroom assignments					
8^{th} grade classroom ICC, 7^{th} grade scores					
	Math	0.52	0.17	77.3	22.7
	ELA	0.50	0.18	52.0	48.0
Inclusiveness					
% College Prep or higher					
	Math	0.85	0.12	60.2	39.8
	ELA	0.63	0.27	49.8	50.2
Mobility					
% fall from college prep 8^{th}-9^{th}					
	Math	0.41	0.16	51.2	48.8
	ELA	0.34	0.27	49.2	50.8
Scope					
Correlation: Math to ELA track		0.67	0.16	53.3	46.7

Figure 2: Variation in homogeneous 8th grade mathematics classroom assignments: Distribution of 7th grade math CST scores by 8th grade classrooms in 1 school, 2010 and 2012

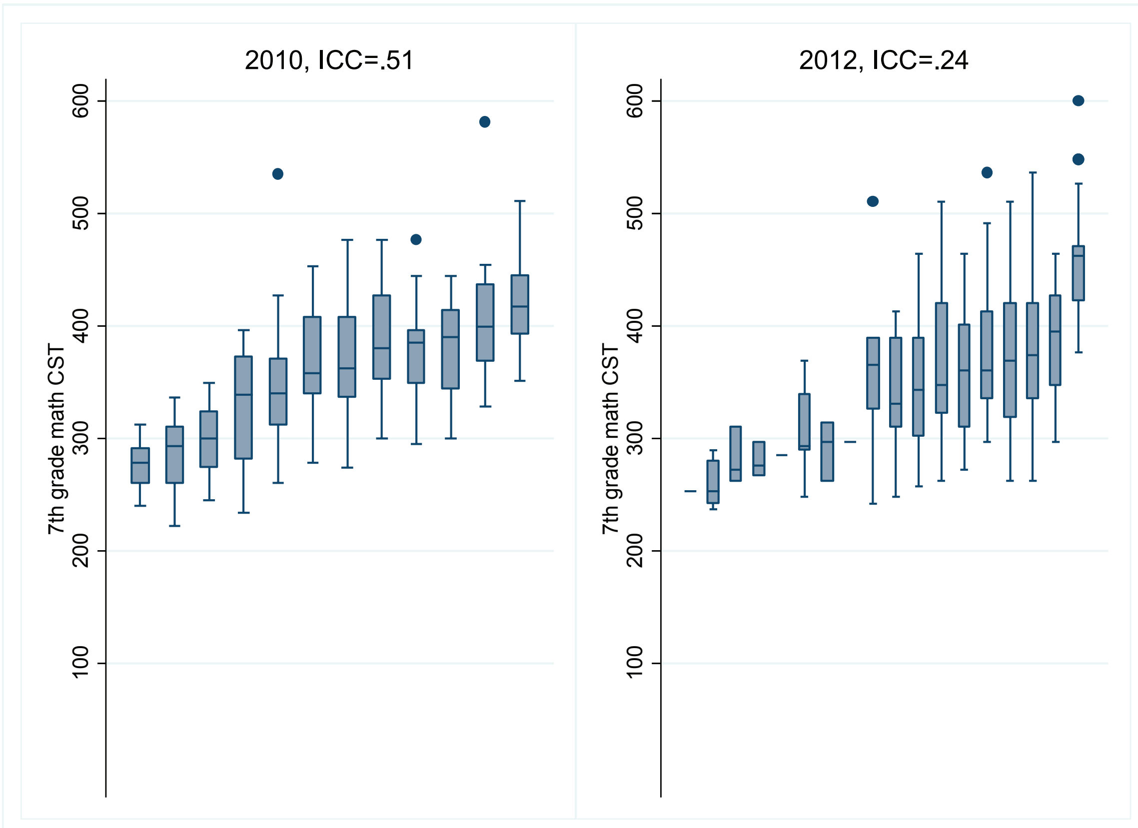


Table 3: Correlation of school-level measures of dimensions of organizational differentiation in 3 Southern California public school districts, 2009-10—2011-12 school years

	Differentiation (Math)	Differentiation (ELA)	Homogeneity (Math)	Homogeneity (ELA)	Inclusiveness (Math)	Inclusiveness (ELA)	Mobility (Math)	Mobility (ELA)	Scope
Differentiation (Math)	1.00								
Differentiation (ELA)	0.00	1.00							
Ability grouping (Math)	0.36	0.08	1.00						
Ability grouping (ELA)	0.20	0.44	0.30	1.00					
Inclusiveness (Math)	-0.42	0.30	-0.18	-0.23	1.00				
Inclusiveness (ELA)	-0.07	0.16	0.39	-0.09	0.23	1.00			
Mobility (Math)	-0.45	0.55	-0.17	0.02	0.71	0.35	1.00		
Mobility (ELA)	-0.04	0.31	0.40	-0.01	0.25	0.95	0.46	1.00	
Scope	0.03	0.05	-0.12	-0.20	0.56	-0.13	0.16	-0.17	1.00

Table 4: Multilevel model, school-level predictors of dimensions of organizational differentiation measures for all District A, B, and C middle schools 2010-2012 (School-year level data, with school-level random effects. Outcomes as well as % Female, % Disadv, x CST and, Enrollment are z-scored)

	Differentiation (Math)	Differentiation (ELA)	Homogeneity (Math)	Homogeneity (ELA)	Inclusiveness (Math)	Inclusiveness (ELA)	Mobility (Math)	Mobility (ELA)	Scope
% Female	-0.09	-0.05	-0.02	-0.05	0.1	0.01	0.03	-0.02	0
% Disadv	-0.26	0.32*	-0.31	0.14	-0.25	-0.22*	-0.1	-0.42***	-0.63*
$\bar{\bar{x}}$ CST	-0.33	-0.38*	-0.48	-0.61**	0.27	0.08	-0.67***	-0.50***	-0.23
Enrollment	0.34	0.18	0.26	0.06	0.49**	0	0.26*	0.16**	0.69***
SD CST (Math)	1.04	1.67	1.52	1.37	-1.91	0.49	-0.86	0.4	-0.38
SD CST (ELA)	0.3	-1.29	1.05	-0.99	-0.78	-0.01	-0.71	-0.37	0.73
2011	0.34	-0.16	-0.03	0.12	-0.26	0.06	-0.11	0.05	-0.27*
2012	0.71**	-0.32	0.21	0.23*	-0.74***	0.18***	-0.58***	0.14***	0.02
District B	0.81*	-0.37	0.19	0.58	-0.18	-1.69***	0.06	-0.97***	1.33*
District C	-0.15	-1.12***	0.09	-0.61	-0.01	0.34	-0.06	0.60***	-0.88*
Constant	-1.89	0.17	-2.47*	-0.6	2.81**	0.25	1.66*	0.24	-0.73
N=	69	69	69	69	69	69	68	66	69

Table 5: Selected coefficients, multilevel model, relationship between dimensions of 8th grade school tracking system and 10th grade mathematics achievement, for students in District A, B, and C middle schools 2010-2012

	Model 1 (Unconditional)	Model 2 (Controls)	Model 3 (Controls, school-mean centered)
Differentiation	0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.02)
Homogeneity	-0.05***	-0.03	-0.03*
	(0.01)	(0.01)	(0.01)
Inclusiveness	0.01	0.01	-0.08*
	(0.02)	(0.01)	(0.05)
Mobility	-0.06**	-0.03*	0.08***
	(0.02)	(0.02)	(0.03)
Scope (level)	0.02	0.01	0.00
	(0.01)	(0.01)	(0.01)
Demographic controls	No	Yes	Yes
Prior achievement	No	Yes	Yes
School-mean centered	No	No	Yes
District FE	No	Yes	Yes
Cohort FE	No	Yes	Yes
School RE	Yes	Yes	Yes
Classroom RE	Yes	Yes	Yes
N=	22,067	20,938	20,921

Table 6: Multilevel model, relationship between dimensions of 8th grade school tracking system and 10th grade ELA achievement, for students in District A, B, and C middle schools 2010-2012

	Model 1 (Unconditional)	Model 2 (Controls)	Model 3 (Controls, school-mean centered)
Differentiation	-0.03*	0.00	0.02
	(0.01)	(0.01)	(0.01)
Homogeneity	0.01	0.02	0.00
	(0.02)	(0.01)	(0.01)
Inclusiveness	0.03	0.10***	-0.00
	(0.05)	(0.02)	(0.05)
Mobility	-0.15**	-0.11**	0.03
	(0.06)	(0.03)	(0.03)
Scope (level)	0.02	0.01	0.02
	(0.02)	(0.01)	(0.01)
Demographic controls	No	Yes	Yes
Prior achievement	No	Yes	Yes
School-mean centered	No	No	Yes
District FE	No	Yes	Yes
Cohort FE	No	Yes	Yes
School RE	Yes	Yes	Yes
Classroom RE	Yes	Yes	Yes
N=	22,067	20,938	20,921

Figure 3: Predicted 10th grade mathematics achievement scores for students in schools with high and low levels of skills-homogeneous assignment in 8th grade mathematics classrooms

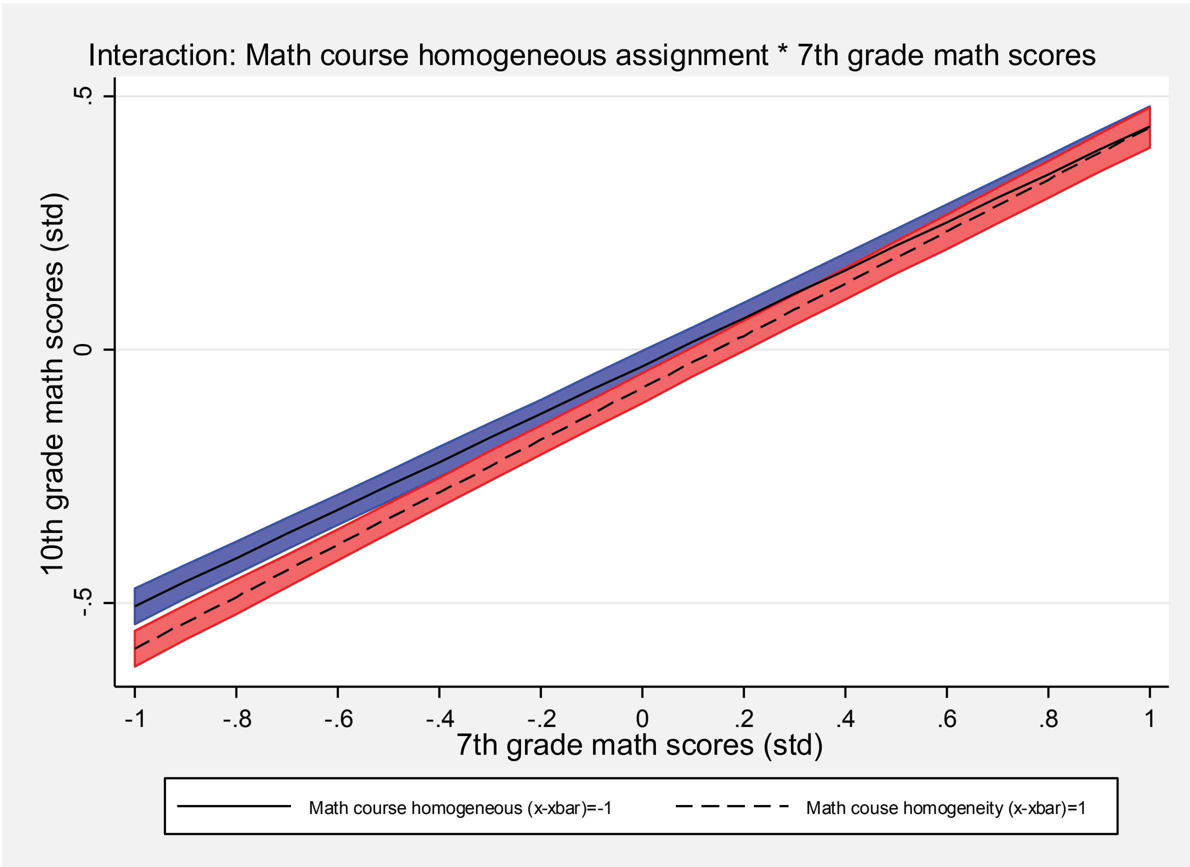
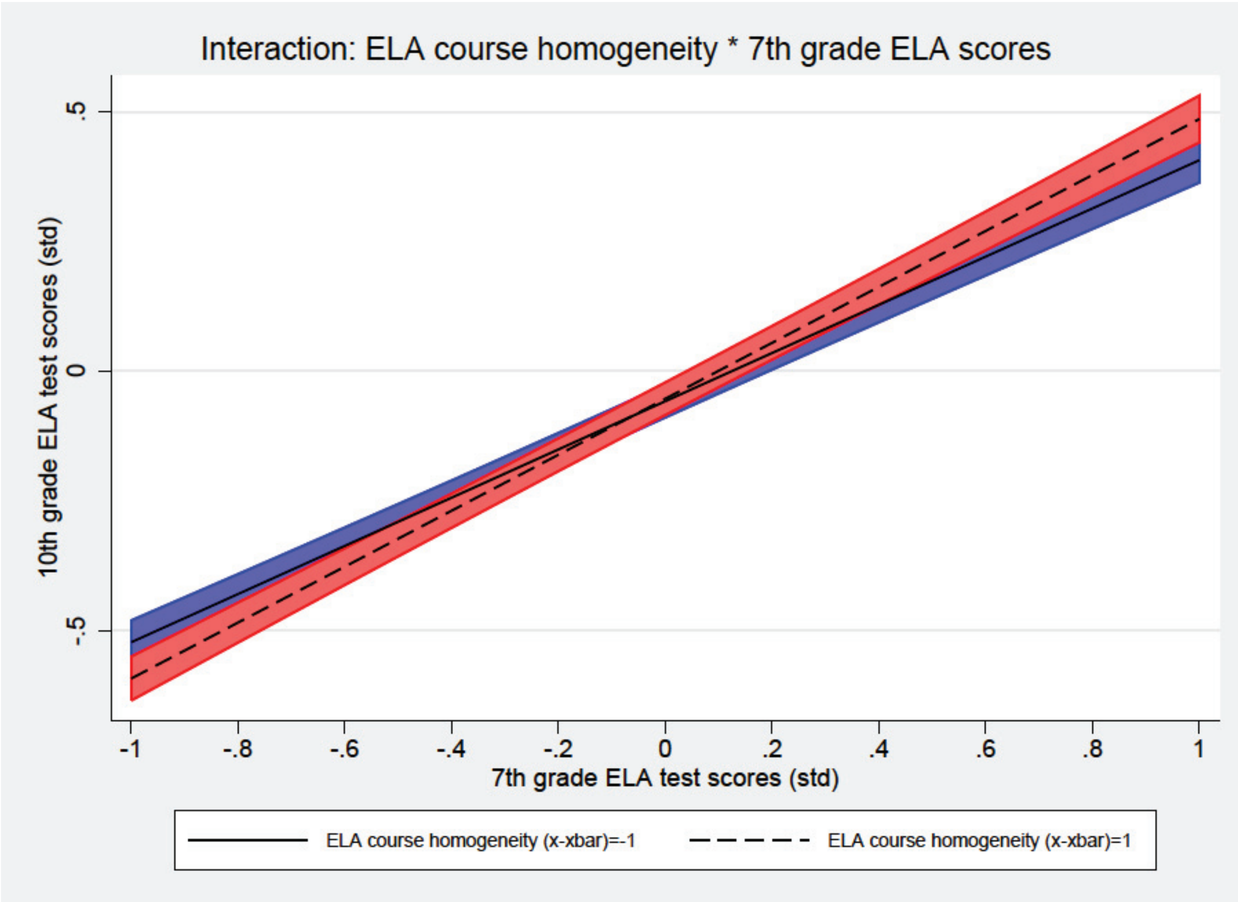


Figure 4: Predicted 10th grade ELA achievement scores for students in schools with high and low levels of skills-homogeneous assignment in 8th grade ELA classrooms



Appendices

Appendix Table 1: Full multilevel models, relationship between dimensions of 8th grade school tracking system and 10th grade mathematics achievement, for students in District A, B, and C middle schools 2010-2012

Math Model

	Model 1	Model 2	Model 3
Homogeneity	-0.05***	-0.03**	
	(0.01)	(0.01)	
Inclusiveness	0.01	0.01	
	(0.02)	(0.01)	
Mobility	-0.06**	-0.03*	
	(0.02)	(0.02)	
Differentiation	0.01	-0.01	
	(0.01)	(0.01)	
Scope (level)	0.02	0.01	
	(0.01)	(0.01)	
Female		-0.13***	-0.13***
		(0.01)	(0.01)
Black		-0.13***	-0.13***
		(0.03)	(0.03)
Asian		0.14***	0.14***
		(0.02)	(0.02)
Hispanic		-0.13***	-0.12***
		(0.02)	(0.02)
Free or reduced lunch		-0.03**	-0.03*

	(0.01)	(0.01)
English language learner	-0.09***	-0.09***
	(0.02)	(0.02)
Reclassified English Proficient	0.10***	0.10***
	(0.01)	(0.01)
Special Education	-0.20***	-0.20***
	(0.02)	(0.02)
CST (ELA)	0.16***	0.16***
	(0.01)	(0.01)
CST (math)	0.49***	0.49***
	(0.01)	(0.01)
Year 2011	-0.01	-0.01
	(0.01)	(0.01)
Year 2012	-0.03	-0.02
	(0.02)	(0.02)
District 2	-0.16***	-0.25***
	(0.03)	(0.05)
District 3	-0.06	-0.08
	(0.04)	(0.04)
Homogeneity (school mean)		-0.04
		(0.02)
Homogeneity (school mean centered)		-0.03*
		(0.01)
Inclusiveness (school mean)		-0.01
		(0.03)
Inclusiveness (school mean centered)		-0.08***
		(0.02)

Mobility (school mean)			-0.10***
			(0.02)
Mobility (school mean centered)			0.08***
			(0.03)
Differentiation (school mean)			-0.01
			(0.02)
Differentiation (school mean centered)			-0.01
			(0.01)
Scope (school mean)			0.06**
			(0.02)
Scope (school mean centered)			0.00
			(0.01)
Constant	-0.26***	0.15***	0.18***
	(0.03)	(0.02)	(0.03)

Appendix Table 2: Full multilevel models, relationship between dimensions of 8th grade school tracking system and 10th grade ELA achievement, for students in District A, B, and C middle schools 2010-2012

	Model 1	Model 2	Model 3
Homogeneity	0.02	0.01	
	(0.01)	(0.01)	
Inclusiveness	0.10***	0.02	
	(0.02)	(0.02)	
Mobility	-0.11**	-0.08*	
	(0.03)	(0.03)	
Differentiation	0.00	-0.01	
	(0.01)	(0.01)	
Scope	0.01	0.02**	
	(0.01)	(0.01)	
Female	0.08***	0.08***	0.08***
	(0.01)	(0.01)	(0.01)
Black	-0.03	-0.03	-0.03
	(0.04)	(0.04)	(0.04)
Asian	0.09***	0.09***	0.09***
	(0.02)	(0.02)	(0.02)
Hispanic	-0.08***	-0.08***	-0.07***
	(0.02)	(0.02)	(0.02)
Free or reduced lunch	-0.05***	-0.06***	-0.05***
	(0.01)	(0.01)	(0.01)
EL	-0.26***	-0.23***	-0.26***
	(0.02)	(0.02)	(0.02)
RFEP	0.05***	0.03**	0.05***

	(0.01)	(0.01)	(0.01)
Special Education	-0.29***	-0.29***	-0.29***
	(0.02)	(0.02)	(0.02)
CST (ELA)	0.50***	0.47***	0.50***
	(0.01)	(0.01)	(0.01)
CST (math)	0.15***	0.14***	0.15***
	(0.01)	(0.01)	(0.01)
Year 2011	0.05***	0.05***	0.06***
	(0.01)	(0.01)	(0.01)
Year 2012	-0.07***	-0.06***	-0.04*
	(0.01)	(0.01)	(0.02)
District 2	-0.12*	-0.17***	-0.19**
	(0.05)	(0.05)	(0.06)
District 3	-0.05	-0.05	-0.13**
	(0.03)	(0.03)	(0.04)
Advanced track (ELA)		0.29***	
		(0.02)	
College prep track (EPA)		0.11***	
		(0.02)	
Homogeneity (school mean)			0.06**
			(0.02)
Homogeneity (school mean centered)			0.00
			(0.01)
Inclusiveness (school mean)			0.09**
			(0.03)
Inclusiveness (school mean centered)			0.00
			(0.05)

Mobility (school mean)			-0.11*
			(0.05)
Mobility (school mean centered)			0.03
			(0.06)
Differentiation (school mean)			-0.08***
			(0.02)
Differentiation (school mean centered)			0.02
			(0.01)
Scope (school mean)			0.01
			(0.01)
Scope (school mean centered)			0.02
			(0.01)
Constant	0.08**	-0.02	0.10**
	(0.03)	(0.03)	(0.03)

Appendix Table 3: Multilevel models, relationship between dimensions of 8th grade school tracking systems and 10th grade achievement, for students in District A, B, and C middle schools 2010-2012 with interactions

	cammci	caemci
Homogeneity (school mean)	-0.04	0.06**
	(0.02)	(0.02)
Homogeneity (school mean centered)	-0.02*	0.00
	(0.01)	(0.01)
Inclusiveness (school mean)	-0.01	0.09**
	(0.03)	(0.03)
Inclusiveness (school mean centered)	-0.08***	0.01
	(0.02)	(0.05)
Mobility (school mean)	-0.10***	-0.11*
	(0.02)	(0.05)
Mobility (school mean centered)	0.09***	0.03
	(0.03)	(0.06)
Differentiation (school mean)	-0.01	-0.08***
	(0.02)	(0.02)
Differentiation (school mean centered)	-0.01	0.02*
	(0.01)	(0.01)
Scope (school mean)	0.07**	0.01
	(0.02)	(0.01)
Scope (school mean centered)	0.00	0.02
	(0.01)	(0.01)
Female	-0.13***	0.08***
	(0.01)	(0.01)
Black	-0.13***	-0.03
	(0.03)	(0.04)
Asian	0.14***	0.08***
	(0.02)	(0.02)
Hispanic	-0.12***	-0.07***
	(0.02)	(0.02)
Free or reduced lunch	-0.03*	-0.05***
	(0.01)	(0.01)
EL	-0.09***	-0.26***
	(0.02)	(0.02)
RFEP	0.10***	0.05***
	(0.01)	(0.01)
Special Education	-0.20***	-0.29***
	(0.02)	(0.02)

CST (ELA)	0.16***	0.50***
	(0.01)	(0.01)
CST (math)	0.49***	0.15***
	(0.01)	(0.01)
Year 2011	-0.01	0.07***
	(0.01)	(0.01)
Year 2012	-0.01	-0.04*
	(0.02)	(0.02)
District 2	-0.25***	-0.19**
	(0.05)	(0.06)
District 3	-0.08	-0.13**
	(0.04)	(0.04)
Homogeneity (school mean centered)*CST	0.02*	0.04**
	(0.01)	(0.01)
Constant	0.18***	0.09**
	(0.03)	(0.03)