American Educational Research Journal
June 2023, Vol. 60, No. 3, pp. 447–485
DOI: 10.3102/00028312221134769
Article reuse guidelines: sagepub.com/journals-permissions

Uneven Progress: Recent Trends in Academic Performance Among U.S. School Districts

Kaylee T. Matheny
Stanford University
Marissa E. Thompson
Columbia University
Carrie Townley-Flores
sean f. reardon
Stanford University

We use data from the Stanford Education Data Archive to describe district-level trends in average academic achievement between 2009 and 2019. Although on average school districts' test scores improved very modestly (by about 0.001 standard deviations per year), there is significant variation among districts. Moreover, we find that average test score disparities between nonpoor and poor students and between White and Black students are growing; those between White and Hispanic students are shrinking. We find no evidence of achievement-equity synergies or trade-offs: Improvements in overall achievement are uncorrelated with trends in achievement disparities. Finally, we find that the strongest predictors of achievement disparity trends are the

Kaylee T. Matheny is a PhD candidate in sociology of education and education policy at Stanford University; email: *kmatheny@stanford.edu*. Her research draws on mixed methods to understand socioeconomic inequality in educational opportunities.

Marissa E. Thompson is an assistant professor of sociology at Columbia University; email: *marissa.thompson@columbia.edu*. Her research focuses on the causes and consequences of racial and socioeconomic inequality in educational access.

Carrie Townley-Flores is a PhD candidate in education policy at Stanford University; email: *ctflores@stanford.edu*. Her research uses quantitative methods to examine how innovative social policies affecting young children can reduce racial, ethnic, and socioeconomic inequality.

sean f. reardon is the professor of poverty and inequality in education and professor (by courtesy) of sociology at Stanford University; email: *sean.reardon@stanford.edu*. His research focuses on the causes, patterns, trends, and consequences of social and educational inequality; the effects of educational policy on educational and social inequality; and applied statistical methods for education research.

levels and trends in within-district racial and socioeconomic segregation and changes in differential access to certified teachers.

Keywords: achievement, achievement gaps, equity, hierarchical linear modeling

Average standardized test scores in the United States have risen since the 1970s, particularly among elementary and middle school students (National Center for Education Statistics [NCES], 2013; Shakeel & Peterson, 2022). The average fourth grader in 2019, for example, had a mathematics test score that was nearly a standard deviation (SD) higher than those of fourth graders in her parents' generation in 1996 (NCES, 2020). These increases reflect improved educational opportunities: They imply that children growing up today have, on average, more resources and opportunities—in their homes, neighborhoods, preschools, and elementary and middle schools—to learn the math and reading skills measured by standardized tests than did children 50 years ago. In addition, they reflect changing curricula and a greater emphasis on academic skills in earlier grades than previous cohorts experienced (Bassok et al., 2016).

The last few decades' increases in average test scores are evident for all racial/ethnic groups, although the increases have been larger among Black and Hispanic students than among White students. As a result, the national White-Black and White-Hispanic achievement disparities have narrowed substantially in the last 50 years as well (Page et al., 2008; Reardon et al., 2014). This suggests that increases in overall educational opportunities have been accompanied, at least over the long term, by growing racial equity in educational opportunities. In contrast, the achievement disparity between nonpoor and poor students (as measured by free/reduced-price lunch eligibility) has been relatively stable for several decades (NCES, 2020), while the disparity between affluent and very poor students (those at the 90th and 10th percentiles of the income distribution, respectively) has widened substantially (Reardon, 2011; but see also Hashim et al., 2020; Reardon, 2021). Equitable access to educational opportunities across socioeconomic backgrounds, then, does not seem to have improved along with overall educational outcomes.

These stylized facts provide insight into the national trends in achievement. However, we do not know how much these trends in achievement and achievement disparities vary among school districts. The U.S. educational system is highly decentralized, with nearly 13,000 traditional public school districts, each of which has considerable control over staffing, curriculum, instruction, and budget allocation decisions (Roza, 2010). This variation in local policies and practices suggests trends in test scores and test score disparities may vary substantially among school districts, but researchers have not yet had sufficient data to explore this possibility. Further, because districts are critical policymakers and resource distributors (Blazar & Schueler, 2022), it is useful to know the extent to which changes in district-level inputs shape trends in achievement. Knowing more about patterns of local achievement trends may (a) provide insight into the mechanisms facilitating these trends, (b) help us understand whether and under what conditions increasing achievement can be accompanied by narrowing between-group disparities, and (c) suggest potential levers for improving educational opportunities and increasing achievement.

In this paper, we provide a detailed descriptive account of trends in achievement patterns among U.S. school districts, extending the recent work of Atteberry et al. (2021). We first measure recent trends in students' academic performance in almost every public school district in the United States, using 11 years of data (2009–2019) on math and reading test scores for Grades 3–8. We then describe how these trends vary among school districts and how they differ between economic and racial/ethnic student subgroups within school districts. Third, we examine whether increasing overall performance and reducing achievement disparities are synergistic or conflicting processes. Finally, we estimate the extent to which local demographic changes and changes in school characteristics predict trends in average performance and economic and racial/ethnic achievement disparities. This exploratory analysis helps identify potential factors to examine further as levers for increasing educational opportunity and student achievement.

We find that in the average school district, test scores changed very little from 2009 to 2019. However, test score changes vary considerably among districts. Our analyses show that this variation has not been driven by demographic changes in the local population. Rather, we find that districts with more experienced teachers, a lower proportion of frequently absent teachers, and higher average achievement at the start of the study period experienced the greatest improvement.

We also find that district-level trends in academic performance vary significantly by subgroup. In particular, socioeconomic disparities and White-Black disparities in the average district have widened, and White-Hispanic disparities have narrowed, although there is considerable variation in these trends among districts. Notably, we find that trends in overall achievement and trends in disparities are uncorrelated, indicating that there is neither synergy nor trade-off between performance and equity over time.

Finally, the strongest predictors of increasing achievement disparities are measures of inequality: Achievement disparities have grown most rapidly, on average, in school districts with increasing levels of school segregation. Our findings, in conjunction with other research (see Johnson, 2019), suggest that reducing within-district segregation may lead to more equitable educational opportunities.

Background

Improving academic performance has been the focus of major federal education policy over the past 20 years. Both the No Child Left Behind Act and its successor, the Every Student Succeeds Act, focus on improving standardized test scores and, importantly, narrowing achievement disparities. Ironically, however, the standardized testing regime driven by these policies has not been implemented in a way that produces comparable measures of academic performance at the school or district level over time, making it difficult to measure progress in academic performance. Not only do the tests and proficiency standards vary among states; they also vary within states over time as tests, proficiency standards, and reporting practices change. Changes in the tests or proficiency levels over time break the measurement of trends in performance and disparities (Bandeira de Mello et al., 2019; Ho, 2008; Yee & Ho, 2015). As a result, after almost 20 years of standardized testing in every public school in the United States, we lack clear descriptive information on how average test scores and test score disparities have changed in individual schools and districts.

The logic of national policy efforts—holding schools and districts accountable for raising achievement of all economic and racial/ethnic groups to a common level—positions raising achievement and narrowing achievement disparities as compatible, even synergistic goals (Every Student Succeeds Act, 2015; No Child Left Behind Act, 2002). If the goal of improving all students' test scores incentivizes schools and districts to allocate resources primarily to under-resourced (and lower-performing) students, then test-based accountability systems may increase average test scores and decrease achievement disparities, making improving achievement and equity synergistic goals. Conversely, it may instead be the case that school districts and policies most effective at raising achievement may not be those most effective at reducing test score disparities. An accountability system based on test scores may have disproportionately negative effects on schools serving under-resourced student groups, which may lead to widening achievement disparities even if overall district achievement were to improve. There may be little synergy or even some trade-off between allocating resources to increase average achievement and allocating resources to support historically disadvantaged students' achievement. However, we know little about trends in district-level academic performance and these potential synergies or trade-offs.

Prior Research on Test Score Trends Overall and by Race/Ethnicity and Income

Despite the lack of information on trends in test scores at the district level, we have evidence from the National Assessment of Educational Progress (NAEP) that math and reading scores for Grades 4 and 8 at the national level increased substantially from the 1970s until about 2013, at a rate of roughly 0.05–0.20 SD/decade (Shakeel & Peterson, 2022). Since 2013, however, math and reading scores have declined modestly in most grades and subjects (Hussar et al., 2020; NCES, 2013, 2018a).

NAEP also provides insight into how trends in achievement vary among states. The average state's NAEP scores improved by about 0.12 SD/decade in math and by 0.08 SD/decade in reading from 2003 to 2019, but the trends range widely. The 10th and 90th percentiles of the math trend distributions were, respectively, 0.00 and 0.22 SD/decade in math and –0.01 and 0.18 in reading (Hussar et al., 2020; NCES, 2018a, 2018b).

The national White-Black and White-Hispanic achievement disparities have also narrowed substantially over the last few decades. The disparities declined at a rate of roughly 0.10 SD/decade during the 1990s and early 2000s (NCES, 2020; Page et al., 2008). Progress in reducing White-Black disparities has reversed in the last decade, but the White-Hispanic disparity continued to narrow through 2019.[1] Nonetheless, the rate at which these disparities have narrowed is very small relative to their size, and the disparities remain extremely large.[2]

The trend in the disparity between higher- and lower-income students is somewhat less clear. NAEP data show that the test score disparity between students eligible and not eligible for free lunch—an imperfect measure of family income—has remained roughly stable in recent decades, although there have been modest to sizeable increases since the 1980s in the disparity between students with college-educated parents and those whose parents have not attended college (NCES, 2020; Reardon et al., 2014). Likewise, Reardon (2011) finds that the disparity between affluent and very poor students grew substantially for cohorts of students born from the 1970s through the 1990s, although Reardon and Portilla (2016) suggest this trend may have reversed for more recent cohorts, at least with respect to the disparity at kindergarten entry. In contrast, several new working papers challenge Reardon's (2011) finding (see Hanushek et al., 2020; Hashim et al., 2020), although Reardon (2021) identifies several methodological shortcomings in these papers.[3]

The NAEP data provide evidence on the national- and state-level trends in average academic performance and achievement disparities. The substantial heterogeneity among state-level trends suggests local forces shape educational opportunity patterns. However, until recently, we have had no reliable information about trends in achievement or achievement disparities within more local units, such as school districts.[4] One recent study, however, offers an exception. Atteberry et al. (2021) use 2009–2016 data from the Stanford Education Data Archive (SEDA) to describe recent trends in racial/ethnic achievement disparities. They find substantial variation among and within districts, with average White-Black disparities widening slightly and White-Hispanic disparities narrowing slightly.

Our analyses build upon this work in several ways. First, we analyze not only trends in racial/ethnic disparities but also trends in overall achievement and in economic disparities, investigating potential synergy (or trade-offs) between achievement and equity. We also identify correlates of these trends, which allows us to describe which kinds of districts have experienced the largest improvements. Further, we use updated SEDA data (spanning 11 years, 2009–2019). In the methods and discussion sections of this paper, we discuss in more detail how our study aligns with and extends the work of Atteberry et al. (2021).

Predictors of Local Test Score Trends

One way to conceptualize test scores is to think of them as indicators of cumulative educational opportunities. In this vein, the average test scores of students in a given school district and year reflect the total set of educational opportunities and resources the students have had from birth through the time they take the test. These opportunities include experiences and resources in their homes, neighborhoods, preschools, peer groups, and schools. A change in average scores from one cohort of students to another within a district then reflects changes in educational opportunities and resources. Broadly put, changes in average scores may result from differences between cohorts in out-of-school experiences—such as changes in family resources or differences in preschool or neighborhood conditions—or from differences between cohorts in school characteristics, practices, or resources. Further, if changes in a district lead to changes in average scores, then systematic differences in experiences of those changes between student groups will predict changes in disparities between those groups.

Experiences outside of school affect students' academic performance. This is evident from the substantial variation in cognitive skills even when children are just starting kindergarten (Garcia, 2015; Lee & Burkam, 2002). In fact, even at school entry, one of the biggest predictors of academic performance is socioeconomic status (SES) (Garcia, 2015). This relationship persists across grades and across cohorts (Hanushek et al., 2020; Hashim et al., 2020; Reardon, 2011; Sirin, 2005). Therefore, changes in average family socioeconomic status across cohorts may lead to corresponding changes in average test scores. Moreover, racial/ethnic achievement disparities are expected to widen (narrow) when racial/ethnic family socioeconomic status disparities widen (narrow; Berends et al., 2008; Garcia, 2015).

School conditions, too, influence student achievement. Teachers, in particular, can have large effects on student outcomes: Teacher quality is consistently considered the most important school-based factor shaping student achievement (McCaffrey et al., 2004; Rivkin, 2000; Wright, 2005). Although there is considerable debate regarding what constitutes "teacher quality," some key factors that shape student achievement include teacher experience, attendance, and certification status (Clotfelter et al., 2009, 2010; Goldhaber & Brewer, 2000; Miller et al., 2008; Podolsky et al., 2016; Roby, 2013).

Other classroom, school, and district characteristics affect achievement, as well. For example, class sizes affect the amount of teacher time, focus, and attention available to the average student. Studies show that shrinking class sizes is associated with increasing achievement, particularly for low-income students (Angrist & Lavy, 1999; Krueger, 1999; Rivkin et al., 2005). Additionally, given the rapid increase in charter school enrollment over the last decade as well as the improving public opinion toward charter schools, it is important to understand the relationship between growing charter enrollment and changing achievement (Betts & Tang, 2011). On the one hand, if charter schools spur competition that leads traditional public schools to improve (Gilraine et al., 2019), a growing share of students enrolled in charter schools may increase district achievement; on the other hand, if the local charter options are no more effective than traditional public schools in the same district (Abdulkadirğlu et al., 2011), they may neither inspire competition nor spur achievement gains for other local schools. Overall, the evidence on these processes is mixed (Cohodes & Parham, 2021). Finally, a number of rigorous studies in the last decade have found clear evidence that increases in school funding lead to improved academic achievement and other educational outcomes (for a detailed review, see Jackson, 2020).

To the extent that there are differences in school conditions across student groups, then, school conditions will also influence achievement disparities. Indeed, in some places, racial/ethnic school segregation is as high today as it was in the 1960s (Orfield & Frankenberg, 2014), and segregation has hardly changed over the last 25 years (Johnson, 2019; Reardon & Owens, 2014). For the 100 districts enrolling the most Asian, Black, and Hispanic students, however, school segregation has substantially increased (Owens et al., 2022). Racial/ethnic segregation is one of the strongest predictors of racial/ethnic achievement disparities, primarily due to the disproportionate concentration of Black and Hispanic students in high-poverty schools (Card & Rothstein, 2007; Gamoran & An, 2016; Owens, 2018; Reardon et al., 2014; Reardon, Kalogrides, et al., 2019; Reardon, Weathers, et al., 2019). Further, although racial/ethnic school segregation may not have changed much, socioeconomic school segregation has grown substantially since the 1980s (Duncan & Murnane, 2011; Owens et al., 2016). This has marked consequences for students, as socioeconomic segregation between schools is strongly associated with racial/ethnic and socioeconomic disparities in test scores (Owens, 2018; Reardon, 2016). Because school segregation may concentrate students differentially in schools with higher or lower levels of educational resources and opportunities (Bischoff & Owens, 2019), districts with growing between-school segregation may widen achievement disparities more than similar districts with stable or declining between-school segregation (Reardon, Weathers, et al., 2019). Finally, expanding charter school options may also lead to widening disparities, as charter schools are associated with increased school segregation (Rich et al., 2021).

One of the primary mechanisms through which school segregation shapes student opportunities is through how schools with higher concentrations of historically disadvantaged students also tend to have fewer and lower-quality resources. For example, Black and Hispanic students are more likely to attend schools with lower-quality teachers and larger class sizes than are White students (Boozer et al., 1995; Clotfelter et al., 2005; Kalogrides et al., 2013). Changes in the distribution of these resources, then, may shape relative achievement patterns. Further, Sosina and Weathers (2019) show that growing between-district racial segregation within a state is associated with decreasing Black/White per-pupil funding ratios. Similar disparities exist between poor and nonpoor students (Clotfelter et al., 2007; Jacob et al., 2016; Kalogrides et al., 2013; Morgan & Amerikaner, 2018). This suggests that as segregation patterns change, changes in the distribution of financial, teacher, and other resources may lead to changes in achievement disparities.

Present Study

Our first goal in this paper is to provide a descriptive overview of trends in academic achievement and achievement disparities at the school district level, using population-level data from almost every school district in the country.

Second, we investigate whether the current policy realm includes a trade-off between achievement and equity. Here, we examine whether improvements in overall achievement are generally shared by all student groups in a district and how often increasing achievement coincides with narrowing achievement disparities. The answer here helps shed light on whether the current policy regime facilitates a virtuous synergy or a vicious choice between equity and achievement.

Third, we investigate how district demographic changes, community characteristics, and school district characteristics shape these trends. These analyses do not provide clear causal evidence but do provide useful descriptive evidence and identify associations that may generate hypotheses for future research. Moreover, the answers to these questions have important policy implications. Knowing the conditions under which average test scores and disparities in test scores generally improve may help states and districts determine how to allocate resources and direct school improvement efforts.

Data and Methods

We use data from several sources. Like Atteberry et al. (2021), we use data from SEDA to measure district-level trends in average test scores, although we use a newer version of the SEDA data that includes 11 years of information. We use additional data from the Common Core of Data (CCD), the Civil Rights Data Collection (CRDC), and the American Community Survey (ACS) to construct covariates used in our models.

District-Level Trends

Our unit of analysis is the school district. Although it might be useful to examine trends in academic performance and achievement disparities at both the district and school levels, we focus in this paper on districts as the unit of analysis for three reasons. First, and most importantly, districts play a key role in the distribution of federal and state school revenue (Roza, 2010). District-level decisions determine, in large part, how financial and other resources (such as experienced teachers) are distributed among schools, potentially contributing to within-district, between-school inequality in opportunity (e.g., Atteberry et al., 2021; Sosina & Weathers, 2019). District-level decisions and policies affect many types of resources that shape student performance, including local funding initiatives, class sizes, staffing decisions, curricular foci, and course offerings (Blazar & Schueler, 2022; Roza, 2010). That is not to say that school-level practices and decisions play no role, but the context in which school leaders operate is shaped in part by district-level actions. As a result, we view school districts as a consequential unit of interest in examining achievement and disparity trends.

Second, districts are more stable units of analysis than are schools. School attendance zone boundaries often change. Even when they do not, student assignment policies (such as school choice policies) and changes in school programs (such as the availability of gifted/talented programs) may alter the mix of students attending a given school. Such changes may confound the measurement of school-level trends in average academic achievement. Of course, the mix of students in a district also may change over time due to local demographic changes or changes in private school enrollment patterns. But between-district student moves and transfers are much rarer than between-school, within-district student transfers (Reardon, Papay, et al., 2019). Moreover, demographic changes are more easily and more comprehensively measured at the district than the school level: The U.S. census provides socioeconomic data for families with children tabulated at the school district level, but not the school level; at the school level, the only available measure of socioeconomic composition is the proportion of students eligible for free and reduced-price lunch, which has grown less reliable over time (Fahle et al., 2021; Greenberg et al., 2019). Finally, although SEDA (our source for test score data) includes estimates of overall trends in test scores at the school level, it does not include data on school-level achievement disparities or trends in disparities. For these reasons, we focus here on measuring and describing district-level, rather than school-level, trends in academic achievement and achievement disparities.

SEDA

SEDA is based on approximately 430 million Grades 3–8 math and reading test scores in the 90,150 U.S. public schools that enrolled students in at least one of Grades 3–8 in at least one year from the spring of 2009 to the spring of 2019. In SEDA, these schools are aggregated into 12,849 "geographic school districts," each of which contains all traditional public schools, charter schools, and Bureau of Indian Education schools within the geographic boundaries of a traditional school district. In other words, charter schools and Bureau of Indian Education schools are treated as being part of the local traditional public district in which they are physically located. Virtual schools, which enroll fewer than half of 1% of all students, are excluded from SEDA (Fahle et al., 2021).

SEDA includes estimates of average performance in each district by grade, cohort, and subject. It also includes estimates of average scores disaggregated by economic status and race/ethnicity as well as estimates of economic and racial/ethnic test score disparities. The test scores are linked to a common scale across states, grades, and years, making comparisons possible across districts in different states and across grades and years.[5] The scores are standardized within grades and subjects relative to the corresponding national student-level test score distribution for a reference cohort. That is, within each grade-subject, the same national means and SDs are used to standardize the estimated scores across all years, so changes in standardized scores across years reflect absolute changes, in national SD units, rather than relative changes.[6]

Estimating Achievement Trends by Using Precision-Weighted Random Effects Models

In each district, SEDA includes up to 132 estimates of average test scores, one for each grade-year-subject cell in the data (6 grades, 11 years, 2 subjects). In some cases, data for a given cell are not included in SEDA. This can occur for one of several reasons, most commonly because the state did not administer tests in a particular year (many states did not administer tests in 2014 as they transitioned to new Common Core–aligned assessments) or because fewer than 95% of enrolled students in the cell were tested. On average, each district includes estimated average test scores in 111 of 132 possible cells. In addition, SEDA includes up to 132 estimates of nonpoor-poor, White-Black, and White-Hispanic test score disparities in each district. More information on the underlying data can be found in the SEDA technical documentation (Fahle et al., 2021).

We pool the (up to) 132 observations within districts and use a precision-weighted multilevel model (described below) to estimate each district's average within-grade and -subject trend in test scores over the period of 2009–2019. We estimate the trend in average scores and the trend in economic and racial/ethnic achievement disparities in each district. In supplemental analyses, we disaggregate trends by subject, given the stark domain differences between math and reading. However, we privilege the pooled estimates for several reasons. First, the correlates we consider, including changes in student-teacher ratios, access to high-quality teachers, and segregation, affect students' broad educational opportunities. We do not include domain-specific correlates, such as classroom time dedicated to reading or class sizes in mathematics. For this reason, we do not expect the correlates in our study to have different relationships with math and reading/language arts. Indeed, districts with high average scores in one subject have high average scores in the other, and the opposite is also true; further, places with unequal resources across groups see similar White-Black or White-Hispanic disparities in both subjects, with little variation across grades within districts (Reardon, Kalogrides, et al., 2019). Methodologically, pooling across subjects yields more reliable estimates of average district trends (the grade-by-subject-specific trends can be very noisy, with substantially larger standard errors). We consider pooled-subject achievement to be a measure of average general academic performance in a district. This approach differs from that used by Atteberry et al. (2021), who estimate trends separately in each district-grade-subject. Although their approach allows for a detailed description of trends across grades and subjects, it does not examine the correlates of trends, which is a primary aim of this paper.

The SEDA-provided estimated average test score and its standard error for students in district d for year y, grade g, and subject b are designated by $\hat{\mu}_{dygb}$ and $\hat{\omega}_{dygb}$, respectively. For models where the outcome is a subgroup or disparity trend, $\hat{\mu}_{dygb}$ refers to a subgroup-specific average test score or a disparity estimate in the relevant models. We define *cohort* = *year* − *grade*, so that the *cohort* variable indicates the spring of the year that a cohort of students was in kindergarten, following Reardon (2019). Because the data span the period of 2009–2019 and Grades 3–8, there are 16 cohorts represented in the data, ranging from those entering kindergarten from 2001 to 2016. The test subject is indicated by the binary variable *math* (1 = math; 0 = reading). In all models, we center the *grade*, *cohort*, and *math* variables at their midpoints in the data (i.e., at 5.5, 2008.5, and 0.5, respectively).

Because the data include multiple observations (across grades, years, and subjects) nested within districts, we estimate trends in achievement by using the following precision-weighted multilevel model and data from districts in the full sample:

Within-district model:

$$\hat{\mu}_{dgyb} = \beta_{0d} + \beta_{1d}(grade_{dgyb}) + \beta_{2d}(cohort_{dgyb}) + \beta_{3d}(math_{dgyb}) + e_{dgyb} + \varepsilon_{dgyb}$$

Between-district model:

$$\beta_{0d} = \gamma_{00} + u_{0d}$$

$$\beta_{1d} = \gamma_{10} + u_{1d}$$

$$\beta_{2d} = \gamma_{20} + u_{2d}$$

$$\beta_{3d} = \gamma_{30} + u_{3d}$$

$$\varepsilon_{dgyb} \sim N\left(0, \hat{\omega}^2_{dgyb}\right); e_{dgyb} \sim N\left(0, \sigma^2_{dgyb}\right);\ [u_{0d},\ u_{1d},\ u_{2d},\ u_{3d}] \sim MVN(0, \tau). \quad (1)$$

The level-one (within-district) model includes two error terms, one (e_{dygb}) representing the residual of the true mean achievement, conditional on *grade*, *cohort*, and *math*; and the other (ε_{dgyb}) representing the SEDA-provided estimation error in $\hat{\mu}_{dygb}$. The estimation error variance $\hat{\omega}^2_{dygb}$ is treated as a known parameter—the error variance of the estimated mean score $\hat{\mu}_{dygb}$—while σ^2_{dgyb} and τ are estimated. We fit the model via maximum likelihood, using the HLM v7 software.

In this model, β_{2d} is the parameter of interest: the pooled average within-grade and -subject (cohort-to-cohort) change in average test scores in SDs in district d. A hypothetical trend of 0.01 in a district would imply that, on average, scores in that district increased by 0.01 SD per year over the period from 2008–2009 to 2018–2019, or 0.1 SD over the study period. In addition, the matrix τ is informative: Its diagonal element τ_{22} indicates the variance of trends across districts. Other estimates obtained from this model are the average district-level test scores, β_{0d}; the estimated within-cohort growth from Grades 3–8, β_{1d};[7] and the estimated difference in math and reading scores (within grade and cohort) for the district, β_{3d}.

We estimate β_{2d} for all districts with sufficient data to produce an estimate by using Model 1, above. In some cases, data are insufficient to produce an estimate; in other cases, an estimate is available but is insufficiently reliable to be useful. Following the procedures used to construct the SEDA data, we exclude from our analyses cases where (a) the estimate is based on test scores of fewer than 20 unique students (or 20 unique students of both groups in the case of disparity estimates) or (b) the standard error of $\hat{\beta}_{2d}$ is larger than 0.015.[8] Almost all districts meet these criteria for the overall trend estimates. We refer to the set of all districts with sufficient data to reliably estimate a pooled trend as the "full sample" (12,194 districts, 95% of all districts, enrolling 99.8% of all public school students not in virtual schools); we refer to the set of all districts with estimated trends and non-missing covariate data (see below) as the "analytic sample" (12,087 districts, enrolling 99.6% of all students). This restriction excludes fewer than 1% of districts from each analysis, primarily due to excluding districts with missing community SES data or missing per-pupil instructional expenditures. Note that the sample sizes are smaller for the subgroup and disparity trend estimates because many districts do not meet the student sample size criteria for one or more student groups and because there are more cases where the estimates do not meet the precision criterion.

It is useful to consider how the estimated trends depend on the underlying NAEP and state test data. Because the SEDA data are linked using NAEP data, the estimated trend in a given district depends on its statewide trend (derived from NAEP data) and the district's within-state deviation from its statewide trend (derived from state test data). As a result, the average trend among all school districts (weighted by average annual enrollment) will necessarily be roughly the same as the overall national trend in performance. (The equivalence is not exact because SEDA does not have test scores for every district and because district enrollments change over time.) Therefore, the weighted average trend across districts is, as it should be, a reflection of the average national trend in NAEP scores. The variation in weighted average trends among states reflects variation in NAEP trends across states. However, the variation in trends among districts within states reflects variation in trends in performance on state tests. That is, the scaling of SEDA using NAEP data determines the national- and state-level average trends that we see in SEDA estimates but does not determine the within-state variation in trends.

Covariates

We use data from a variety of sources to identify demographic, community, and school characteristics. The bulk of these covariates are drawn from the CCD, the CRDC, and the ACS. From the CCD, we obtain district-level economically disadvantaged and racial/ethnic student composition; estimates of within-district, between-school segregation;[9] school resource variables, including pupil-teacher ratios and instructional expenditures; and enrollment patterns, including overall district enrollment and charter school enrollment. From the CRDC, we obtain district-level measures of teacher experience and quality, including the proportion of teachers in their first 2 years of teaching, the proportion of teachers absent from school more than 10 days in a given school year, and the proportion of certified teachers. From the ACS, we include measures of district-level SES following the approach used by Reardon (2019). Additionally, we follow the approach of Shores and Steinberg (2019) to construct a measure of intensity of the impacts of the 2008 recession by using data from the Bureau of Labor Statistics. (See Online Supplement Table 1 for descriptive statistics of the covariates.)

Table 1

Descriptive Statistics of District-Level Achievement Trends, 2009–2019

	Full sample				Analytic sample				
Student subgroup	Trend mean	Trend SD	Trend reliability	Obs.	Trend mean	Trend SD	% Variance within states	Obs.	Trend analysis incorporating student subgroup
All students	0.001	0.023	0.914	12,194	0.001	0.023	85.2%	12,087	Overall, see Table 3
Poor students	0.002	0.024	0.883	11,351	0.001	0.023	76.8%	10,507	Nonpoor-Poor, see Table 4
Nonpoor students	0.006	0.024	0.881	11,200	0.006	0.024	82.5%	10,507	Nonpoor-Poor, see Table 4
Black students	–0.001	0.025	0.766	4,516	–0.001	0.025	84.1%	4,375	White-Black, see Table 5
Hispanic students	0.006	0.025	0.790	6,711	0.006	0.025	81.8%	6,449	White-Hispanic, see Table 6
White students	0.002	0.023	0.892	11,667	0.002	0.020	82.9%	4,375	White-Black, see Table 5
					0.002	0.021	81.8%	6,449	White-Hispanic, see Table 6
Test score disparity	**Trend mean**	**Trend SD**	**Trend reliability**	**Obs.**	**Trend mean**	**Trend SD**	**% Variance within states**	**Obs.**	**Trend-trend correlation**
Nonpoor-Poor	0.005	0.016	0.742	10,578	0.005	0.016	84.6%	10,507	0.794
White-Black	0.003	0.018	0.741	4,391	0.003	0.018	90.9%	4,375	0.752
White-Hispanic	–0.005	0.018	0.721	6,478	–0.005	0.018	90.6%	6,449	0.768

Note. *Trend mean* refers to average annual changes in achievement in the period of 2009–2019. *Trend-trend correlation* indicates the correlation between the two subgroup trends (e.g., the correlation between the White and Black trends). The analytic samples are smaller than the full sample because they only include districts with enough students from each subgroup to estimate trends or disparities in trends as well as full information regarding each of the covariates in the analysis. All student subgroups are included in the analytic sample of overall trends (see Table 3). The subgroups of poor and nonpoor students appear in the analytic samples of nonpoor-poor gap trends (see Table 4). The subgroups of Black and White students appear in the analytic samples of White-Black gap trends (see Table 5). The subgroups of Hispanic and White students appear in the analytic samples of White-Hispanic gap trends (see Table 6). SD = standard deviation.

Modeling Trends in Achievement and Achievement Disparities

We model the trends in achievement and achievement disparities in each district (the estimated $\hat{\beta}_{2d}$'s from Equation 1) as a function of district characteristics. We include models with and without state-level fixed effects. The models have the following form:

$$\hat{\beta}_{2d} = \Lambda + \mathrm{Z}_{\boldsymbol{d}}\Gamma + v_d + \varphi_\delta$$

$$v_d \sim \mathrm{N}(0, \tau_d);\ \ \varphi_\delta \sim \mathrm{N}(0,\ \theta_d), \tag{2}$$

where $\hat{\beta}_{2d}$ is the ordinary least squares (OLS) estimate of the trend in district d, estimated from Model 1 above, and $\boldsymbol{Z}_d$ is a vector of district-level covariates. The models with state fixed effects have the same form (with the addition of state fixed effects). For each district's characteristics, we include the average and the change in the measure over the period (for example, we include a measure of average segregation and the change in segregation in the district from 2008–2009 to 2018–2019). The parameters of interest are the coefficients in the vector Γ, which describe the partial associations between district achievement (or disparity) trends and district characteristics. Because the precision of $\hat{\beta}_{2d}$ varies substantially among districts (as a result of varying district size and variation in the number of district-grade-year-subject cells available), we fit Equation 2 by using precision-weighting, which downweights low-precision estimates relative to high-precision estimates. Thus, the models include two error terms, one (v_d) representing the residual of the true trend in achievement, conditional on $\boldsymbol{Z}_d$, and the other (φ_δ) representing the estimation error in $\hat{\beta}_{2d}$ (i.e., $\hat{\beta}_{2d} = \beta_{2d} + \varphi_\delta$). The error variance of φ_δ is assumed to be known (it is set to equal the squared standard error of $\hat{\beta}_{2d}$), and the true conditional variance τ_d^2 is estimated (note that τ_d is smaller than τ_{22} from Model 1 above, as τ_d is the residual variance of β_{2d}, net of variance explained by the covariates $\boldsymbol{Z}_d$).

Results

Variation in Trends Among School Districts

We find that over the period of 2009–2019, test scores improved annually by 0.001 SD in cohort-scale units in the typical school district. However, trends vary considerably among districts: The SD of the annual trend is 0.023, meaning that over the study period, roughly one-sixth of U.S. school districts saw improvements of more than 0.24 SD unit, and one-sixth had declines in average scores of more than 0.22 SD unit.[10] Table 1 shows average annual changes in achievement for the period of 2009–2019, disaggregated by demographic group.

There is also considerable variation in trends in group-specific academic performance and in achievement disparities (note the SDs of trends in

Table 1). Among the roughly 11,000 districts with enough poor and nonpoor students to report disparities, nonpoor students saw average achievement gains on the order of 0.006 SD annually. Average gains for poor students, however, were not as high, with average district-level increases of 0.002 SD annually. This means that between 2009 and 2019, the average within-district test score disparity between poor and nonpoor students has grown by approximately 0.05 SD unit (or about 11%).[11] This was largely driven by the high average achievement gains of nonpoor students.

For White and Black students, the patterns are slightly different. From 2009 to 2019, average district-level test scores for White students improved by 0.002 SD per year. Among the roughly 4,400 districts with enough Black and White students to report disparities in test scores, the average district's Black students' scores declined by 0.001 SD per year. This results in a disparity that has increased by about 0.003 SD per year. In the average district, then, the White-Black disparity has grown by about .03 SD (about 6%) over the last decade.

Although economic and White-Black disparities have widened, on average, the White-Hispanic disparity has narrowed. In districts with enough White and Hispanic students to report disparities, the test scores of White students improved by 0.002 SD annually, while those of Hispanic students grew by an average of 0.006 SD annually. This resulted in a disparity trend of –0.005 SD/year, or about –.05 SD (about 13%) over the study period.

The vast majority of the variance in trends is explained by differences among districts within states. Only 10%–20% of differences in trends in overall scores and score disparities are attributable to differences between states, and 80%–90% are attributable to differences within states.

There are clear trends in the relative extent of educational opportunity over time. Although test scores are improving for poor students, the average within-district nonpoor-poor disparity continues to grow. In contrast, although all other subgroups have improved over time, scores for Black students have declined slightly. This has led to an increasing disparity between White and Black students in the average district. Finally, the average within-district disparity between White and Hispanic students has decreased over the last decade.

Variation Across the 20 Largest Districts

To illustrate the extent of the variation in trends, we show the variation in overall trends across the 20 school districts with the highest enrollment in the United States (which enroll approximately 11% of all students represented in our analytic sample) in Figure 1. The trends in achievement and disparities vary widely across school districts. Average test scores in Philadelphia, Pennsylvania, declined by approximately 0.024 SD/year, while scores in San Diego, California, improved by 0.026 SD/year.

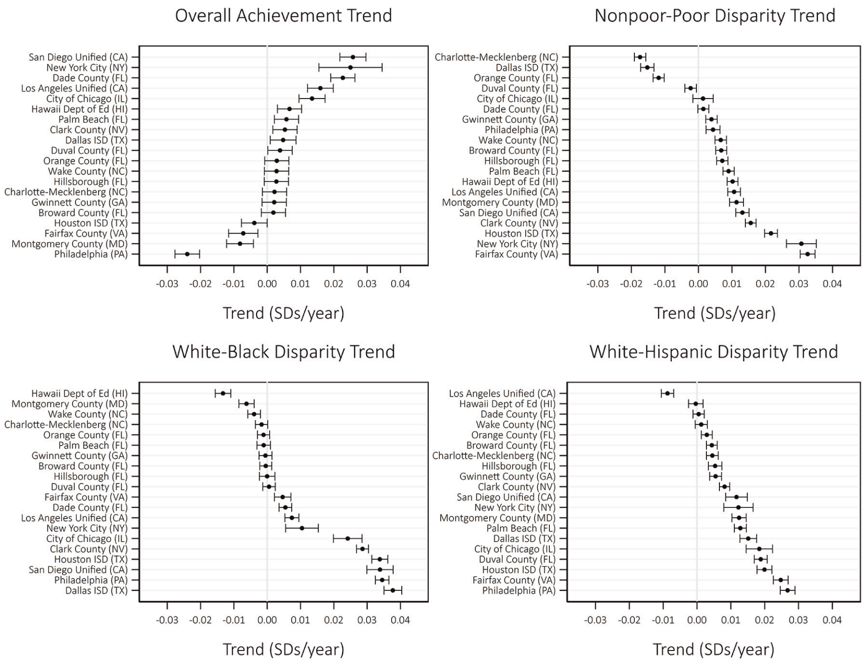

Figure 1 **Achievement trends and trends in disparities for the 20 largest districts by enrollment**

Note. Trends measure the average change in performance per year in standard deviations, standardized to student performance in 2009. Each plot shows the 20 largest geographic school districts by enrollment.

Achievement disparity trends also vary widely among the 20 largest districts. The district where economic disparities have narrowed the fastest is Charlotte-Mecklenberg, North Carolina, where the nonpoor-poor disparity has declined at a rate of 0.017 SD/year. Conversely, inequality is growing rapidly in Fairfax County, Virginia, where the economic disparity has grown almost twice as fast as the disparity in Charlotte-Mecklenberg has shrunk. In some districts, improving achievement is accompanied by narrowing achievement disparities: In Los Angeles Unified School District, in California, the White-Hispanic disparity has narrowed modestly, while overall achievement has improved substantially. In other places, the opposite is true: In Philadelphia, Pennsylvania, not only have average scores declined substantially; the White-Black and White-Hispanic disparities have widened significantly at the same time. The data indicate that even among large school

Table 2
Correlations Between Overall Achievement Trend and Trends in Achievement Disparities

	1	2	3	4
1. Overall achievement trend	—			
2. Nonpoor-poor disparity trend	0.032	—		
3. White-Black disparity trend	–0.013	0.509	—	
4. White-Hispanic disparity trend	–0.011	0.492	0.506	—

districts, the patterns of change in achievement and achievement disparities vary widely.

Achievement-Equity Trade-Off?

Incentives for improving overall educational opportunities may be at odds with incentives for equitably improving educational opportunities, or providing more resources to students with greater need. To identify patterns in the extent to which achievement gains and changes in achievement disparities are related, we estimate correlations between subgroup trends and between overall and disparity trends. If subgroup trends are positively correlated, this suggests changes in educational opportunities are generally shared across student groups; if they are negatively correlated, this suggests changes that increase opportunity for one group may decrease opportunity for another. Here, we find that subgroup trends are strongly and positively correlated (see trend-trend correlations in Table 1). That is, in districts where the structurally advantaged group (e.g., White students) experiences increasing educational opportunity, so does the structurally disadvantaged group (e.g., Black students), on average. Further, disparity trends are themselves moderately correlated, such that when the White-Black disparity increases, so, too, do the nonpoor-poor and White-Hispanic disparities (see Table 2). In part, these correlations are mechanical because both the White-Black and White-Hispanic trends depend on the trend for White students.

In general, if there were synergy between growing achievement and narrowing achievement disparities, we would expect to see negative correlations between trends in achievement and trends in disparities; if there were trade-offs between growing achievement and narrowing achievement disparities, we would expect to see positive correlations between trends in achievement and trends in disparities. We find almost no correlation between overall achievement trends and trends in nonpoor-poor, White-Black, and White-Hispanic achievement disparities (with correlations of 0.032, –0.013, and –0.011, respectively; see Table 2). This suggests there has been no systematic synergy between achievement gains and equity gains over the last decade.

Figure 2 displays the relationships between trends in the performance of different student groups. Each of the three graphs shows information for a different pair of student groups (nonpoor-poor, White-Black, and White-Hispanic). Each graph plots the trend in one group's achievement against the trend in the other group's achievement (e.g., comparing the trends for White students vs. the trends for Black students). Each district is represented by a bubble. There is wide variation in group achievement trends (evident from the range of district trends along the axes) and in disparity trends (evident from the dispersion of the data around the 45-degree line). In districts on the 45-degree line, disparities are unchanging; districts above or below the line have disparities that are becoming more or less favorable, respectively, to poor, Black, or Hispanic students. The scatterplot is divided into six regions: In districts in regions A and B, both student groups' scores improved over time; in regions D and E, both groups' scores declined; in C and F, one group's scores improved, while the other group's declined. In virtually all cases, disparities narrowed for districts above the 45-degree line (in regions A, F, and E) and widened for those below the line (regions B, C, and D).[12]

To the extent that there is synergy between improving achievement and narrowing achievement disparities, we would expect to see most districts fall in regions A (improving achievement accompanied by narrowing disparities) or D (declining achievement accompanied by widening disparities). To the extent there is a trade-off, we would expect more districts in regions B and E. The figure suggests that in general, achievement gains or losses are common to all student groups in a district (there is a high correlation between group trends, evident in clustering along the 45-degree line in the scatterplot and as reported in Table 1). However, there is no correlation between trends in achievement and trends in achievement disparities: Districts where achievement is growing for both groups are generally as likely to fall in region A as in B; and districts where achievement is declining for both groups are as likely to fall in D as in E.

Changes in Demographic, Neighborhood, and District Characteristics

Table 3 reports estimates from the regression models predicting overall trends. We display models with different covariate sets and with and without state fixed effects, although our preferred model includes the full slate of covariates and state fixed effects (Model 6). We find that demographic changes have only a very small impact on the overall trends: Models 1 and 2 of Table 3 show that changes in districts' demographic compositions explain only 4.7% of the variance in trends (6.5% of the within-state variance). Thus, although districts' *average* achievement is strongly associated with economic and demographic characteristics, very little of the different trends between districts is attributable to changing demographic characteristics.

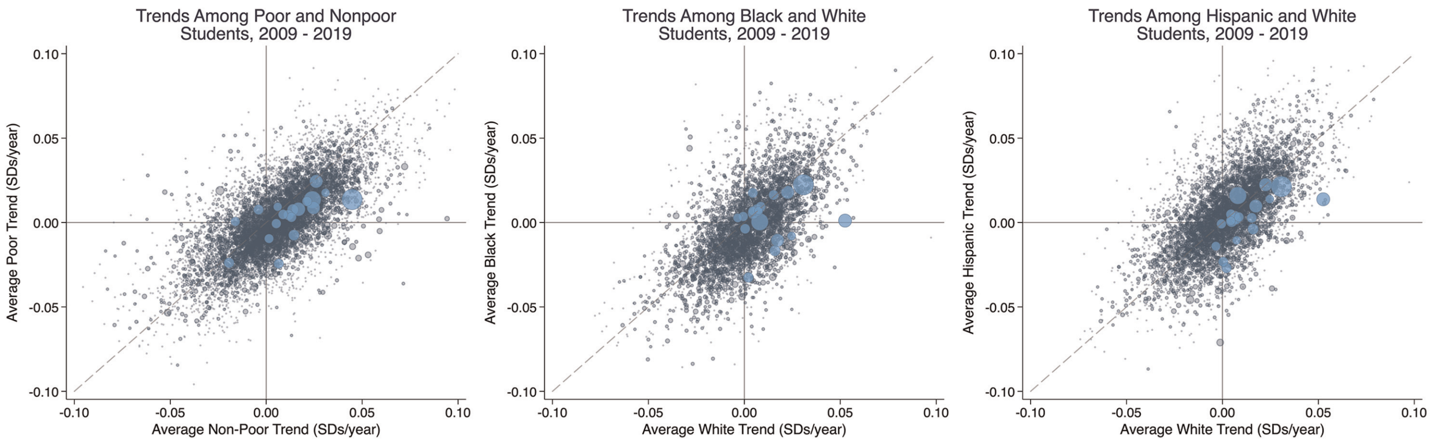

Figure 2 **District-level trends, by student subgroup**

Note. Light blue circles represent the 20 largest school districts. Dashed lines represent 45-degree lines. Trends measure the average change in performance per year in SD units. Outlying districts where trends fall above 0.10 SD/year or below –.10 SD/year were omitted from figures. SD = standard deviation.

Table 3

Coefficient Estimates From Models Predicting Overall Achievement Trends

	Demographic changes		+ Resource changes		+ Levels	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Change in exposure to novice teachers (average student)			−0.067***	−0.062***	−0.073***	−0.040**
			(0.010)	(0.012)	(0.011)	(0.012)
Change in exposure to absent teachers (average student)			−0.020***	−0.019**	−0.024***	−0.017*
			(0.006)	(0.007)	(0.006)	(0.006)
Change in % teaching staff with certification (average student)			0.007	0.027*	0.010	0.016
			(0.013)	(0.013)	(0.015)	(0.014)
Change in pupil-teacher ratio in average student's school			−0.000	0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)
Change in % charter school enrollment			0.059	−0.027	0.093*	0.044
			(0.033)	(0.036)	(0.038)	(0.045)
Change in per-pupil expenditures (ln)			0.060***	0.058	0.055***	0.027
			(0.010)	(0.033)	(0.011)	(0.019)
Change in enrollment grades 3–8 (ln)			0.017	0.018	−0.033***	−0.021
			(0.010)	(0.014)	(0.010)	(0.012)
Average performance					0.023***	0.031***
					(0.001)	(0.004)
Constant	0.001***	0.000	0.001***	0.000	0.001***	0.000
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Covariate set: Changes in demographics	Yes	Yes	Yes	Yes	Yes	Yes
Covariate set: Changes in resources	No	No	Yes	Yes	Yes	Yes
Covariate set: Means of demographics and resources	No	No	No	No	Yes	Yes
State fixed effects	No	Yes	No	Yes	No	Yes
Adj. R-squared	0.047	0.065	0.055	0.065	0.145	0.174
Observations	12087	12087	12087	12087	12087	12087

Note. This table presents descriptive statistics showing the relationship between different predictors and the average annual change in achievement in standard deviations. For models with state fixed effects, the within-state R-squared is reported. All covariates are mean-centered.
* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Table 4
Coefficient Estimates From Models Predicting Nonpoor-Poor Trends

	Demographic changes		+ Resource changes		+ Levels	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Change in % poor in average poor student's school			−0.405***	–0.291	–0.342**	–0.199
			(0.106)	(0.195)	(0.110)	(0.181)
Change in the difference in poor student exposure (poor-nonpoor)			0.394***	0.379**	0.363***	0.330**
			(0.048)	(0.119)	(0.051)	(0.111)
Change in exposure to novice teachers (poor)			0.029**	0.018	0.023*	0.016
			(0.009)	(0.009)	(0.010)	(0.009)
Change in the difference in exposure to novice teachers (poor-nonpoor)			0.040	0.041	0.025	0.015
			(0.037)	(0.059)	(0.038)	(0.059)
Change in exposure to absent teachers (poor)			−0.001	–0.001	–0.003	–0.001
			(0.005)	(0.005)	(0.005)	(0.005)
Change in the difference in exposure to absent teachers (poor-nonpoor)			0.014	0.009	0.040	0.036
			(0.031)	(0.040)	(0.033)	(0.046)
Change in exposure to certified teachers (poor)			0.032**	0.029*	0.035**	0.025
			(0.011)	(0.011)	(0.013)	(0.013)
Change in the difference in exposure to certified teachers (nonpoor-poor)			0.044	0.034	0.077*	0.063*
			(0.029)	(0.027)	(0.031)	(0.026)
Change in pupil-teacher ratio in average poor student's school			−0.000	0.000	–0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)
Change in the difference in the pupil-teacher ratio (poor-nonpoor)			0.000	0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)
Change in % poor enrolled in charter schools			−0.116***	–0.117**	–0.156***	–0.148***
			(0.028)	(0.039)	(0.033)	(0.042)
Change in the difference % enrolled in charter schools (nonpoor-poor)			−0.045	–0.019	–0.067	–0.042
			(0.031)	(0.073)	(0.035)	(0.073)
Change in per-pupil expenditures (ln)			0.003	0.011	–0.014	0.001
			(0.010)	(0.014)	(0.011)	(0.018)
Change in enrollment grades 3–8 (ln)			0.052***	0.044*	0.039***	0.027
			(0.009)	(0.017)	(0.009)	(0.015)

(continued)

Table 4 **(continued)**

	Demographic changes		+ Resource changes		+ Levels	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Average nonpoor-poor gap					0.004**	0.007*
					(0.001)	(0.003)
Constant	0.005***	0.004***	0.005***	0.004***	0.005***	0.004***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Covariate set: Changes in demographics	Yes	Yes	Yes	Yes	Yes	Yes
Covariate set: Changes in resources	No	No	Yes	Yes	Yes	Yes
Covariate set: Means of demographics and resources	No	No	No	No	Yes	Yes
State fixed effects	No	Yes	No	Yes	No	Yes
Adj. *R*-squared	0.011	0.000	0.031	0.000	0.046	0.045
Observations	10507	10507	10507	10507	10507	10507

Note. This table presents descriptive statistics showing the relationship between different predictors and the average annual change in the nonpoor-poor disparity in standard deviations. For models with state fixed effects, the within-state *R*-squared is reported. All covariates are mean-centered.
*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Table 5

Coefficient Estimates From Models Predicting White-Black Trends

	Demographic changes		+ Resource changes		+ Levels	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Change in % poor in average Black student's school			0.087**	–0.002	0.099***	0.003
			(0.028)	(0.043)	(0.029)	(0.043)
Change in the difference in poor student exposure (Black-White)			0.254**	0.265*	0.301***	0.334**
			(0.078)	(0.121)	(0.078)	(0.107)
Change in Black students' exposure to novice teachers			0.015	0.001	–0.004	–0.017
			(0.018)	(0.023)	(0.019)	(0.024)
Change in the difference in exposure to novice teachers (Black-White)			0.199**	0.153*	0.095	0.045
			(0.065)	(0.070)	(0.066)	(0.067)
Change in Black students' exposure to absent teachers			0.007	0.000	0.003	–0.003
			(0.009)	(0.008)	(0.009)	(0.007)
Change in the difference in exposure to absent teachers (Black-White)			0.078	0.067	0.120*	0.097
			(0.048)	(0.053)	(0.050)	(0.054)
Change in Black students' exposure to certified teachers			0.017	0.029	0.001	0.013
			(0.019)	(0.019)	(0.021)	(0.021)
Change in the difference in exposure to certified teachers (White-Black)			0.119*	0.099*	0.108*	0.090*
			(0.050)	(0.047)	(0.051)	(0.044)
Change in pupil-teacher ratio in average Black student's school			0.000	0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)
Change in the difference in the pupil-teacher ratio (Black-White)			0.001	0.001	–0.001	–0.001
			(0.000)	(0.000)	(0.001)	(0.001)
Change in % Black enrolled in charter schools			0.089	0.101*	–0.025	0.001
			(0.046)	(0.048)	(0.053)	(0.048)
Change in the difference % enrolled in charter schools (White-Black)			0.212***	0.185***	0.087	0.078
			(0.050)	(0.044)	(0.056)	(0.063)
Change in per-pupil expenditures (ln)			−0.029	0.005	–0.027	0.019
			(0.019)	(0.032)	(0.022)	(0.029)
Change in enrollment grades 3−8 (ln)			0.070***	0.036	0.047*	0.015
			(0.019)	(0.025)	(0.020)	(0.022)

(continued)

Table 5 **(continued)**

	Demographic changes		+ Resource changes		+ Levels	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Average White-Black gap					0.014***	0.017***
					(0.002)	(0.003)
Constant	0.003***	0.004***	0.003***	0.004***	0.003***	0.003**
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Covariate set: Changes in demographics	Yes	Yes	Yes	Yes	Yes	Yes
Covariate set: Changes in resources	No	No	Yes	Yes	Yes	Yes
Covariate set: Means of demographics and resources	No	No	No	No	Yes	Yes
State fixed effects	No	Yes	No	Yes	No	Yes
Adj. R-squared	0.011	0.000	0.036	0.033	0.107	0.100
Observations	4375	4375	4375	4375	4375	4375

Note. This table presents descriptive statistics showing the relationship between different predictors and the average annual change in the White-Black disparity in standard deviations. For models with state fixed effects, the within-state R-squared is reported. All covariates are mean-centered.
$^*p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Table 6

Coefficient Estimates From Models Predicting White-Hispanic Trends

	Demographic changes		+ Resource changes		+ Levels	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Change in % poor in average Hispanic student's school			−0.029	–0.047	–0.005	–0.021
			(0.024)	(0.031)	(0.024)	(0.026)
Change in the difference in poor student exposure (Hispanic-White)			0.395***	0.395***	0.512***	0.513***
			(0.080)	(0.076)	(0.080)	(0.076)
Change in Hispanic students' exposure to novice teachers			−0.007	–0.005	–0.009	–0.015
			(0.014)	(0.013)	(0.015)	(0.014)
Change in the difference in exposure to novice teachers (Hispanic-White)			0.085	0.042	0.049	0.008
			(0.065)	(0.058)	(0.066)	(0.063)
Change in Hispanic students' exposure to absent teachers			0.000	–0.004	–0.001	–0.005
			(0.007)	(0.006)	(0.007)	(0.007)
Change in the difference in exposure to absent teachers (Hispanic-White)			0.031	0.013	0.051	0.044
			(0.051)	(0.064)	(0.054)	(0.051)
Change in Hispanic students' exposure to certified teachers			0.011	0.008	0.001	0.003
			(0.017)	(0.017)	(0.019)	(0.018)
Change in the difference in exposure to certified teachers (White-Hispanic)			0.181***	0.154**	0.146**	0.138**
			(0.048)	(0.045)	(0.049)	(0.047)
Change in pupil-teacher ratio in average Hispanic student's school			0.000	0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)
Change in the difference in the pupil-teacher ratio (Hispanic-White)			0.000	0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)	(0.000)
Change in % Hispanic enrolled in charter schools			−0.046	–0.022	–0.066	–0.057
			(0.036)	(0.061)	(0.043)	(0.041)
Change in the difference % enrolled in charter schools (White-Hispanic)			0.100*	0.089	0.065	0.062
			(0.051)	(0.054)	(0.055)	(0.051)
Change in per-pupil expenditures (ln)			−0.003	0.008	–0.037*	–0.01
			(0.015)	(0.023)	(0.016)	(0.019)

(continued)

Table 6 **(continued)**

	Demographic changes		+ Resource changes		+ Levels	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Change in enrollment grades 3−8 (ln)			0.063***	0.065**	0.058***	0.044**
			(0.013)	(0.022)	(0.014)	(0.014)
Average White-Hispanic gap					–0.001	0.000
					(0.002)	(0.002)
Constant	−0.004***	−0.002*	−0.004***	–0.002*	–0.005***	–0.002*
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Covariate set: Changes in demographics	Yes	Yes	Yes	Yes	Yes	Yes
Covariate set: Changes in resources	No	No	Yes	Yes	Yes	Yes
Covariate set: Means of demographics and resources	No	No	No	No	Yes	Yes
State fixed effects	No	Yes	No	Yes	No	Yes
Adj. R-squared	0.011	0.035	0.026	0.035	0.077	0.069
Observations	6449	6449	6449	6449	6449	6449

Note. This table presents descriptive statistics showing the relationship between different predictors and the average annual change in the White-Hispanic disparity in standard deviations. For models with state fixed effects, the within-state R-squared is reported. All covariates are mean-centered.

$^*p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Models 3 and 4 incorporate resource changes within districts, including changes to each district's size, staffing, expenditures, and charter enrollments. We find that, net of other factors in the models (including state fixed effects), districts with increasing student exposure to novice or frequently absent teachers tend to experience declines in test scores; districts with increasing proportions of certified teachers tend to have rising scores. However, most changes in district policies and characteristics offer little predictive power in understanding patterns of changing educational opportunity across districts. Finally, Models 5 and 6 incorporate average demographic and resource levels. Districts with higher average achievement from the outset experienced greater improvement over the study period. As a result, average achievement varied more across districts in later cohorts than in earlier cohorts, with a 16% increase in between-district variation for the 2016 cohort (the cohort of children in kindergarten in the spring of 2016, whom we observe in Grade 3 in 2019) compared to the 2001 cohort (those who were kindergarteners in 2001 and observed in our data in Grade 8 in 2009). The full set of variables in the model together account for about 17% of the within-state variation in achievement trends. To view the full set of coefficients for these models, see Online Supplement Table 2.

Correlates of Trends in Achievement Disparities

We next examine the bivariate correlations between district characteristics and group-specific trends in achievement. Figure 3 illustrates these bivariate correlations graphically. To create this figure, we first tested whether each covariate was significantly associated with the relevant achievement disparity; for those covariates, we then estimated the bivariate association of the covariate with each of the two groups' trends separately and tested whether those associations were significant. Figure 3 plots these pairs of bivariate regression coefficients, with symbols indicating whether the coefficients were significant predictors of one or both groups' trends. For these figures, each of the covariates are standardized, so the lengths of the lines indicate the relative strength of the associations. Using the same region labels as in Figure 2, the direction of each line indicates the sign of the associations between the covariate and each group's achievement trends and the achievement disparity. Measures pointing into region B are associated with increasing disparities because although they advantage both groups, they advantage White students the most. Those in region C are associated with improving trends for White students and declining trends for Black students. Region D includes measures that are associated with declining trends for both groups.[13]

For example, in the middle panel, the longest line is labeled "Difference in SES (White-Black)." The line points down into region D of the graph, indicating that places where White families have greater socioeconomic resources than Black families tend to have declining White achievement and even more

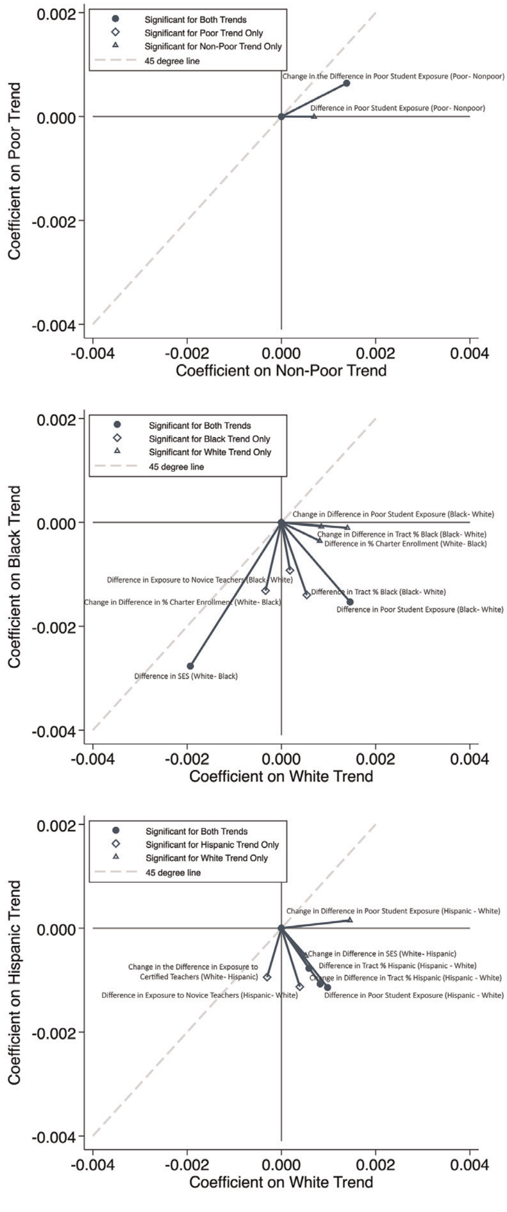

Figure 3 **Factors associated with equity and performance in trends**

Note. All coefficients are standardized. All coefficients are significantly different across trends. $p < 0.05$.

rapidly declining Black achievement; together, these result in widening White-Black achievement disparities. In other words, White and Black achievement tend to decline in more socioeconomically unequal school districts, but the associated declines in achievement are stronger for Black students' average achievement than for White students' average achievement.

One of the key findings evident in Figure 3 is that achievement disparities tend to widen the most in places where average school segregation is higher or increasing. Indeed, socioeconomic school segregation (measured as the difference in two groups' exposure to poor schoolmates) is associated with positive trends for White and non-poor students but negative trends for Black and Hispanic students; it is unrelated to trends for poor students. Further, as socioeconomic school segregation increases, so do disparities. As White and non-poor students are exposed to smaller relative shares of poor schoolmates, their achievement increases relative to Black, Hispanic, and poor students' achievement.

In addition, White-Black and White-Hispanic differences in exposure to novice teachers are associated with increasing achievement disparities, and increases in Black and poor students' relative exposure to novice teachers are associated with increases in White-Black and economic disparities, respectively.

Tables 4–6 report the estimates from multivariate regression models predicting the trends in each of the between-group achievement disparities. We focus here on the district characteristics that are significantly associated with at least two of the three types of achievement disparity trends in the preferred model (Model 6). Three key district characteristics stand out.

First, changes in school segregation—specifically, changes in the between-group difference in the average proportion of poor students in the schools of each student group—are consistently associated with achievement disparity trends. In districts where Black, Hispanic, and poor students are increasingly concentrated in high-poverty schools relative to their White and nonpoor peers, achievement disparities have increased, on average. Second, changes in differential access to certified teachers are consistently associated with changes in achievement disparities. In districts where Black, Hispanic, or poor students are in schools with increasingly fewer certified teachers compared to White and nonpoor students, achievement disparities widened, on average. And third, nonpoor-poor and White-Black disparities grew the most, on average, in places where those disparities were already wide. To view the full set of coefficients from our preferred model for each disparity trend, see Online Supplement Table 3.

To determine whether these patterns differ by tested subject, we also estimate subject-specific trends and the correlations between them. We find, first, that the overall trends in math and reading are moderately correlated ($r = 0.68$, $p < 0.001$) and can be distinguished from one another with modest reliability (reliability of the difference in trends is 0.66). The achievement disparity trends in math and reading, however, cannot be reliably distinguished from one another (the reliabilities of the differences between subjects for the nonpoor-poor, White-Black, and White-Hispanic disparities are 0.16, 0.21, and 0.25, respectively). When we fit Model 6 from Table 3 separately by subject, we find similar results across subject trends as well as no significant differences in how changes in school inputs are associated with math and reading trends. These results are summarized in Online Supplement Table 4.

Discussion

Overall, in the decade following the Great Recession, school districts' standardized test scores improved, albeit at a very modest pace compared to the rate of improvement over the prior few decades. We find that in the *average* school district, test scores changed very little (increasing by only 0.001 SD/year) from 2009 to 2019. These changes in average achievement are generally small in comparison to the sharp increases in achievement that occurred in the decades prior. One interpretation of this finding is that, on average, students' cumulative educational opportunities have not changed substantially since 2009. However, achievement trends vary considerably among districts: Test scores improved by more than 0.024 SD/year in the fastest-improving one-sixth of districts and declined by more than 0.022 SD/year in the one-sixth with the fastest-decreasing performance. These changes are substantial; for example, trends of 0.024 SD/year indicate that test scores improved by nearly a quarter of an SD over a decade.

In some places, educational opportunities have increased substantially in the last decade, while in others, opportunities have precipitously declined. In addition, average achievement has improved the most in districts that started the decade with the highest average test scores. As a result, between-district variation in average test scores has widened over the last decade. Similar patterns are described by Atteberry et al. (2021). Extending this prior work by using additional years of data, pooling across grades and subjects, and focusing on the correlates of the trends, we are able to identify how the changing composition and characteristics of districts are related to changing achievement and opportunity. Our work has implications for how districts may improve performance and equity simultaneously.

One question posed by Atteberry et al. (2021) is to what degree variation in trends is driven by population movement, resulting in changes in demographics. To be clear, our analyses show that this variation has not been driven by demographic changes in the local population. Changes in districts' economic and racial/ethnic composition explain very little of the variation in academic performance trends. We show instead that achievement improved more in districts that started the decade with higher achievement as well as districts where the proportions of novice teachers and frequently absent teachers have declined. These associations are descriptive, and they corroborate studies showing that regularly present, experienced teachers promote student learning (Clotfelter et al., 2009, 2010; Miller et al., 2008).

However, we also find that district-level trends in academic performance vary significantly by subgroup. Over the last decade, within-district nonpoor-poor and White-Black achievement disparities have widened modestly, while White-Hispanic disparities have declined. Still, there is wide variation among school districts in achievement disparity trends. Racial/ethnic and economic achievement disparities are narrowing in some districts but widening rapidly in others, which aligns with prior work (Atteberry et al., 2021).

We find that the strongest correlates of achievement disparity trends are factors related to school resources and socioeconomic segregation. Increasing economic and racial/ethnic school segregation is predictive, net of other observed factors, with growing disparities between nonpoor and poor students and between White and Black or Hispanic students. All the associations described here are descriptive and are by themselves insufficient evidence on which to make strong causal claims. However, these patterns corroborate studies showing that increasing segregation is associated with unequal educational opportunities and outcomes (Fahle et al., 2020; Reardon, Weathers, et al., 2019; Sosina & Weathers, 2019). Research on the effects of school segregation makes clear that it is harmful for low-income and minority students, in part because it concentrates Black, Hispanic, and poor students in high-poverty and under-resourced schools (Bischoff & Owens, 2019; Guryan, 2001; Johnson, 2019; Reardon, Weathers, et al., 2019).

Many high-profile education policies focus on improving either overall performance or racial, ethnic, and economic equity, sparking debates as to whether the two goals are compatible (Condron, 2011; Powers, 2004). Notably, we find that district trends in overall achievement and trends in disparities are uncorrelated, indicating that there is neither a systematic synergy nor a necessary trade-off between performance and equity over time. This means that places with increasing achievement trends are just as likely to see improvement in racial, ethnic, and economic equity as they are to see declines. Although our results indicate achievement and equity are neither incompatible nor synergistic, our district-level analysis provides an existence proof that it is possible to improve achievement for all groups and reduce achievement disparities at the same time.

Although we do not have definitive evidence regarding the best way to achieve this synergistic pattern, our analysis points to several common factors associated with improving achievement: reducing teacher absenteeism and retaining experienced teachers. In addition, several factors consistently predict narrowing achievement disparities: reducing racial/ethnic and socioeconomic inequality and segregation and reducing inequality in access to certified teachers. In other words, districts with more experienced and present teachers—and where those teachers are more equitably distributed—tend to be the districts where achievement and equity are improving the most.

A promising implication of this is that a focus on policies to improve school districts along these dimensions may lead to improvements in achievement and equity. Broader policy changes outside schools designed to reduce racial and socioeconomic inequality and segregation will likely reduce racial/ethnic and economic disparities between students as well. On the other hand, our analyses suggest that policies that lead to increasing segregation and inequality will increase educational disparities in the future.

There is much more to be learned about how to improve average achievement and equity in America's schools. Our models account for, at best, only about one-sixth of the variation in trends among districts, indicating that many other factors are at work. A number of characteristics that might be correlated with improving test scores were not available with these data. For example, differences in resources and curricula across districts and states might drive some of the changes that we observe. Likewise, variation in teachers' skills likely drives some of the changes, but the measures of teacher quality we have included may not fully capture teaching skill. In addition, changes in quality of preschool and early-childhood experiences may lead to changes in kindergarten readiness and subsequent academic performance, but such data are not widely available. More generally, it would be useful to use the SEDA data to identify districts that have shown rapid improvements in achievement and equity (and similar districts that have not) and then conduct case studies of these districts. Such work could lead to new insights and hypotheses to inform future research and education policy.

ORCID iDs

Kaylee T. Matheny https://orcid.org/0000-0002-9939-4155
Marissa E. Thompson https://orcid.org/0000-0002-9497-1400
Carrie Townley-Flores https://orcid.org/0000-0002-7464-4840

Supplemental Material

Supplemental material for this article is available online.

Notes

Kaylee T. Matheny, Marissa E. Thompson, and Carrie Townley-Flores contributed equally and share first authorship.

The article would not have been possible without the assistance of Belen Chavez and Demetra Kalogrides. The research reported here was supported in part by the Bill and Melinda Gates Foundation, the William T. Grant Foundation, the Spencer Foundation Award, the Russell Sage Foundation, the Overdeck Family Foundation, and the Institute of Education Sciences, U.S. Department of Education, through Grant R305B140009 to the Board of Trustees of the Leland Stanford Junior University and Grant R305B170015 to the University of Michigan. The opinions expressed are those of the authors and do not represent the views of the Bill and Melinda Gates Foundation, the William T. Grant Foundation, the Spencer Foundation, the Russell Sage Foundation, the Overdeck Family Foundation, the Institute of Education Sciences, the U.S. Department of Education, the University of Michigan, or the Board of Trustees of the Leland Stanford Junior University. Publicly available data related to this manuscript can be accessed at the website of The Educational Opportunity Project at Stanford University: https://edopportunity.org/get-the-data/. Direct correspondence and comments to Kaylee T. Matheny, email: *kmatheny@stanford.edu*.

[1]Based on authors' calculations using NAEP data, from 2003–2019, the average state's White-Black disparity increased by 0.02 SD, but the SD of the trend was 0.19 SD. The average White-Hispanic disparity state trend decreased by 0.04 SD/decade, with an SD of 0.09 SD/decade (NCES, 2020).

[2]The national White-Black disparity was 0.7–0.9 SD in 2019, depending on tested subject and grade; the White-Hispanic disparity was 0.50–0.65 SD.

[3]In brief, Reardon (2021) reanalyzes Hanushek et al.'s (2020) data and finds that the socioeconomic disparity in test scores has grown across cohorts born from the 1970s to 2000 in three of four data sources (NAEP-Long-Term Trends, Main NAEP, and the Trends in International Mathematics and Science Study, or TIMSS) and narrowed in the other (the Programme for International Student Assessment, or PISA). Hanushek et al.'s analysis overweights PISA data and erroneously concludes there has been no change. Hashim et al. (2020) rely on an estimator that is subject to considerable potential bias and that yields estimates inconsistent in size with other simpler estimators based on the same NAEP data.

[4]The NAEP Trial Urban District Assessment provides trend data for roughly two dozen large urban districts, but these constitute only a small fraction of the 13,000 + school districts in the United States.

[5]The linking process that was used to compile SEDA is detailed in recent papers (Kuhfeld et al., 2019; Reardon et al., 2021). Briefly, SEDA is based on the EDFacts data files collected by the U.S. Department of Education. These files contain counts of students' test scores—disaggregated by school, year, grade, subject, and subgroup—in each of a set of usually four or five coarse proficiency categories (often labeled "Below Basic," "Basic," "Proficient," and "Advanced," or similarly). These counts are used to estimate the mean test scores in each district-grade-year-subject cell for all students; for White, Black, and Hispanic students; for economically disadvantaged and non-disadvantaged students separately, using heteroskedastic ordered probit models (Ho & Reardon, 2012; Reardon & Ho, 2015). The estimated means are then scaled to the NAEP scale (Reardon et al., 2021).

[6]To standardize the NAEP-scaled estimated means, a reference pseudo-cohort is defined: the combined cohorts of students who were in Grade 4 in 2009, 2011, 2013, and 2015 (and in Grade 8 in 2013–2019). The national means and SDs of NAEP math and reading scores in each of Grades 3–8 for these cohorts are estimated by using 2009–2019 NAEP data. For grades other than 4 and 8, national means and SDs are interpolated/extrapolated for these cohorts. The means and SDs are averaged across cohorts (within grades and subjects), yielding grade- and subject-specific means and SDs that are then used to standardize the district-level scores. The resulting scores (which are those reported in the "cohort scale" in SEDA) can be interpreted as standardized relative to the average grade- and subject-specific national student-level test score distributions for the cohorts that were in Grade 4 in 2009 through 2013. Note that because test scores were standardized to the same distribution in all years (within grades and subjects), changes across years (our focus in this paper) can be interpreted as absolute rather than relative changes.

[7]For more information on estimated within-cohort growth (β_{1d}), see Reardon (2019).

[8]The publicly available SEDA pooled data do not include estimates where the OLS reliability of the parameter estimate is below 0.7—that is, when the standard error of the OLS estimate is greater than $\sqrt{\frac{3}{7}\hat{\tau}}$, where $\hat{\tau}$ is the variance of the parameter across districts. We estimate the variance of the overall trend estimate to be 0.023, meaning that we exclude trend estimates whose standard error is greater than 0.0154.

[9]We measure racial economic school segregation as the difference in the proportion of poor students in the average Black (Hispanic, poor) student's school and in the average White (White, nonpoor) student's school. We use this measure of racial economic segregation because of research showing that racial differences in the exposure to poor students is the dimension of segregation that most strongly predicts racial test score disparities (Reardon, 2016; Reardon, Weathers, et al., 2019).

[10]We calculate this as follows: Roughly two-thirds (68%) of districts' trends fall within + /– 1 SD of the average trend (so between −0.022 and + 0.024). We multiply by 10 years, which yields a range from −0.22 to + 0.24 grade levels. One-sixth of districts have achievement changes that are less than –0.22, and one-sixth have changes greater than 0.24 (in cohort-scale units).

[11]We calculate percentage changes in disparities by using the trends reported in Table 1 and the means reported in Online Supplement Table 1 to project achievement disparities in 2009 and 2019. For example, the nonpoor-poor disparity in 2014 (the middle year in our data) is .48 SD. Because the disparity grows by .005 SD each year, this means that it was approximately .48 – (5 * .005) = .46 SD in 2009 and .48 + (5 * .005) = .51 SD in 2019, an increase of .05 SD—approximately 11% of the disparity in 2009 (.05 / .46 = .11).

[12]There is a very small number of districts where mean achievement disparities are either statistically indistinguishable from zero or negative (such that they favor the structurally disadvantaged group). In these districts, improvements in the disadvantaged group's trend would then not indicate a narrowing disparity (and vice versa). However, this is true for only a very small subset of districts; more than 99% of districts in the analytic sample have White-Black and nonpoor-poor disparities that are statistically significant and favor the advantaged group, and more than 95% of districts have White-Hispanic disparities that favor White students (an additional 4% have disparities that are statistically indistinguishable from zero).

[13]Note that we created measures of inequality such that they would theoretically yield a positive coefficient in the disparity trend models in Table 4. For example, the disproportionate concentration of poor students and students of color in high-poverty schools contributes substantially to achievement disparities (Bischoff & Owens, 2019; Owens, 2018), so our measure is the extent to which the historically disadvantaged group (poor, Black, or Hispanic students) is disproportionately exposed to factors negatively associated with achievement and to which the historically advantaged group (nonpoor or White students) is disproportionately exposed to factors positively associated with achievement. Thus, all covariates in Figure 3 fall under the 45-degree line, indicating a relationship with growing disparities. However, if we had constructed our measures differently—for example, the difference in poor student exposure (White-Black) rather than (Black-White)—these measures would fall *above* the 45-degree line.

References

Abdulkadirğlu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *Quarterly Journal of Economics*, *126*(2), 699–748.

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, *114*(2), 533–575.

Atteberry, A., Bischoff, K., & Owens, A. (2021). Identifying progress toward ethnoracial achievement equity across U.S. school districts: A new approach. *Journal of Research on Educational Effectiveness*, *14*(2), 410–441.

Bandeira de Mello, V., Rahman, T., Fox, M. A., & Ji, C. S. (2019). *Mapping state proficiency standards onto NAEP scales: Results from the 2017 NAEP reading and mathematics assessments*. No. 2019–040; NCES. Institute of Education Sciences, National Center for Education Statistics.

Bassok, D., Latham, S., & Rorem, A. (2016). *Is kindergarten the new first grade?* (Vol. 2). https://doi.org/10.1177/2332858415616358

Berends, M., Lucas, S. R., & Peñaloza, R. V. (2008). How changes in families and schools are related to trends in Black-White test scores. *Sociology of Education*, *81*(4), 313–344.

Betts, J. R., & Tang, Y. E. (2011). *The effect of charter schools on student achievement: A meta-analysis of the literature*. Center on Reinventing Public Education.

Bischoff, K., & Owens, A. (2019). The segregation of opportunity: Social and financial resources in the educational contexts of lower-and higher-income children, 1990–2014. *Demography*, *56*(5), 1635–1664.

Blazar, D., & Schueler, B. (2022). Why do school districts matter? An interdisciplinary framework and empirical review. *EdWorkingPapers*, *22*(581). https://doi.org/10.26300/58M4-FS65

Boozer, M. A., & Rouse, C. E. (2001). Intraschool variation in class size: Patterns and implications. *Journal of Urban Economics*, *50*(1), 163–89. https://doi.org/10.1006/juec.2001.2216.

Card, D., & Rothstein, J. (2007). Racial segregation and the Black–White test score gap. *Journal of Public Economics*, *91*(11–12), 2158–2184. https://doi.org/10.1016/j.jpubeco.2007.03.006

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review*, *24*(4), 377–392.

Clotfelter, C. T., Ladd, H., & Vigdor, J. (2007). *How and why do teacher credentials matter for student achievement?* No. 12828; NBER Working Paper Series. National Bureau of Economic Research. http://www.nber.org/papers/w12828.pdf

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2009). Are teacher absences worth worrying about in the United States? *Education Finance and Policy*, *4*(2), 115–149. https://doi.org/10.1162/edfp.2009.4.2.115

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources*, *45*(3), 655–681. https://doi.org/10.3368/jhr.45.3.655

Cohodes, S., & Parham, K. (2021). *Charter schools' effectiveness, mechanisms, and competitive influence*. No. w28477; p. w28477. National Bureau of Economic Research. https://doi.org/10.3386/w28477

Condron, D. J. (2011). Egalitarianism and educational excellence: Compatible goals for affluent societies? *Educational Researcher*, *40*(2), 47–55. https://doi.org/10.3102/0013189X11401021

Duncan, G. J., & Murnane, R. J. (Eds.). (2011). *Whither opportunity? Rising inequality, schools, and children's life chances*. Russell Sage Foundation, Spencer Foundation.

Every Student Succeeds Act, Pub. L. No. 114–95, 20 USC (2015). https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf

Fahle, E. M., Chavez, B., Kalogrides, D., Shear, B. R., Reardon, S. F., & Ho, A. D. (2021). *Stanford Education Data Archive Technical documentation (Version 4.1)*. https://stacks.stanford.edu/file/druid:db586ns4974/seda_documentation_4.1 .pdf

Fahle, E. M., Reardon, S. F., Kalogrides, D., Weathers, E. S., & Jang, H. (2020). Racial segregation and school poverty in the United States, 1999–2016. *Race and Social Problems*, *12*(1), 42–56. https://doi.org/10.1007/s12552-019-09277-w

Gamoran, A., & An, B. P. (2016). Effects of school segregation and school resources in a changing policy context. *Educational Evaluation and Policy Analysis*, *38*(1), 43–64. https://doi.org/10.3102/0162373715585604

Garcia, E. (2015). Inequalities at the starting gate: Cognitive and noncognitive skills gaps between 2010–2011 kindergarten classmates. Report. *Economic Policy Institute*.

Gilraine, M., Petronijevic, U., & Singleton, J. D. (2019). *Horizontal differentiation and the policy effect of charter schools*. Working Paper No. 19–80; EdWorkingPaper. Annenberg Institute at Brown University. http://www.edworkingpapers.com/ai19-80

Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, *22*(2), 129–145.

Greenberg, E., Blagg, K., & Rainer, M. (2019). *Measuring student poverty: Developing accurate counts for school funding, accountability, and research*. Center on Education Data and Policy Research Report. Urban Institute.

Guryan, J. (2001). *Does money matter? Regression-discontinuity estimates from education finance reform in Massachusetts*. Working Paper No. 8269. National Bureau of Economic Research. http://www.nber.org/papers/w8269

Hanushek, E., Peterson, P., Talpey, L., & Woessmann, L. (2020). *Long-run trends in the U.S. SES-achievement gap*. No. w26764; p. w26764. National Bureau of Economic Research. https://doi.org/10.3386/w26764

Hashim, S. A., Kane, T. J., Kelley-Kemple, T., Laski, M. E., & Staiger, D. O. (2020). *Have income-based achievement gaps widened or narrowed?* Working Paper No. 27714; NBER Working Paper Series. http://www.nber.org/papers/w27714

Ho, A. D. (2008). The problem with "proficiency": Limitations of statistics and policy under No Child Left Behind. *Educational Researcher*, *37*(6), 351–360. https://doi.org/10.3102/0013189X08323842

Ho, A. D., & Reardon, S. F. (2012). Estimating achievement gaps from test scores reported in ordinal "proficiency" categories. *Journal of Educational and Behavioral Statistics*, *37*(4), 489–517. https://doi.org/10.3102/1076998611411918

Hussar, B., Zhang, J., Hein, S., Wang, K., Roberts, A., Cui, J., Smith, M., Mann, F. B., Barmer, A., & Dilig, R. (2020). *The condition of education 2020*. NCES 2020-144. National Center for Education Statistics.

Jackson, C. K. (2020). Does school spending matter? The new literature on an old question. In L. Tach, R. Dunifon, & D. L. Miller (Eds.), *Confronting inequality: How policies and practices shape children's opportunities* (pp. 165–186). American Psychological Association. https://doi.org/10.1037/0000187-008

Jacob, B., Berger, D., Hart, C., & Loeb, S. (2016). Can technology help promote equality of educational opportunities? *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *2*(5), 242–271.

Johnson, R. C. (2019). *Children of the dream: Why school integration works* (1st ed.). Basic Books and Russell Sage Foundation Press.

Kalogrides, D., Loeb, S., & Béteille, T. (2013). Systematic sorting: Teacher characteristics and class assignments. *Sociology of Education*, *86*(2), 103–123.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, *114*(2), 497–532.

Kuhfeld, M., Domina, T., & Hanselman, P. (2019). Validating the SEDA measures of district educational opportunities via a common assessment. *AERA Open*, *5*(2), 233285841985832. https://doi.org/10.1177/2332858419858324

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Economic Policy Institute.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, *29*(1), 67–101. https://doi.org/10.3102/10769986029001067

Miller, R. T., Murnane, R. J., & Willett, J. B. (2008). Do teacher absences impact student achievement? Longitudinal evidence from one urban school district. *Educational Evaluation and Policy Analysis*, *30*(2), 181–200.

Morgan, I., & Amerikaner, A. (2018). *Funding gaps 2018: An analysis of school funding equity across the US and within each state*. Education Trust.

National Center for Education Statistics. (2013). *The nation's report card: Trends in academic progress 2012*. NCES 2013-456. Institute of Education Sciences, U.S. Department of Education. https://nces.ed.gov/nationsreportcard/subject/publications/main2012/pdf/2013456.pdf

National Center for Education Statistics. (2018a). *2017 mathematics and reading assessments: Highlighted results at grades 4 and 8 for the nation, states, and districts*. NCES 2018-037. Institute of Education Sciences, U.S. Department of Education. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018037

National Center for Education Statistics. (2018b). *Trial urban district assessment*. Institute of Education Sciences, U.S. Department of Education. https://nces.ed.gov/nationsreportcard/tuda/

National Center for Education Statistics. (2020). *NAEP data explorer*. Institute of Education Sciences, U.S. Department of Education. https://www.nationsreportcard.gov/ndecore/xplore/nde

No Child Left Behind Act of 2001 (NCLB), Pub. L. No. 107–110, 20 USC (2002).

Orfield, G., & Frankenberg, E. (2014). Increasingly segregated and unequal schools as courts reverse policy. *Educational Administration Quarterly*, *50*(5), 718–734.

Owens, A. (2018). Income segregation between school districts and inequality in students' achievement. *Sociology of Education*, *91*(1), 1–27. https://doi.org/10.1177/0038040717741180

Owens, A., Reardon, S. F., & Jencks, C. (2016). Income segregation between schools and school districts. *American Educational Research Journal*, *53*(4), 1159–1197. https://doi.org/10.3102/0002831216652722

Owens, A., Reardon, S. F., Kalogrides, D., Jang, H., & Tom, T. (2022). *Trends in racial/ethnic and economic school segregation, 1991–2020*. The Segregation Index Research Brief. https://socialinnovation.usc.edu/wp-content/uploads/2022/05/Trends-in-Racial-Ethnic-Segregation_FINAL.pdf

Page, L., Murnane, R., & Willett, J. (2008). *Trends in the Black-White achievement gap: Clarifying the meaning of within- and between-school achievement gaps*. No. w14213; p. w14213. National Bureau of Economic Research. https://doi.org/10.3386/w14213

Podolsky, A., Kini, T., Bishop, J., & Darling-Hammond, L. (2016). *Solving the teacher shortage how to attract and retain excellent educators*. Learning Policy Institute. https://learningpolicyinstitute.org/sites/default/files/product-files/Solving_Teacher_Shortage_Attract_Retain_Educators_REPORT.pdf

Powers, J. M. (2004). High-stakes accountability and equity: Using evidence from California's public schools accountability act to address the issues in *Williams v. State of California*. *American Educational Research Journal*, *41*(4), 763–795. https://doi.org/10.3102/00028312041004763

Reardon, S. F. (2011). The widening academic achievement gap between rich and poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–115). Russell Sage Foundation.

Reardon, S. F. (2016). School segregation and racial academic achievement gaps. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *2*(5), 34–57. https://doi.org/10.7758/RSF.2016.2.5.03

Reardon, S. F. (2019). Educational opportunity in early and middle childhood: Using full population administrative data to study variation by place and age. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *5*(2), 40. https://doi.org/10.7758/rsf.2019.5.2.03

Reardon, S. F. (2021). *The economic achievement gap in the US, 1960–2020: Reconciling recent empirical findings*. Working Paper No. 21–09; CEPA Working Paper. Stanford Center for Education Policy Analysis. https://cepa.stanford.edu/sites/default/files/wp21-09-v112021.pdf

Reardon, S. F., & Ho, A. D. (2015). Practical issues in estimating achievement gaps from coarsened data. *Journal of Educational and Behavioral Statistics*, *40*(2), 158–189. https://doi.org/10.3102/1076998615570944

Reardon, S. F., Kalogrides, D., & Ho, A. D. (2021). Validation methods for aggregate-level test scale linking: A case study mapping school district test score distributions to a common scale. *Journal of Educational and Behavioral Statistics*, *46*(2), 138–167. https://doi.org/10.3102/1076998619874089

Reardon, S. F., Kalogrides, D., & Shores, K. (2019). The geography of racial/ethnic test score gaps. *American Journal of Sociology*, *124*(4), 1164–1221. https://doi.org/10.1086/700678

Reardon, S. F., & Owens, A. (2014). 60 years after *Brown*: Trends and consequences of school segregation. *Annual Review of Sociology*, *40*(1), 199–218. https://doi.org/10.1146/annurev-soc-071913-043152

Reardon, S. F., Papay, J. P., Kilbride, T., Strunk, K. O., Cowen, J., An, L., & Donohue, K. (2019). *Can repeated aggregate cross-sectional data be used to measure average student learning rates? A validation study of learning rate measures in the Stanford Education Data Archive*. No. 19–08; CEPA Working Paper. Stanford Center for Education Policy Analysis.

Reardon, S. F., & Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open*, *2*(3), 1–18. https://doi.org/10.1177/2332858416657343

Reardon, S. F., Robinson-Cimpian, J. P., & Weathers, E. S. (2014). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. In H. F. Ladd & M. E. Goertz (Eds.), *Handbook of research in education finance and policy* (2nd ed.). Routledge.

Reardon, S. F., Weathers, E. S., Fahle, E. M., Jang, H., & Kalogrides, D. (2019). *Is separate still unequal? New evidence on school segregation and racial academic achievement gaps*. Working Paper No. 19–06; CEPA Working Paper. Stanford Center for Education Policy Analysis. https://cepa.stanford.edu/sites/default/files/wp19-06-v092021.pdf

Rich, P., Candipan, J., & Owens, A. (2021). Segregated neighborhoods, segregated schools: Do charters break a stubborn link? *Demography*, *58*(2), 471–498.

Rivkin, S. G. (2000). School desegregation, academic attainment, and earnings. *Journal of Human Resources*, *35*(2), *333*. https://doi.org/10.2307/146328

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458.

Roby, D. (2013). Teacher attendance effects on student achievement: Research study of Ohio schools. *Education*, *134*(2), 201–206.

Roza, M. (2010). *Educational economics: Where do school funds go?* Urban Institute Press.

Shakeel, M. D., & Peterson, P. E. (2022). A half century of progress in US student achievement: Agency and flynn effects, ethnic and SES differences. *Educational Psychology Review*. https://doi.org/10.1007/s10648-021-09657-y

Shores, K., & Steinberg, M. P. (2019). Schooling during the great recession: Patterns of school spending and student achievement using population data. *AERA Open*, *5*(3). https://doi.org/10.1177/2332858419877431

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*, 417–453.

Sosina, V. E., & Weathers, E. S. (2019). Pathways to inequality: Between-district segregation and racial disparities in school district expenditures. *AERA Open*, *5*(3).

Wright, B. D. (2005). A history of social science measurement. *Educational Measurement: Issues and Practice*, *16*(4), 33–45. https://doi.org/10.1111/j.1745-3992.1997.tb00606.x

Yee, D., & Ho, A. D. (2015). Discreteness causes bias in percentage-based comparisons: A case study from educational testing. *American Statistician*, *69*(3), 174–181.

Manuscript received April 14, 2021
Final revision received July 13, 2022
Accepted September 16, 2022