



## Academic mobility in U.S. public schools: Evidence from nearly 3 million students

Wes Austin<sup>a</sup>, David Figlio<sup>b</sup>, Dan Goldhaber<sup>c</sup>, Eric A. Hanushek<sup>d</sup>, Tara Kilbride<sup>e</sup>, Cory Koedel<sup>f,\*</sup>, Jaeseok Sean Lee<sup>g</sup>, Jin Lou<sup>h</sup>, Umut Özek<sup>i</sup>, Eric Parsons<sup>g</sup>, Steven G. Rivkin<sup>j</sup>, Tim R. Sass<sup>k</sup>, Katharine O. Strunk<sup>l</sup>

<sup>a</sup> Environmental Protection Agency, United States

<sup>b</sup> University of Rochester, United States

<sup>c</sup> CALDER at the American Institutes for Research (AIR) and CEDR at the University of Washington, United States

<sup>d</sup> Hoover Institution, Stanford University, United States

<sup>e</sup> Michigan State University, United States

<sup>f</sup> University of Missouri, Department of Economics, 615 Locust Street, #E231, Columbia, MO 65211, United States

<sup>g</sup> University of Missouri, United States

<sup>h</sup> University of Texas at Dallas, United States

<sup>i</sup> RAND, United States

<sup>j</sup> University of Illinois-Chicago, United States

<sup>k</sup> Georgia State University, United States

<sup>l</sup> University of Pennsylvania, United States

### A B S T R A C T

We use administrative panel data from seven states covering nearly 3 million students to document and explore variation in “academic mobility,” a term we use to describe the extent to which students’ ranks in the distribution of academic performance change during their public schooling careers. We find that student ranks are highly persistent during elementary and secondary education—that is, academic mobility is limited in U.S. schools on the whole. Still, there is non-negligible variation in the degree of upward mobility across some student subgroups as well as individual school districts. On average, districts that exhibit the greatest upward academic mobility serve more socioeconomically advantaged populations and have higher value-added to student achievement.

### 1. Introduction

An effective and equitable education system can be viewed as a form of social insurance against a poor birth endowment—even in the face of considerable obstacles, access to effective schools can provide a pathway to success. However, the performance of the U.S. education system in this regard leaves much to be desired. Students from different socioeconomic backgrounds enter K-12 schools already exhibiting large achievement gaps and these gaps persist, or even widen, as they progress through school (Haskins and Rouse, 2005; Jang and Reardon, 2019; Reardon, 2011). It would be a mistake to conclude U.S. public schools do not contribute to social equity (counterfactual equity conditions would almost surely be worse in their absence), but the inability of schools to narrow achievement gaps during elementary and secondary education is an ongoing policy concern.

In this article, we introduce the concept of “academic mobility” to study the persistence of student placements in the distribution of academic performance during elementary and secondary education. An education system with high academic mobility is one where students’ early-grade ranks are less predictive of their later-grade ranks. We estimate academic mobility using administrative panel data from seven states covering almost 3 million students.

Our estimation procedures borrow from tools developed in a related literature on economic mobility including Chetty, Hendren, Kline, and Saez (2014) and Chetty, Hendren, Jones, and Porter (2018).<sup>1</sup> We assess students’ initial performance levels using test scores in the third grade, which is the earliest grade we have universal data on test performance in public schools. Then, we use four long-term outcomes to estimate academic mobility during K-12 education: eighth-grade test performance, high-school test performance, on-time high school graduation, and high

\* Corresponding author.

E-mail address: [koedelc@missouri.edu](mailto:koedelc@missouri.edu) (C. Koedel).

<sup>1</sup> In turn, these studies build on a large prior literature on economic mobility—for reviews see Black and Devereux (2011) and Solon (1999).

school graduation within one year of on-time. For most of our analysis we focus on upward mobility among initially low-achieving students. Again following the recent literature on economic mobility, we define “absolute upward mobility” as academic mobility measured at the 25th percentile of the distribution of initial performance ranks.

We find that students’ ranks in the distribution of academic performance are highly persistent during K-12 education. It follows that absolute upward mobility is low. For example, on average across our seven states, a student at the 25th percentile of the academic performance distribution in the third grade can be expected to perform at roughly the 30th percentile by high school. Moreover, conditional on beginning with a low rank, students from more advantaged backgrounds generally have greater upward mobility than their less advantaged peers. Asian students in particular have very high upward mobility. These results buttress existing research on the persistence, and even widening, of achievement gaps during K-12 schooling.

We also explore variation in academic mobility across school districts. Despite finding academic mobility is low on average, we document statistically and economically significant variance in upward mobility across districts. We decompose the variance into two components: “baseline mobility” and “relative mobility.” Districts with high baseline mobility promote gains throughout the performance distribution; initial low achievers are caught in a rising tide that lifts all boats in these districts. In districts with high relative mobility, initial low achievers gain on their higher-achieving peers as they progress through school—i.e., the within-district achievement gap narrows over time.<sup>2</sup> Variation across districts in both components contributes to the total variance in absolute upward mobility, but we find that most of the variance in upward mobility is driven by cross-district differences in baseline mobility. While our mobility metrics are descriptive and should not be interpreted causally, these results are informative about the ways in which districts are likely to help low-achieving students improve. The low variation in relative mobility suggests limited differences in success across districts at reducing within-district achievement gaps.

We also explore the correlates of absolute upward mobility at the district level and show mobility is largest in districts serving more socioeconomically advantaged students. For instance, mobility is higher in districts where local-area incomes, education levels, and residential stability are higher, and where more Asian and White families live. Independent of these attributes, district value-added to student achievement is also a strong predictor of high upward mobility. When we estimate district-level academic mobility separately for Black, Hispanic, and low-income students to allow for heterogeneity in the correlates of mobility, we generally find the same factors predict upward academic mobility for all students.

## 2. Data and measurement of academic mobility

### 2.1. Data

We use state administrative panel data from public schools in seven states—Georgia, Massachusetts, Michigan, Missouri, Oregon, Texas, and Washington. We assemble cohorts of all students with standardized test scores in math and English language arts (ELA) in the third grade and follow them through high school.

<sup>2</sup> Both a student’s absolute position in the performance distribution and a student’s relative position within a class, school, or district are important outcomes of interest. A student’s absolute position is important given causal evidence on the link between test scores and later life outcomes (Goldhaber and Özek, 2019). There is also increasing evidence that a student’s relative rank has independent effects on student behaviors and outcomes, as social comparisons help to shape ability beliefs. See, for instance, Cicala et al. (2018), Denning et al. (2020), Elsner and Isphording (2017a, 2017b), Elsner et al. (2019), and Murphy and Weinhardt (2020).

Table 1 reports descriptive information for the third-grade cohorts in each state, as well as for K-12 students in the entire U.S. for comparison. We track academic mobility for two to four cohorts of students in the sample states between 2005 and 06 and 2008–09 (hereafter, including in Table 1, we identify school years by the spring year; e.g., 2006 for “2005–06”).<sup>3</sup> Our analysis includes about 2.9 million students. The sample states exhibit substantial heterogeneity in their populations. For example, the shares of Black and Hispanic students across states range from 3.0 to 38.1 and 4.0 to 47.7, respectively. There is also considerable variation across states in the shares of students: receiving free or reduced-price lunch (FRL), identified for an Individualized Education Program (IEP), and who are geographically mobile.<sup>4</sup> In addition, the structure of the education system differs across states in terms of the shares of schools located in urban/suburban/rural areas, and the numbers of districts and schools, both in absolute and per-capita terms. While our sample is not representative of the United States, the seven states are diverse and provide substantively different evaluation contexts.<sup>5</sup>

Under the No Child Left Behind and Every Student Succeeds Acts, all students in public schools are tested in math and English Language Arts (ELA) annually in grades 3–8. As a result, our analysis of academic mobility between grades 3 and 8 is fairly uniform across states (although each state administers its own tests). At the high school level, students must be tested at least once but there is flexibility over the grade and subject. To assess academic mobility based on high-school achievement, in each state we identify the exam with the highest coverage rate administered in a common grade. These tests are shown in Table 2.<sup>6</sup> With the exception of Michigan, which has a universal ACT/SAT policy, the common-grade requirement is such that the test subject is ELA-based. This is because the high school English curriculum is more rigid than in other subjects. The focal tests are administered mostly in the tenth and eleventh grades (except in Georgia), have high coverage rates, and are overwhelmingly taken in a common grade. In Oregon there is no test given overwhelmingly in a common grade in high school, so we omit Oregon from the portion of the analysis focused on high school test ranking.

In addition to assessing test-based academic mobility, we also assess mobility in terms of the likelihood of high school graduation. We consider both on-time graduation and graduation within one year of on-time.

### 2.2. Measuring academic mobility

#### 2.2.1. Overview

Our methodological approach follows the framework developed by Chetty, Hendren, Kline, and Saez (CHKS, 2014) and Chetty, Hendren, Jones, and Porter (CHJP, 2018) to study intergenerational economic

<sup>3</sup> The earliest cohort is from 2006 because this is the first year of consistent testing in grades 3–8 in most states, and the latest cohort is from 2009 because this is the oldest cohort for whom we can track graduation outcomes (within one year of on-time graduation) using our data panels.

<sup>4</sup> Geographic mobility is defined by students who are enrolled in more than one school during the year in which they took the third grade test. States differ in terms of the frequency of collecting school enrollment information, which may account for some of the heterogeneity across states in this variable. The FRL data used for these cohorts pre-date the option of schools to use the Community Eligibility Provision (Koedel and Parsons, 2021).

<sup>5</sup> Appendix Table A1 further shows that enrollment shares in charter schools in our third-grade cohorts is small in all states, ranging from 0 to 7.7 percent, with a median value of 1.8 percent.

<sup>6</sup> The requirement of a common grade limits concerns about the confounding effect of test timing on our cross-district measures of academic mobility, which has come up most often with respect to studies of Algebra-I end-of-course exam performance (Clotfelter, Ladd, and Vigdor, 2015; Domina et al., 2015; Parsons et al., 2015).

**Table 1**  
Definition of the analytic sample and descriptive statistics at panel entry for each state.

	Cohort Years	N (entry cohorts)	Pct. Black	Pct. Hispanic	Pct. FRL	Pct. IEP	Pct. Mobile	Pct. Urban	Pct. Suburban	# of Districts	# of Schools	Private Schl Enrl % (2008)
Georgia	2007–2009	376,427	38.08	12.72	56.02	12.47	8.12	8.81	39.66	182	1255	8.71
Massachusetts	2007–2008	139,337	7.83	13.94	31.65	17.30	2.32	20.11	68.19	304	1,116	13.60
Michigan	2006–2009	453,946	19.03	5.72	40.99	10.92	12.19	20.96	44.39	755	2,039	8.59
Missouri	2006–2009	264,612	18.17	4.00	46.34	15.16	6.62	18.79	30.87	548	1,200	12.05
Oregon	2006–2008	123,833	3.03	16.83	47.59	15.37	4.03	30.69	25.60	208	1,086	10.49
Texas	2006–2009	1,309,114	13.54	47.68	57.84	5.86	6.68	42.27	27.90	1,173	4,338	5.96
Washington	2006–2008	218,051	5.70	15.80	42.26	11.44	1.04	26.12	45.30	296	1,254	9.17
Entire U.S.	2008	–	17.04	21.13	42.95	12.35	–	29.03	35.10	–	–	–

Table Notes: “Cohort Years” refers to the years of panel entry for the cohorts included in the analytic sample; i.e., the years in which the students were in the third grade. The spring year is used to indicate the academic year (e.g., 2009 = 2008–09 school year). Students who took both the Math and ELA third-grade state tests are included in the core sample. For Washington and Massachusetts, in earlier years of data, enrollment surveys were not conducted frequently, which likely contributes to the low reported mobility rates in those two states. In more recent data, the mobility rates in Massachusetts and Washington are around 5 and 8–9 percent, respectively. Note that the numbers of schools and districts indicate the numbers of unique schools and districts included in the analysis in each state. Data for the “Entire U.S.” are reported in the bottom row of the table for context and taken from the 2008 common core of data and are for students in public K-12 elementary and secondary grades. Note that we do not report a mobility percentage because a comparable variable is not available in the common core of data.

**Table 2**  
High school exams by state.

	HS Exam	Grade Typically Taken	Pct. Of Cohort Students Taking the Exam On-Grade	Pct. Of Cohort Students Taking the Exam Within 1 Year of On-Grade
Georgia	Literature EOC	9	97.7	2.0
Massachusetts	MCAS ELA	10	99.5	0.2
Michigan	ACT/SAT	11	99.3	0.7
Missouri	English II EOC	10	93.1	3.8
Oregon	Not Applicable			
Texas	Reading/English II EOC	10	94.1	5.7
Washington	HSPE ELA, SBAC ELA	10, 11	98.3	1.4

Notes: In Washington, a test change led to the change in the grade in which the third-grade cohorts took their high school exit exams (from grade 10 to 11), as shown in the Table. Michigan transitioned from the ACT to the SAT in the 2016–17 school year. The first two analysis cohorts took the ACT in 11th grade; the second two cohorts took the SAT in 11th grade. In Oregon, there is no single high school test given to more than 90 percent of students in a fixed grade to support our analysis of mobility using HS test achievement.

mobility. The test-based mobility metrics we use link percentile rankings in the test distribution at different points in the schooling career. Like CHKS and CHJP, we have sufficiently rich data to describe the joint distribution of early- and late-career student performance nonparametrically in the form of 100x100 percentile matrices for each outcome and state. However, a key insight from CHKS that permits a more parsimonious presentation is that the rank-rank relationship between intergenerational economic outcomes is functionally linear. This is also true in our application, allowing us to summarize academic mobility with just the slope and intercept parameters from a linear regression.

We illustrate the linearity of the rank-rank relationships using binned scatterplots of students’ entry and late-outcome test ranks in each state. Fig. 1 shows scatterplots from Georgia as an example. The top two graphs are for test scores. The entry ranks on the horizontal axis are the average rank in math and ELA in the third grade. The outcome ranks on the vertical axes are: (1) the average rank on math and English Language Arts (ELA) tests in the eighth grade and (2) the rank on the high school test listed in Table 2. Similar graphs for all sample states are provided in Appendix Figure A1.

The test-based rank-rank relationships are linear, at least to a close

approximation, in all states for all tests (we discuss the scatterplots for graduation outcomes below). Given this, the mapping between students’ early- and late-career outcomes can be summarized by equation (1):

$$O_i = \alpha + \beta R_i + \varepsilon_i \tag{1}$$

where  $O_i$  is a late-career outcome rank for student  $i$  and  $R_i$  is the initial rank in the third grade.

When the rank-rank relationship is estimated on the entire population (in our case, an entire state),  $\alpha$  and  $\beta$  are mechanically linked. To see why, note the estimated regression line must pass through the mean of the data, which in a percentiles-on-percentiles model means it must pass through (50, 50). As a result, the mobility relationship is fully captured by the slope coefficient,  $\beta$ , which also defines the y-intercept,  $\alpha$ .

When we disaggregate the data below the state level—i.e., for subpopulations of students within a state or for individual school districts—the parameters  $\alpha$  and  $\beta$  are separately identified and provide unique information about baseline and relative mobility, respectively. This is because the rank-rank regression line need not pass through the point (50, 50) for each subpopulation. Consider the following versions of equation (1) that permit subgroup analyses:

$$O_{is} = \alpha_s + \beta_s R_{is} + \varepsilon_{is} \tag{2}$$

$$O_{id} = \alpha_d + \beta_d R_{id} + \varepsilon_{id} \tag{3}$$

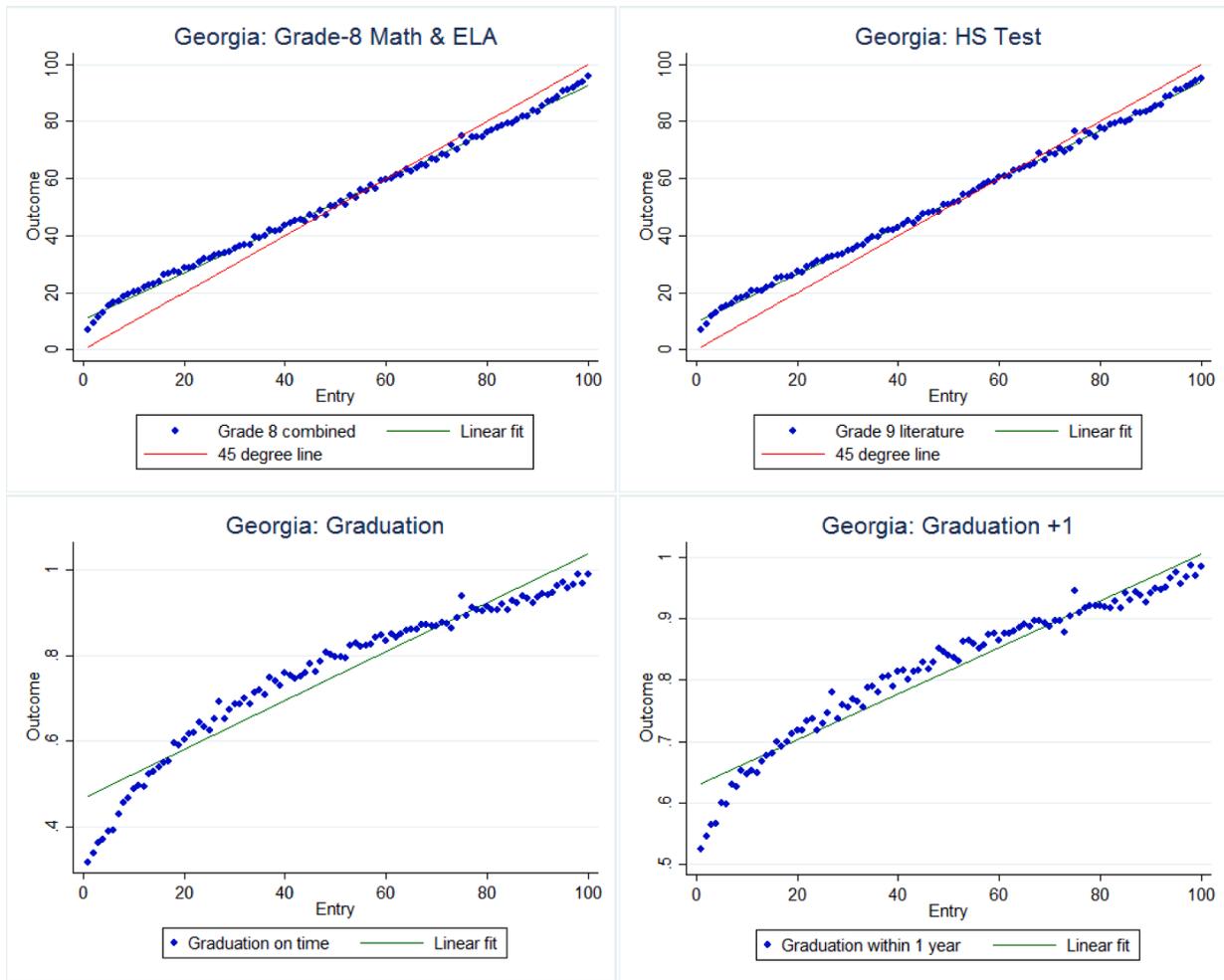
In equation (2), the subscript  $s$  indicates group membership for student  $i$ . We define groups  $s$  by race/ethnicity, FRL eligibility, and the urbanicity of the school attended in the third grade (urban, suburban, or rural). In equation (3), the subscript  $d$  identifies students who attend district  $d$  in the third grade. As long as the dependent and independent variables in equations (2) and (3)—which are in percentiles—continue to be calculated from the full statewide distributions of test scores, the intercepts and slopes for the groups indexed by  $s$  and  $d$  are separately identified and provide unique information about the nature of academic mobility.

Total academic mobility at initial percentile  $p$ , inclusive of baseline and relative mobility, can be expressed for district  $d$  as follows:

$$\bar{O}_{pd} = \alpha_d + \beta_d p \tag{4}$$

Similarly,  $\bar{O}_{ps}$  gives the student-subgroup analog. Like in CHKS, we focus on total mobility of students at the 25th percentile of the initial performance distribution to produce measures of absolute upward mobility for initially low-achieving students, denoted by  $\bar{O}_{25}$ . From equation (4),  $\bar{O}_{25}$  for students in district  $d$  is estimated by  $\hat{\alpha}_d + \hat{\beta}_d * 25$ .

It is straightforward to interpret a higher value of baseline mobility ( $\alpha$ ) as a positive attribute, but the same is not true of relative mobility



**Fig. 1.** Binned scatter plots with percentile ranks on the 3rd grade test on the horizontal axis (averaged across math and ELA), and either test-outcome percentiles or graduation rates on the vertical axis, in Georgia. Notes: This figure shows binned scatterplots of the raw (binned) entry and outcome ranks in Georgia. Appendix A shows similar scatterplots for all other states and all outcomes.

( $\beta$ ). For instance, consider two districts where initially low-achieving students perform similarly. The district with higher relative mobility will be the one where initial high achievers perform relatively worse. Many researchers and education systems use the district achievement gap as a measure of performance, but this is an insufficient measure due to the potential tradeoff between inequality and the outcome level for initial low achievers. A rightward shift in the entire achievement distribution that is more pronounced at higher achievement percentiles would increase both the achievement gap and the expected outcome percentiles for initial low achievers, while a leftward shift in the entire distribution that is again more pronounced at the upper percentiles would reduce both the achievement gap and the expected outcome percentiles for initial low achievers.

Finally, we turn to the application of this framework to analyze graduation outcomes. Although graduation is a binary outcome, the academic mobility parameters are conceptually similar in the graduation models (CHKS, 2014). For example,  $O_{25d}$  for on-time graduation indicates the likelihood of high-school graduation for a student at the 25th percentile of the third grade performance distribution from district  $d$ . This likelihood can be compared to the likelihood of graduation for a student in the 25th percentile in district  $c$  to compare mobility across districts measured by graduation.

The bottom graphs in Fig. 1 show binned scatterplots mapping students' entry percentiles to their graduation outcomes in Georgia and Appendix Figure A1 shows similar scatterplots for the other states. The plots are roughly linear throughout most of the initial rank distribution

(about the upper 80 percent). The nonlinearity at lower entry percentiles is due to a combination of strong floor effects and the fact that graduation is a binary outcome.

2.2.2. Estimation details

2.2.2.1. Measurement error in Students' initial test scores. The initial percentile ranks are based on students' third grade scores on high-stakes state tests. These tests meet the highest standards of test publishers in terms of their reliability, but they are not error-free.<sup>7</sup> Measurement error in these tests comes from two broad sources (Boyd et al., 2013; Lockwood and McCaffrey, 2014): (1) the tests rely on a finite number of questions to assess student knowledge, making student scores subject to test-item sampling variance, and (2) idiosyncratic factors associated with student or test circumstances on the day of the test (e.g., the proverbial dog barking in the parking lot).

Measurement error in students' initial scores will result in mean reversion. If left unaccounted for, this will lead us to overstate academic mobility. To illustrate consider an extreme scenario where initial scores are comprised entirely of error. Under the standard assumption that the error is uncorrelated with the outcome, the expected value of  $\beta$  would be

<sup>7</sup> The test reliability estimates in our sample states are consistently around 0.90 or above, which is at the upper end of the recommended range in the psychometrics literature (Tavakol and Dennick, 2011).

zero, implying perfect academic mobility. More generally, measurement error in students' initial scores will attenuate our estimates of  $\beta$  and correspondingly inflate our estimates of  $\alpha$ .<sup>8</sup>

Within the context of a latent-ability framework, we use two approaches to address the measurement error problem. Both leverage the fact that we observe two different measures of skill in the third grade from the math and ELA tests. Each measure can be written as a function of general skill, an orthogonal subject-specific skill, and an error. Formally, we write:

$$S_i^M = S_i + M_i + e_i^M \tag{5}$$

$$S_i^E = S_i + E_i + e_i^E \tag{6}$$

where  $S_i^M$  and  $S_i^E$  are observed test scores for student  $i$  in math ( $M$ ) and ELA ( $E$ ), respectively,  $S_i$  is general skill,  $M_i$  and  $E_i$  are subject specific skills constructed to be orthogonal to  $S_i$ , and  $e_i^M$  and  $e_i^E$  are test-specific random measurement errors.  $e_i^M$  and  $e_i^E$  are assumed to be mean-zero and independent.

Our preferred error correction is a two-step procedure where we first average the third-grade ranks in math and ELA to set the initial rank. By averaging the two noisy measures, the error variance is reduced. Then, we make an additional correction to remove error deriving from the testing instruments themselves—i.e., due to sampling variance in the items that appear on the tests. Our correction is based on test reliability ratios reported by test publishers, which we incorporate into our models using a standard errors-in-variables (EIV) regression framework. The EIV models disattenuate  $\beta$  by the expected value of the attenuation bias caused by the measurement error (and correspondingly shrink  $\alpha$ ). Procedurally, the error variance is subtracted from the total variance in the initial ranks to calculate the error-adjusted parameter  $\beta$  (Fuller, 1987; Lockwood and McCaffrey, 2014).

A complication is that we use the average of the entry ranks as the initial rank variable, but the reliability ratios from test publishers are for the individual tests. Define  $r_m$  and  $r_e$  as the reliability ratios for the third grade math and ELA tests individually, and  $\theta_{m,e}$  as the correlation of performance on the two tests. Following Wang and Stanley (1970), the reliability of average performance across the two tests is given by:  $r_c = \frac{0.25r_m + 0.25r_e + 0.50\theta_{m,e}}{0.50 + 0.50\theta_{m,e}}$ . We use reliabilities based on this equation in our EIV models for each state.<sup>9,10</sup>

Our approach to adjust for measurement error has strengths and weaknesses. A key strength is that it is more efficient than available alternatives—namely instrumental variables (IV), which we discuss below—resulting in more precise estimation. This is especially important when we estimate mobility parameters for each district individually. However, it is not a comprehensive correction and has two notable limitations. First, the averaging strategy, while conceptually appealing, is incomplete since we observe just two scores (Ashenfelter and Krueger, 1994). And while the EIV correction helps, it only addresses measurement error associated with the testing instruments and ignores other sources of error. Therefore, we do not expect our estimates of  $\beta$  to be fully disattenuated; rather, they are lower-bound estimates. This means

the EIV specifications will overstate relative academic mobility to some degree (recall that a lower value of  $\beta$  corresponds to more academic mobility). The second limitation is that the EIV approach does not allow for subgroup heterogeneity in the magnitude of measurement error (the publisher-reported test reliabilities are averages across all students and not available for student subgroups). For instance, below we compare academic mobility between FRL and non-FRL students. If the magnitude of measurement error is larger for one of these groups, it could confound the comparison. A similar problem exists for other comparisons, including across school districts.

We assess the severity of these concerns by replicating our entire analysis using instrumental variables to correct for measurement error. We estimate our IV models by first specifying the math score as the primary independent variable, with the ELA score serving as the instrument. Next, we duplicate the dataset and flip the positions of the third-grade tests so the ELA score is the independent variable and the math score is the instrument. Then we stack the duplicated dataset and estimate two-stage-least-squares models on the stacked dataset, with standard errors clustered by student to account for the duplication. This procedure yields first-stage error correction coefficients that are averages of the coefficients obtained when each test is individually specified as the independent variable and instrument, respectively.

The IV approach addresses the two main limitations of the EIV approach. First, the IV error correction is not confined to addressing a specific type of measurement error, making it more comprehensive than the EIV correction based on the test reliability ratios. Second, the IV approach allows for subgroup heterogeneity in measurement error—for instance, if FRL students have more error in their test scores than non-FRL students, the IV models for FRL students will make a stronger correction.

But while these theoretical benefits make the IV approach appealing, it also has limitations. The exclusion restriction in the specification where the 3rd grade ELA score is an instrument for the 3rd grade math score relies on the assumption that 3rd grade ELA score is only related to the outcome through its relationship with the general skill component of the 3rd grade math score. However, if the subject specific ELA skill in 3rd grade ( $E_i$  in Equation (6)) also has a direct relationship with the outcome, the exclusion restriction would be violated. This would likely inflate the estimates of  $\beta$  because the subject-specific skill in third grade would be positively related to skill in the future; the same holds in specifications where the 3rd grade math test serves as the instrument for the 3rd grade ELA test. In addition, the IV models are substantially less efficient.<sup>11</sup>

The IV analogs to all results that follow in the main text are reported in Appendix B. At a high level, two themes emerge in comparing the EIV and IV results. First, as expected, the EIV estimates of  $\beta$  are smaller, on average, than the IV estimates, though they are at least 0.80 for the eighth grade and high school tests in all states. Thus, despite providing lower-bound estimates of  $\beta$ , the EIV coefficients show limited mobility. Second, while the mobility parameters differ to some degree across methods, our comparative findings are upheld substantively using either approach. That is, the gaps in academic mobility by student characteristics, and the variance in academic mobility across school districts, are similar. This suggests that while conceptually concerning, in practice the potential for measurement-error heterogeneity to confound our comparisons is limited.

Finally, for completeness, we also estimate models where we set the initial rank using the average of the third-grade math and ELA ranks but do not use the reliability ratios to correct for test measurement error. As expected, our estimates of  $\beta$  are consistently smallest using this approach, although even in these models our comparative findings remain similar. Results using this third approach are shown in

<sup>8</sup> Test measurement error in the post-tests is not of concern because it passes through to the error term (assuming it is uncorrelated with prior test performance and prior measurement error, which is standard).

<sup>9</sup> The applicability of test-reliability ratios when the data undergo a monotonic transformation (in our case, from scale scores to percentile ranks) depends on several factors and is the subject of some debate in the literature (May and Nicewander, 1994; de Gruijter, 1997). May and Nicewander (1994), who examine percentile-transformed data specifically, find that reliability ratios translate poorly only in cases where tests are especially easy or difficult, which is not the case in any sample states during the time period we study.

<sup>10</sup> The reliability ratios for each state test in each subject vary slightly from year-to-year. We use one reliability ratio for each state and subject, which we calculate using the average subject-specific ratio across all cohorts in a state.

<sup>11</sup> In fact, when we estimate the district-level IV models, we must drop small districts from the analysis. See Appendix B for details.

Appendix B in tandem with the IV results.

**2.2.2.2. Geographic mobility.** Some students exit the public school system and/or leave their home states before we observe their later-grade outcomes.<sup>12</sup> Table 3 shows system exiters are negatively selected—i. e., the average entry percentiles of students with missing late-grade outcomes are almost always below those of students with observed outcomes.<sup>13</sup> The sample attrition raises two concerns. The first is reference bias and applies to our analysis of test percentiles, which are normed against the population of test takers. The departure of negatively selected students, if left unaddressed, would lead us to understate upward mobility even if there is no unobserved selection (note the reference bias issue is not relevant to our analysis of graduation outcomes because these outcomes are not normed in the distribution). The second concern is the potential for unobserved selection into system exit conditional on the initial rank. This concern applies to both our test-based and graduation-based mobility metrics and may be problematic for subgroup analyses. For example, suppose system exiters are negatively selected conditional on their initial ranks and district A has a higher proportion of exiters than district B. The differential attrition between districts will cause a compositional difference in their comparison and lead to an overstatement of outcome variance.

We address the reference-bias concern by including students with missing outcomes in our analysis via imputation. Our imputation procedure uses all available test information prior to the missing outcome, up to the seventh grade, to impute test percentiles in eighth grade and high school, and both graduation outcomes.<sup>14</sup> The imputed values allow us to preserve the full entry-cohort distributions in each state, mitigating the concern about reference bias.

To address the concern about unobserved selection, we build hypothetical selection scenarios into the imputation framework. The baseline selection scenario, which we maintain throughout our primary analysis, is that students with missing outcomes are negatively selected on unobservables to the same degree as within-state, cross-district movers. We produce imputed values for students with missing outcomes that embody this condition by relying on observed outcomes for district movers within each state to estimate a “mobility selection parameter.” Using this scenario as an anchor, we consider the sensitivity of our findings to four scenarios where the degree of selection into exit is reparameterized relative to baseline are as follows: (1) 25 percent more negative than baseline, (2) 10 percent more negative than baseline, (3) 10 percent less negative than baseline, and (4) 25 percent less negative than baseline. With the selection-adjusted imputed values, we re-estimate our academic mobility models to determine the sensitivity of our findings to different assumptions about the direction and magnitude of unobserved selection into system exit, above-and-beyond selection into district mobility within the public school system of a state. Full details regarding our imputation procedure are in Appendix C.

This sensitivity analysis shows that none of our findings are substantively affected by the different unobserved-selection conditions we test. This is the combined result of several aspects of outcome missingness in our data: (1) even in the most extreme unobserved selection scenario, and noting that we already capture *observed* selection via early-grade performance, the degree of parameterized negative selection into

<sup>12</sup> For the test outcomes, we also lose students who do not take the tests. The tests are meant to be given to all students, minimizing sample attrition for this reason.

<sup>13</sup> There is one exception in Texas. Note that with an underlying continuous distribution of scores, the mean of each rank distribution should be exactly 50. The mean in several states deviates (very) slightly from 50 because of lumpiness in the underlying test-score distributions.

<sup>14</sup> We additionally make an *ad hoc* correction to the variance of the imputed values to avoid complications due to shrinkage. Appendix C describes our imputation procedure in detail.

exit is modest (based on within-state district movers), (2) although the likelihood of outcome missingness is not evenly distributed across student subgroups or districts, the divergence across subgroups and districts is not extreme, and (3) most students do not exit, limiting the scope for attrition to impact our findings.

Finally, we turn to the issue of geographic mobility within states. We assign students to their third-grade districts, which means our estimates of cross-district variability take on an interpretation akin to “intent-to-treat” parameters. Although most students remain in the same district during grades 3–12, many change districts as well. Some of the changes are structural (e.g., a district that ends after the eighth-grade) or opportunistic (e.g., our data cover a period of growth in the charter sector in many states), although moves surely occur for many other reasons as well.<sup>15</sup> Disentangling the reasons for student mobility across districts, and the implications, is a substantial undertaking and natural extension of this work, but here we focus on understanding differences in academic mobility across districts defined by the district attended in third grade.

### 3. Findings

For presentational convenience we focus the discussion of our findings primarily on simple averages of the state-level results.<sup>16</sup> However, we conduct our entire analysis separately for each state and present many of the state-by-state results alongside the state averages. State-by-state results that are suppressed in the main text are reported in Appendix A.

#### 3.1. Broad patterns of academic mobility at the state level

Table 4 reports estimates of  $\beta$  and  $\bar{O}_{25}$  from equation (1); recall that  $\alpha$  is redundant in the statewide models. Panel A shows the cross-state averages of  $\beta$  and  $\bar{O}_{25}$  using each of our three estimation methods (EIV, IV, and uncorrected). Panel B provides state-by-state results using our preferred EIV approach (Appendix Tables B2a and B2b show state-by-state results using the other methods).

Consistent with prior evidence that early measures of achievement are highly predictive of later outcomes, a student’s rank in the test distribution in the third grade is a strong indicator of later-grade achievement. Using the EIV approach, the cross-state average estimates of  $\beta$  in the eighth-grade and high-school test models are 0.84 and 0.82, respectively (these averages reflect relatively homogeneous estimates across the seven states, as shown in Panel B). As high as these estimates are, they likely understate  $\beta$  because of the incomplete correction for measurement error. The analogous IV estimates, which we interpret as bounding  $\beta$  from above, are 0.90 and 0.92 on average across states. Put plainly, where a student starts in the distribution when tested in the third grade is highly predictive of where they are in the distribution in eighth grade and high school.

Turning to the graduation models, the estimates of  $\beta$  are lower and more variable across states. The simple-average values of  $\beta$  for on-time and lagged graduation are 0.35 and 0.27, respectively, in the EIV models. The IV estimates are again larger, but still reflect a much weaker gradient between initial percentile ranks and the likelihood of graduating. The weaker gradient is visually apparent in the scatterplots in Fig. 1 and Appendix Figure A1, and driven by the fact that graduation rates are high over most of the entry-rank distribution. Put another way, because high school graduation is a fairly indiscriminate outcome, early-

<sup>15</sup> An important structural factor is the size of districts within a state, but even in a small-district state like Missouri, about two-thirds of students remain in the same district for grades 3–12.

<sup>16</sup> We prefer simple averages to weighted averages because Texas contributes so many students and districts compared to the other states. Weighted averages would largely reflect the findings from Texas.

**Table 3**

Documentation of sample attrition in each state and for each late-grade outcome.

		Original Cohort Members					
		Panel Entry		Observed with Outcome		Observed without Outcome	
		N	N	Avg. Outcome Pctl. or Grad Rate	N	Avg. Entry Pctl.	Avg. Imputed Outcome Pctl. or Grad Rate
Grade 8 – Combined Math and ELA	Georgia	376,427	308,624	49.58	67,803	41.21	41.55
	Massachusetts	139,337	124,606	49.41	14,731	46.51	47.99
	Michigan	453,946	395,263	49.41	58,683	39.44	41.83
	Missouri	262,366	227,459	50.69	34,907	47.68	46.63
	Oregon	123,833	105,674	50.44	18,159	45.70	44.07
	Texas	1,280,996	1,094,987	48.73	186,009	49.29	53.68
	Washington	218,051	185,501	49.98	32,550	45.24	45.26
High School Exam	Georgia	376,427	310,207	50.43	66,220	44.96	45.27
	Massachusetts	139,337	114,374	49.31	24,963	46.12	47.23
	Michigan	453,946	346,705	50.40	107,241	39.38	39.45
	Missouri	262,366	205,634	51.23	56,732	42.73	40.53
	Oregon	Not Applicable					
	Texas	1,280,996	1,095,603	50.57	185,393	41.11	44.19
	Washington	218,051	172,229	51.02	45,822	42.71	42.69
Graduation (On-Time)	Georgia	376,427	314,346	80.29	62,081	43.75	69.83
	Massachusetts	139,337	114,413	93.92	24,924	46.13	90.23
	Michigan	453,946	392,186	84.97	61,760	45.13	77.79
	Missouri	262,366	210,423	91.08	51,943	46.10	85.99
	Oregon	123,833	101,692	80.99	22,141	47.43	70.51
	Texas	1,280,996	1,129,684	84.27	151,312	41.59	76.79
	Washington	218,051	176,505	82.66	41,546	43.38	70.33
Graduation (Within One Year of On Time)	Georgia	376,427	314,346	83.76	62,081	43.75	76.21
	Massachusetts	139,337	114,413	94.18	24,924	46.13	90.61
	Michigan	453,946	392,186	87.86	61,760	45.13	81.72
	Missouri	262,366	210,423	93.59	51,943	46.10	89.81
	Oregon	123,833	101,692	82.62	22,141	47.43	72.39
	Texas	1,280,996	1,129,684	87.73	151,312	41.59	82.82
	Washington	218,051	176,505	86.66	41,546	43.38	76.29

Notes: Sample sizes and entry percentiles are based on the average of the grade 3 math and ELA percentiles (i.e., percentiles at entry). For the test outcomes, the mean of each rank distribution should be 50, but in several states it deviates (very) slightly because of lumpiness in the underlying test-score distributions. For graduation outcomes, we report the percent of students who graduate among stayers because percentiles are not informative.

**Table 4**

Statewide estimates of  $\beta$  and  $\bar{O}_{25}$  for each outcome.

Panel A. Average statewide estimates of $\beta$ and $\bar{O}_{25}$ across the seven states, using different measurement error approaches.								
	Grade-8 Test		HS Test		Grad		Grad + 1	
	$\beta$	O25	$\beta$	O25	$\beta$	O25	$\beta$	O25
Cross-State Averages								
EIV	<b>0.84</b>	<b>29.66</b>	<b>0.82</b>	<b>30.44</b>	<b>0.35</b>	<b>75.76</b>	<b>0.27</b>	<b>80.59</b>
IV	<b>0.90</b>	<b>28.28</b>	<b>0.92</b>	<b>28.54</b>	<b>0.41</b>	<b>74.46</b>	<b>0.31</b>	<b>79.58</b>
Uncorrected	<b>0.79</b>	<b>30.99</b>	<b>0.79</b>	<b>31.69</b>	<b>0.33</b>	<b>76.16</b>	<b>0.25</b>	<b>80.88</b>
Panel B. Individual statewide estimates of $\beta$ and $\bar{O}_{25}$ using our preferred measurement error correction (EIV).								
	Grade-8 Test		HS Test		Grad		Grad + 1	
	$\beta$	O25	$\beta$	O25	$\beta$	O25	$\beta$	O25
All (Avg)	<b>0.84</b>	<b>29.66</b>	<b>0.82</b>	<b>30.44</b>	<b>0.35</b>	<b>75.76</b>	<b>0.27</b>	<b>80.59</b>
GA	0.86	29.92	0.86	29.78	0.52	66.03	0.39	73.12
MA	0.84	29.56	0.83	29.65	0.19	88.73	0.19	89.23
MI	0.84	28.48	0.80	29.82	0.35	76.02	0.30	80.22
MO	0.87	28.18	0.82	29.97	0.25	83.61	0.17	88.58
OR	0.81	29.86	Not Applicable		0.33	71.03	0.29	73.63
TX	0.85	31.69	0.80	33.16	0.43	73.77	0.26	81.39
WA	0.82	29.94	0.83	29.68	0.39	71.15	0.29	77.94

Notes: EIV = Errors in Variables Regression, IV = Instrumental Variables Regression, Uncorrected = Uncorrected Linear Regression. The top row of Panel B repeats the average values reported in Panel A using EIV. In these statewide regressions corresponding to equation (1),  $\alpha$  and  $\beta$  are not separately identified. O25 is equal to  $\alpha + 25\beta$ . Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results. All  $\beta$  coefficients are statistically significant; standard errors and statistical significance information suppressed for brevity.

career performance is a weaker predictor of success.

Our estimates of  $\bar{O}_{25}$  for the test outcomes are similar across states and tests, ranging from 28.2 to 33.2, with average values around 30 for each test. The IV analogs in Appendix B are lower—ranging from 27.2 to 29.5. The upper and lower bound estimates of  $\bar{O}_{25}$  thus form a fairly

tight window. The graduation-based  $\bar{O}_{25}$  values, which capture on-time and delayed graduation likelihoods for the average 25th percentile student, are 75.8 and 80.6, respectively, on average across the sample states, and again exhibit more state-to-state variability than their test-based analogs. Like with the test-based estimates, the IV estimates for

graduation-based  $\bar{O}_{25}$  in Appendix B are similar to the EIV estimates but slightly smaller.

The similarity across states in the test-based mobility parameters is partly the result of the distributions of test ranks being forced into alignment by the percentiles conversion. This does not happen with graduation outcomes. Thus, one source of differences in the graduation models are differences in statewide graduation rates. Unsurprisingly, states with higher graduation rates have higher graduation-based  $\bar{O}_{25}$  values. This highlights a source of ambiguity in interpreting our findings with respect to graduation. One interpretation of a high  $O_{25}$  value is that it reflects a state's success in pushing initially low-performing students to graduate. But an alternative interpretation is that it reflects low graduation standards (Costrell, 1994). Unfortunately, our data are ill-suited to distinguish between these interpretations, though when we get to the district-level analysis below, we show districts' test-based and graduation-based mobility metrics are positively correlated ( $\rho \approx 0.3$ – $0.4$ ). This provides some support for the more optimistic interpretation of graduation-based mobility, at least measured at the district level.

### 3.2. Academic mobility for student subgroups within states

In Tables 5, 6, and 7 we report results from versions of Equation (2) where we define student subgroups ( $s$ ) by third-grade racial/ethnic designation, FRL designation, and school urbanicity (urban, suburban, rural). The entry and outcome percentile ranks continue to be normed against the full state distributions. This allows for separate identification of  $\alpha_s$  and  $\beta_s$ , with the tradeoff that we may overstate the academic mobility of higher performing subgroups relative to lower performing subgroups in the presence of uncorrected measurement error (Hanush and Rivkin, 2009). Consequently, we place greater emphasis on the IV estimates in this section because of the more comprehensive treatment of measurement error; any failures of the exclusion restriction will tend to inflate the estimates of  $\beta$ , but we are focusing on the differences between subgroups.

Table 5 and Fig. 2 show results by race/ethnicity. We compare mobility for Asian, Black, Hispanic, and White students.<sup>17</sup> Focusing on  $\bar{O}_{25s}$ —marked by a vertical line at the 25th percentile of the entry distribution in each graph in Fig. 2—we find that initially low-performing Asian students have much higher upward mobility than all other racial-ethnic groups: the average  $\bar{O}_{25s}$  value for the eighth grade test is 39.0 for Asians, 27.1 for Blacks, 29.8 for Hispanics, and 30.7 for Whites. The other panels in Table 5 reveal a similar pattern for the other outcomes. Fig. 2 shows an Asian student advantage in outcomes throughout the distribution of initial ranks via higher baseline mobility (i.e., Asian students have a high value of  $\alpha_s$ ). For test scores this translates to an outcome-rank advantage throughout the entry-rank distribution; for graduation, outcomes converge at higher entry percentiles for all racial-ethnic groups because the graduation likelihood approaches 1.0 for students with high entry percentiles.

Using the IV measurement error correction, the racial-ethnic gaps are smaller but substantively similar (Appendix Table B3a). For example, the average values of  $\bar{O}_{25s}$  for the eighth grade test are 37.5 for Asians, 26.2 for Blacks, 29.1 for Hispanics, and 28.9 for Whites. The gap between Asian and other students remains large despite the more comprehensive measurement-error correction, indicating much higher upward mobility for Asian students.

The results in Table 5 (and Appendix B3a) exemplify the outsized influence of baseline mobility ( $\alpha$ ) in driving variation in absolute upward mobility ( $\bar{O}_{25}$ ) across racial-ethnic groups. The differences in relative mobility ( $\beta$ ) are modest in comparison. For instance, consider

<sup>17</sup> There is also an “other race/ethnicity” category in the data to capture all other students, but it is a small group and omitted from our focal comparisons.

the gap between the group with the highest absolute upward mobility—Asian students—and the group with the lowest upward mobility—Black students—on the eighth-grade test. Based on either the EIV or IV results, 90-plus percent of the Asian-Black  $\bar{O}_{25}$  gap is accounted for by the gap in  $\alpha$  between Asian and Black students, with only a small fraction of the gap remaining to be explained the gap in  $\beta$  (which is multiplied by a factor of 25 to map to  $\bar{O}_{25}$ ). The value of  $\alpha$  is mechanically overstated relative to  $\beta$  by focusing at a point in the distribution below the 50th percentile; still, even evaluated at  $\bar{O}_{50}$ ,  $\alpha$  is the dominant explanatory factor. The primary influence of baseline mobility is a recurring theme throughout our investigation of the variance in absolute upward mobility across student subgroups and school districts.

Our finding of negative Black-White mobility gaps aligns with evidence on the widening of Black-White outcome gaps during K-12 education documented previously (Clotfelter, Ladd, & Vigdor, 2009; McDonough, 2015; Todd & Wolpin, 2007). Our mixed findings for Hispanic-White differences (across outcomes and the EIV and IV models) contribute to a mixed literature. For example, Clotfelter, Ladd, and Vigdor (2009) find the Hispanic-White achievement gap narrows during grades 3–8 in North Carolina. Alternatively, Reardon and Galindo (2009) find the gap is flat from grades 1–5 using a nationally representative sample, and Todd and Wolpin (2007) find it remains flat or widens modestly.<sup>18</sup>

Next, Table 6 and Appendix Table B4a follow the structure of Table 5 but divide students by FRL status. Compared to FRL students, non-FRL students have greater absolute upward mobility. For test scores in eighth grade and high school, the average  $\bar{O}_{25}$  gaps in by FRL status in Table 6 are 4.5 and 6.7 percentage points, respectively. The average on-time and lagged graduation gaps are 12.5 and 11.0 percentage points. Again, the more comprehensive treatment of measurement error via IV reduces the magnitudes of these gaps to some degree, but they remain substantial (see Appendix Table B4a).

The last subgroup comparison is by school urbanicity in the third grade, shown in Table 7. Here there is less heterogeneity across groups. The most notable differences in Table 7 are for graduation outcomes: graduation rates for initially low-performing students who attend urban schools are significantly lower than graduate rates for their peers who attend suburban and rural schools (who have similar graduation rates to each other). These results are again replicated substantively in the IV models in Appendix Table B5a.

### 3.3. District-Level variation in mobility and Cross-Outcome, Cross-Cohort correlations

In this section we estimate the within-state, cross-district standard deviations of  $\alpha$ ,  $\beta$ , and  $\bar{O}_{25}$  to explore the extent to which baseline, relative, and absolute upward mobility vary across school districts. For charter schools, we follow the coding conventions of the states to assign district status. In most cases, charter schools (or their networks in instances of multi-site charters) are coded as separate districts, although a small number of charters are intergraded into larger districts, in which case they are coded as part of the larger district. Note, however, that charter enrollment shares in our cohorts are small; across states they range from 0 to 7.7 percent, with a median value of 1.8 percent (see Appendix Table A1).<sup>19</sup>

The raw variances of  $\hat{\alpha}_d$ ,  $\hat{\beta}_d$ , and  $\hat{\bar{O}}_{25d}$  will overstate the true vari-

<sup>18</sup> A more nuanced explanation of Reardon and Galindo's (2009) findings is that their estimates imply a modest shrinking of the gap in math and a modest increase in reading.

<sup>19</sup> Charter enrollment in the U.S. more than doubled during the timespan over which we track our focal third-grade cohorts (National Center for Education Statistics, 2022), so charter enrollment would be expected to account for a larger share of total enrollment in more recent cohorts.

**Table 5**  
Statewide academic mobility estimates by race/ethnicity.

Student Group: Asian	Grade-8 Test			HS Test			Grad			Grad + 1		
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	18.43	0.82	39.00	19.76	0.81	39.91	79.58	0.19	84.93	84.44	0.16	88.36
GA	19.93	0.80	40.02	16.90	0.82	37.45	66.63	0.36	75.74	74.53	0.27	81.33
MA	18.99	0.82	39.50	20.85	0.80	40.81	91.43	0.10	93.84	92.09	0.09	94.3
MI	16.59	0.85	37.75	14.48	0.89	36.73	83.20	0.18	87.64	87.09	0.14	90.53
MO	15.61	0.87	37.40	18.81	0.81	39.03	86.42	0.14	90.05	90.65	0.10	93.07
OR	14.09	0.84	34.99	Not Applicable			75.88	0.23	81.68	78.51	0.20	83.52
TX	28.08	0.77	47.24	32.82	0.71	50.52	81.68	0.18	86.26	89.17	0.08	91.21
WA	15.75	0.81	36.07	14.68	0.81	34.93	71.82	0.3	79.32	79.06	0.22	84.59
<b>Student Group: Black</b>												
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	7.81	0.77	27.06	8.45	0.76	27.36	62.94	0.39	72.73	71.10	0.29	78.32
GA	8.12	0.79	28.09	7.56	0.81	27.69	51.83	0.57	65.95	63.26	0.41	73.55
MA	9.82	0.75	28.49	10.33	0.71	28.18	80.70	0.21	85.89	81.92	0.19	86.70
MI	8.15	0.71	25.88	5.09	0.71	22.73	62.60	0.39	72.35	68.35	0.34	76.86
MO	5.62	0.78	25.09	8.12	0.75	26.82	70.11	0.32	78.18	79.01	0.21	84.29
OR	5.40	0.76	24.49	Not Applicable			57.60	0.32	65.65	63.53	0.26	70.14
TX	10.25	0.80	30.21	12.53	0.75	31.32	61.20	0.51	73.94	74.91	0.29	82.17
WA	7.29	0.79	27.15	7.05	0.81	27.41	56.51	0.42	67.13	66.74	0.31	74.53
<b>Student Group: Hispanic</b>												
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	10.25	0.78	29.76	9.91	0.77	29.25	65.85	0.36	74.86	73.17	0.27	79.90
GA	11.28	0.82	31.67	9.75	0.82	30.42	51.57	0.53	64.71	62.62	0.38	72.24
MA	9.10	0.74	27.59	8.96	0.71	26.77	76.16	0.27	82.98	77.19	0.26	83.69
MI	8.83	0.80	28.75	6.35	0.82	26.96	66.66	0.32	74.69	72.04	0.28	78.94
MO	9.30	0.82	29.85	12.21	0.77	31.37	76.09	0.25	82.42	83.49	0.16	87.57
OR	11.58	0.73	29.77	Not Applicable			67.25	0.27	74.00	70.50	0.24	76.43
TX	11.47	0.78	31.04	13.41	0.73	31.70	61.13	0.50	73.64	74.39	0.30	81.80
WA	10.16	0.78	29.62	8.79	0.78	28.30	62.07	0.38	71.61	71.93	0.27	78.66
<b>Student Group: White</b>												
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	9.93	0.83	30.69	11.24	0.82	31.66	68.79	0.32	76.82	74.79	0.25	81.08
GA	10.29	0.85	31.53	11.34	0.84	32.42	53.52	0.50	66.06	62.40	0.39	72.25
MA	9.06	0.84	29.95	10.09	0.81	30.38	88.40	0.13	91.72	88.86	0.13	92.04
MI	9.16	0.82	29.69	7.89	0.85	29.10	70.58	0.31	78.30	75.86	0.26	82.34
MO	7.75	0.86	29.24	10.80	0.81	30.94	81.14	0.20	86.21	87.41	0.13	90.63
OR	10.18	0.80	30.24	Not Applicable			61.48	0.35	70.21	64.91	0.31	72.72
TX	13.03	0.84	33.97	17.23	0.78	36.69	64.22	0.38	73.60	72.73	0.26	79.26
WA	10.02	0.81	30.20	10.08	0.81	30.45	62.16	0.38	71.61	71.35	0.28	78.31

Notes: These estimates are from mobility regressions estimated separately for each racial-ethnic student group in each state, as shown in equation (2). O25 is equal to  $\alpha + 25\beta$ . Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results. All  $\beta$  coefficients are statistically significant; standard errors and statistical significance information suppressed for brevity.

ances due to sampling variance. We estimate the sampling variance of each parameter by the average of the squared standard errors (Aaronson, Barrow, and Sander, 2007), which we subtract from the total variance to estimate the true variance.<sup>20</sup> Error-corrected standard deviations of the mobility parameters are shown in Tables 8 and B6a for the EIV and IV estimates, respectively. The variances across districts for all parameters, all outcomes, and in all states are statistically significant. Fig. 3 shows the distributions of  $\hat{O}_{25d}$  in two example states, Missouri and Washington, for all outcomes. The distributions are consistently unimodal and smooth, ruling out odd patterns of heterogeneity across districts.

On average across states, the EIV results in Table 8 indicate that one standard deviation in the distribution of absolute upward mobility ( $\hat{O}_{25}$ ) corresponds to a change in student rank on the eighth-grade and high-school tests of about 4.8 percentile points. For on-time and delayed graduation, the analogous average standard deviations are 5.5 and 4.9

<sup>20</sup> The standard errors of  $\hat{O}_{25d}$  are linear combinations of the standard errors of  $\hat{\alpha}_d$  and  $\hat{\beta}_d$ .

percentage points, respectively. The standard deviations of the IV-based estimates are similar. Adding context from Table 4, the estimates in Table 8 indicate that a third-grade entrant at the 25th percentile who attends a district with academic mobility one-standard-deviation above average would be expected to score at the 35.2 percentile in the state distribution of the high school test, compared to the 30.4 percentile at the average district. In terms of on-time graduation, a similar comparison yields a graduation likelihood at the high-mobility district of 81.3 percent, versus 75.8 percent at the average district.

It is also of interest to compare the importance of  $\alpha_d$  and  $\beta_d$  in driving upward mobility. Although separable inference is challenging because  $\alpha_d$  and  $\beta_d$  are negatively correlated within districts, on average; there is ample variation in the data to separately identify the magnitude of variation in both parameters across districts. The comparison is complicated because the importance of  $\alpha_d$  and  $\beta_d$  varies by the initial rank—i.e., at low initial ranks, variation in  $\alpha_d$  will be a more important driver of upward mobility but as the initial rank increases,  $\beta_d$  becomes more important. This dynamic is illustrated in Appendix Figure A2. The key takeaway from the figure is that over most of the distribution of initial ranks, and certainly at lower-valued ranks, variation in  $\alpha$  is the

**Table 6**  
Statewide academic mobility estimates by FRL status.

Student Group: FRL	Grade-8 Test			HS Test			Grad			Grad + 1		
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	9.03	0.76	28.11	8.61	0.75	27.47	62.32	0.35	71.15	69.94	0.26	76.48
GA	8.78	0.79	28.61	7.96	0.79	27.86	48.99	0.53	62.45	60.24	0.39	70.02
MA	8.92	0.73	27.09	9.04	0.69	26.40	77.13	0.23	82.8	78.11	0.22	83.49
MI	8.31	0.75	26.98	5.55	0.78	24.99	60.89	0.34	69.37	66.69	0.30	74.17
MO	6.47	0.81	26.79	8.46	0.76	27.52	73.00	0.26	79.62	81.12	0.17	85.37
OR	10.54	0.72	28.56	Not Applicable			60.12	0.25	66.30	63.83	0.21	69.09
TX	11.06	0.77	30.34	12.74	0.72	30.65	59.14	0.50	71.55	72.63	0.29	79.81
WA	9.13	0.77	28.42	7.89	0.78	27.37	56.96	0.36	65.99	66.97	0.26	73.44
<b>Student Group: non-FRL</b>												
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	12.06	0.82	32.59	14.28	0.80	34.18	77.99	0.23	83.64	83.39	0.16	87.45
GA	13.28	0.82	33.96	15.22	0.81	35.47	67.46	0.35	76.27	75.61	0.25	81.95
MA	11.39	0.82	31.90	12.89	0.79	32.71	92.94	0.08	94.87	93.35	0.07	95.16
MI	10.28	0.82	30.86	8.88	0.85	30.15	79.91	0.21	85.14	84.36	0.17	88.53
MO	9.29	0.86	30.71	14.50	0.78	34.02	86.76	0.15	90.43	91.91	0.08	94.04
OR	12.48	0.80	32.50	Not Applicable			73.42	0.24	79.47	76.50	0.21	81.72
TX	15.10	0.83	35.80	20.48	0.76	39.41	73.39	0.28	80.40	81.97	0.16	86.06
WA	12.59	0.79	32.41	13.68	0.79	33.34	72.08	0.27	78.93	80.00	0.19	84.70

Notes: These estimates are from mobility regressions estimated separately for each FRL student group in each state, as shown in equation (2). O25 is equal to  $\alpha + 25^*\beta$ . Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results. All  $\beta$  coefficients are statistically significant; standard errors and statistical significance information suppressed for brevity.

**Table 7**  
Statewide academic mobility estimates by the urbanicity of the school district in the third grade.

Student Group: Urban	Grade-8 Test			HS Test			Grad			Grad + 1		
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	7.97	0.83	28.73	7.65	0.82	28.22	62.14	0.40	72.24	69.76	0.31	77.53
GA	5.84	0.81	26.34	6.29	0.83	26.97	47.00	0.56	61.20	57.59	0.44	68.50
MA	8.40	0.81	28.53	8.37	0.78	27.97	77.07	0.26	83.62	78.19	0.25	84.38
MI	7.02	0.84	27.99	2.69	0.89	24.91	62.62	0.41	72.90	68.40	0.36	77.31
MO	4.83	0.85	26.17	7.38	0.81	27.54	68.86	0.33	77.11	77.49	0.23	83.17
OR	10.06	0.82	30.66	Not Applicable			61.45	0.34	70.00	65.68	0.30	73.08
TX	10.26	0.84	31.29	12.66	0.80	32.62	60.36	0.48	72.31	73.58	0.28	80.67
WA	9.38	0.83	30.12	7.90	0.85	29.15	57.60	0.44	68.51	67.40	0.33	75.61
<b>Student Group: Suburban</b>												
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	9.11	0.85	30.32	10.30	0.84	31.18	68.75	0.34	77.17	75.49	0.25	81.87
GA	9.10	0.87	30.91	9.39	0.87	31.24	52.35	0.53	65.72	62.36	0.41	72.64
MA	8.91	0.85	30.07	9.76	0.83	30.38	86.81	0.16	90.75	87.36	0.15	91.13
MI	8.21	0.83	29.04	6.69	0.86	28.25	70.82	0.32	78.82	76.29	0.26	82.91
MO	6.29	0.89	28.56	11.55	0.82	32.12	78.43	0.24	84.42	85.35	0.16	89.27
OR	10.08	0.82	30.65	Not Applicable			63.78	0.34	72.34	67.26	0.30	74.85
TX	11.41	0.85	32.68	14.47	0.80	34.48	66.23	0.39	75.96	77.77	0.22	83.38
WA	9.76	0.82	30.31	9.94	0.83	30.58	62.81	0.38	72.21	72.05	0.28	78.93
<b>Student Group: Rural</b>												
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	9.22	0.82	29.77	9.55	0.82	29.99	69.52	0.31	77.43	75.94	0.24	81.90
GA	9.04	0.84	30.23	8.28	0.85	29.64	55.57	0.49	67.84	65.98	0.35	74.89
MA	9.13	0.83	29.78	9.89	0.81	30.15	88.37	0.13	91.60	88.89	0.12	91.96
MI	9.53	0.80	29.46	7.42	0.83	28.28	70.30	0.31	78.07	75.38	0.27	82.03
MO	7.68	0.85	28.99	10.01	0.80	30.00	81.42	0.20	86.50	87.83	0.13	90.98
OR	9.72	0.78	29.10	Not Applicable			63.46	0.31	71.26	66.54	0.28	73.54
TX	10.23	0.85	31.48	13.15	0.79	32.88	64.26	0.40	74.33	74.59	0.25	80.90
WA	9.21	0.81	29.36	8.53	0.82	29.00	63.28	0.36	72.39	72.40	0.26	78.97

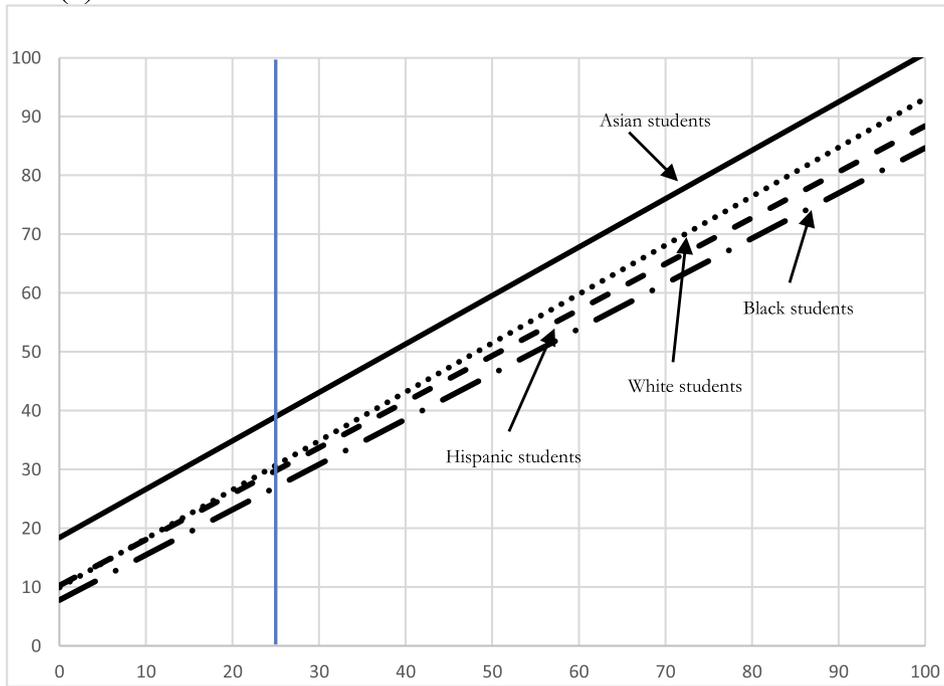
Notes: These estimates are from mobility regressions estimated separately for each urbanicity student group in each state, as shown in equation (2). O25 is equal to  $\alpha + 25^*\beta$ . Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results. All  $\beta$  coefficients are statistically significant; standard errors and statistical significance information suppressed for brevity.

primary driver of variation in upward academic mobility. This only changes at very high levels of the initial outcome percentile—the crossover point is at approximately the 83rd percentile, where

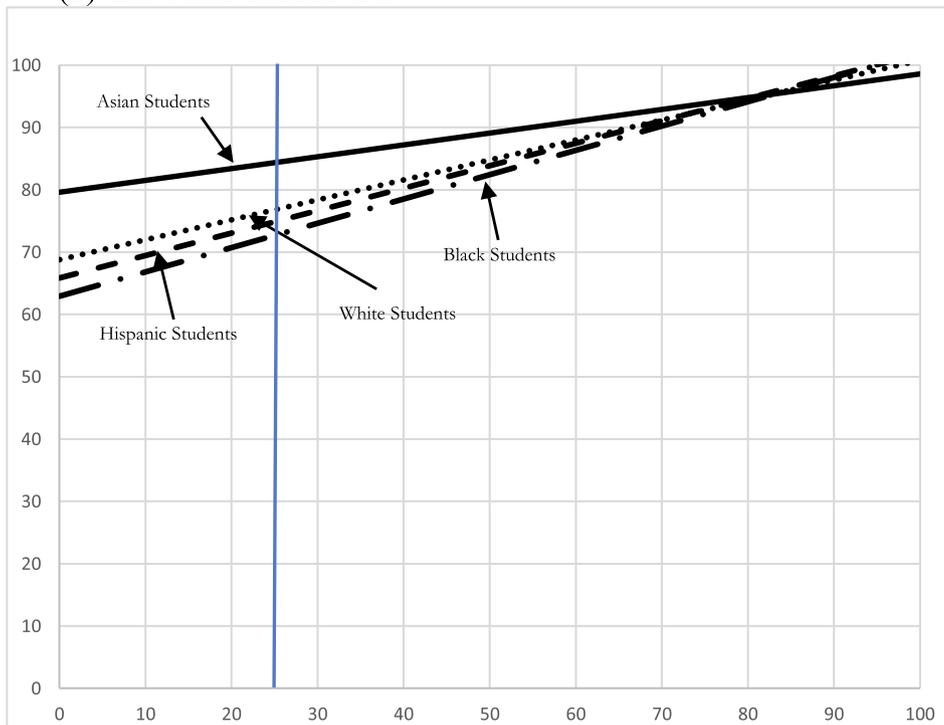
$$\sigma_{\beta^*p} \approx \sigma_{\alpha}.$$

This result is another example of the consistent theme of our analysis that there is more variation in academic-mobility intercepts than slopes,

(a) 8<sup>th</sup>-Grade Test Scores



(b) On-Time Graduation



**Fig. 2.** Illustrations of the linearly estimated rank-rank relationships for 8th-grade test score and on-time graduation outcomes by race-ethnicity, corresponding to the results in Table 5. On average across states. Notes: These graphs illustrate the cross-state averages of the linearly-estimated mobility relationships by race-ethnicity between the 3<sup>rd</sup> grade test rank and (a) the 8<sup>th</sup>-grade test rank and (b) on-time high school graduation. The linear-model parameters for each race-ethnicity (and state) and outcome are shown in Table 5.

in this case across districts. This suggests the potential is greater for districts to improve the outcomes of initially low-performing students through overall improvement rather than by differentially impacting students at different points in the entry distribution. We are mindful in this interpretation that our estimates are not causal, but the ratio of the

variances of  $\alpha_d$  and  $\beta_d$  is suggestive of how districts are likely to affect the trajectories of low performers absent reforms to current practice.

Finally, Tables 9 and 10 show correlations between district estimates of absolute upward mobility across outcomes and cohorts. We adjust for estimation error in the correlation between any two sets of estimates by

**Table 8**  
Estimates of the of the within-state, cross-district standard deviations of  $\bar{O}_{25d}$ ,  $\alpha_d$ , and  $\beta_d$ .

Standard Deviations	Grade-8 Test			HS Test			Grad			Grad + 1		
	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25	$\alpha$	$\beta$	O25
All (Avg)	4.89	0.06	4.79	5.02	0.06	4.84	7.31	0.09	5.51	6.34	0.07	4.86
GA	3.65	0.04	3.54	3.83	0.05	3.72	7.70	0.10	5.59	6.88	0.07	5.14
MA	6.32	0.05	5.91	6.38	0.05	5.96	5.61	0.07	4.00	5.39	0.07	3.84
MI	4.07	0.06	4.06	5.10	0.07	4.98	7.79	0.08	6.38	7.15	0.07	5.69
MO	4.11	0.06	4.33	4.33	0.06	4.28	5.70	0.06	4.18	4.26	0.04	3.20
OR	7.21	0.07	6.62	Not Applicable			9.37	0.10	7.19	9.01	0.10	6.94
TX	5.56	0.06	5.14	5.82	0.06	5.47	7.34	0.12	5.03	5.62	0.08	4.26
WA	3.28	0.06	3.93	4.67	0.06	4.60	7.69	0.09	6.19	6.05	0.08	4.92

Notes: These standard deviations are for the parameters estimated from equation (3) for each district in each state, adjusted for estimation error variance using the average of the squared standard errors of the mobility parameters. Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results.

first estimating the ratio of the true variance to the total variance for each set of estimates. Then we multiply the raw correlation by the inverse of the square root of the product of the ratios, following Spearman (1904). As noted by Kraft (2017), this procedure generates what are best interpreted as upper-bound correlations because it assumes all estimation error is uncorrelated. We also show unadjusted correlations that provide complementary lower bounds. For ease of presentation, we focus on the adjusted correlations in our discussion, and for brevity we show the average values of the correlations across states in the tables. The state-by-state results are reported in Appendix Tables A2 and A3.

Table 9 shows that the mobility metrics are positively correlated across outcomes within districts, on average. The error-adjusted, upper-bound correlations within outcome mode are very high—for test outcomes the average correlation across states is 0.87, and for graduation outcomes it is 1.0. The adjusted correlations across outcome modes are positive but lower, ranging from 0.37 to 0.40.

Table 10 shows analogous correlations within districts and outcomes, but across cohorts. The states that contribute to the average in each cell depend on the cohorts included in the state samples (per Table 1). The contiguous-cohort adjusted correlations are between 0.59 and 0.73 on average, and somewhat larger for test-based mobility than graduation-based mobility. The adjusted correlations for cohorts two- and three-years removed are mostly smaller but still consistently positive, and no adjusted correlation across any cohorts for any outcome is below 0.42.

#### 4. Correlates of academic mobility

##### 4.1. Primary correlates

Next we explore links between academic mobility and the attributes of districts and their local areas. We assemble a database of district and local-area attributes from two sources: (1) our administrative education databases and (2) externally geocoded data from the National Center for Education Statistics (NCES). Using the administrative data, we construct variables for the percentages of students in each district who are (a) Black, (b) Hispanic, (c) FRL enrolled, (d) participants in an individualized education plan (IEP), and (e) geographically mobile. Following CHKS, we also construct a Theil index that captures within-district segregation by race-ethnicity (measured by the segregation of under-represented minority students, who we define as Black and Hispanic),

and a parallel segregation index based on economic status (measured by FRL enrollment) motivated by recent research on economic connectedness (Chetty et al., 2022).<sup>21</sup> All these variables are constructed using data from students in our cohorts in the third grade.

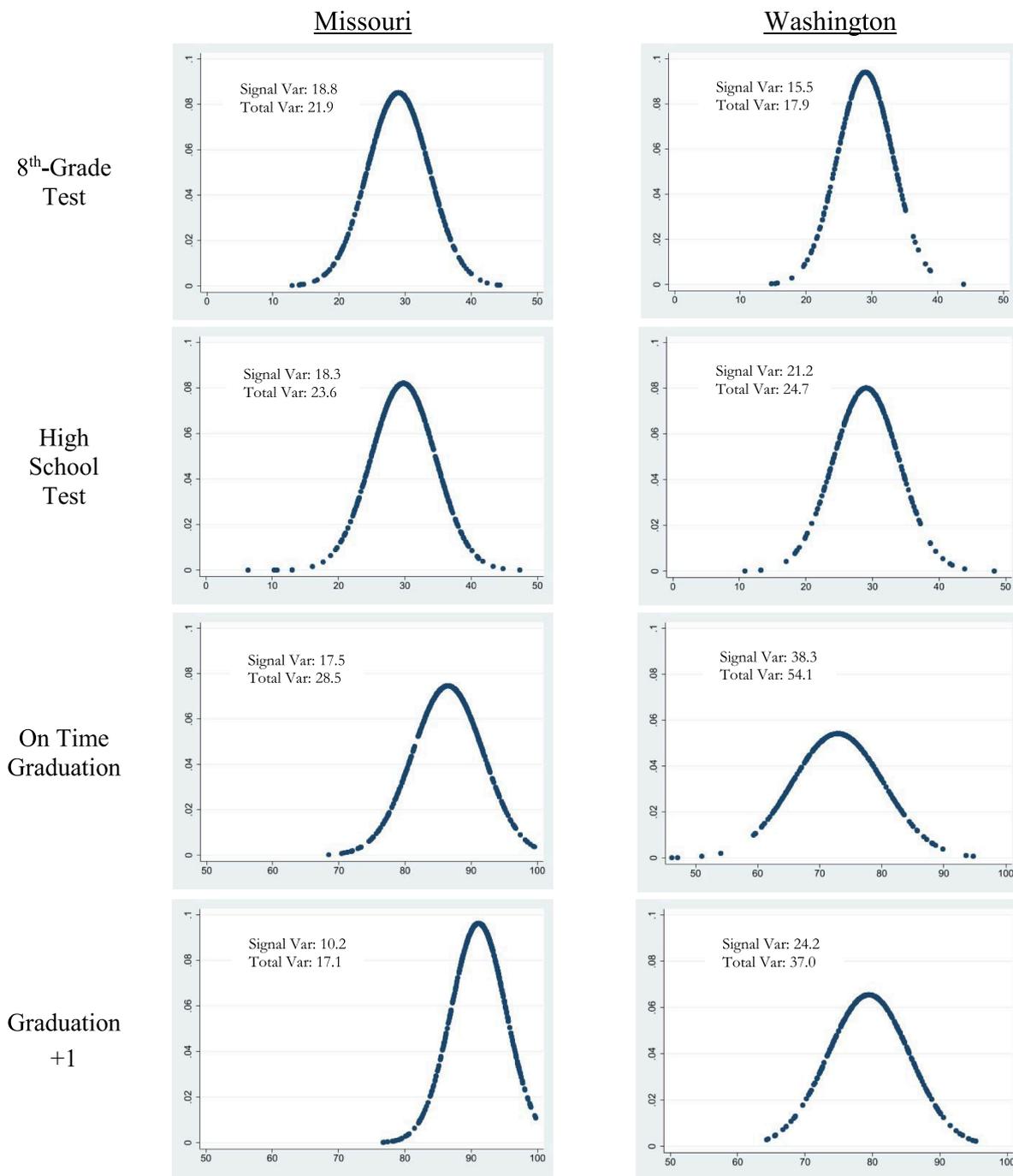
An additional district attribute we construct using our administrative data is value added to student test scores in math and ELA in grades 4–8. Our value-added estimates capture district contributions to student test score growth in both subjects. We estimate value added using data from the same time periods during which we follow the cohorts in each state but jackknife the estimates around our cohorts to remove any mechanical correlation between academic mobility and value added. We construct the value-added estimates to be uncorrelated with student characteristics following Parsons, Koedel, and Tan (2019). Finally, we estimate value-added separately for above- and below-median students based on lagged test scores. Appendix D provides estimation details for the value-added models.

We also correlate academic mobility with local-area attributes geocoded to districts’ catchment areas based on data from the American Community Survey (ACS), made available by the Education Demographic and Geographic Estimates (EDGE) program of NCES. We include variables that capture local-area median household income and the poverty rate, along with the percent of families with school-aged children where the head of household is identified as (a) Black, (b) Hispanic, (c) a high school graduate, (d) a college graduate, (e) speaking a language other than English at home, (f) residentially stable, and (g) never married. These variables are from the population of parents of school-aged children in districts’ catchment areas. Finally, we correlate academic mobility with district per-pupil expenditures taken from the District Finance Survey, also from the NCES.<sup>22</sup>

Fig. 4 shows coefficients from univariate regressions of  $\bar{O}_{25d}$ —estimated for each of the four long-term outcomes—on the district and local-area attributes. We report average coefficients across the seven states for presentational convenience. The independent variables are standardized within each state to have a mean of zero and variance

<sup>21</sup> The Theil indices measure the degree of racial/ethnic or economic segregation in a district. Values range from 0 (where all schools within a district have the same racial/ethnic or economic composition as the district as a whole) to 1 (where racial/ethnic or economic groups are entirely segregated between schools within a district). We drop districts with only one school because the Theil index is undefined.

<sup>22</sup> We use NCES data from the 2010–14 period to construct these variables. The EDGE data are only available over selected periods. Among the available options, the 2010–14 period provides the most overlap with the years over which we estimate academic mobility for the sample cohorts.



**Fig. 3.** Distributions of estimated  $\bar{O}_{25d}$  in two states, Missouri and Washington, for all outcomes. These distributions are visual complements to the results in Table 8. Notes: Distributions of estimated  $\bar{O}_{25d}$  values. The distributions are reported as estimated directly without any adjustments for sampling variance. Each graph reports the total variance and signal variance, where the latter is the total variance minus the sampling variance, to indicate the portion of the variation driven by sampling error.

of one—therefore, the coefficient averages reflect the predicted change in academic mobility associated with a one-standard-deviation move in the distribution of the independent variable, on average across states.<sup>23</sup> Detailed state-by-state regression output underlying Fig. 4 is available in

<sup>23</sup> The value-added measures are shrunk using the approach of Lefgren and Sims (2012), per Appendix D. This means a one-standard-deviation change in the raw data corresponds to more than a one-standard-deviation change in the true (unobserved) distribution of value added (Chetty, Friedman, and Rockoff, 2014a).

Appendix Table A4. Results using IV-based estimates of  $\bar{O}_{25d}$  as dependent variables in place of the EIV-based estimates are broadly similar and reported in Appendix Table B9a.<sup>24</sup>

The preceding analysis offers some predictions about our findings.

<sup>24</sup> We note two caveats to the similarity: (1) the district-level IV regressions are especially noisy in some cases due to the reduced efficiency of the IV models, and (2) it is only feasible to estimate the district-level mobility regressions using IV for a subsample of larger districts. See the discussion in Appendix B for more information.

**Table 9**  
Adjusted and unadjusted correlations of  $\bar{O}_{25d}$  across outcomes, on average across states.

	Adjusted Correlations					Unadjusted Correlations			
	Grade-8 test	HS test	Grad	Grad + 1		Grade-8 test	HS test	Grad	Grad + 1
Grade-8 test	1.00				Grade-8 test	1.00			
HS test	0.87	1.00			HS test	0.76	1.00		
Grad	0.38	0.40	1.00		Grad	0.31	0.31	1.00	
Grad + 1	0.38	0.37	1.00	1.00	Grad + 1	0.30	0.29	0.92	1.00

Notes: Correlation matrix entries are simple, cross-state averages of the correlations of the district-level mobility parameters across outcomes. State-by-state correlation matrices are reported in [Appendix Table A2](#). The adjusted correlations are best interpreted as upper bounds because they assume estimation error in  $\bar{O}_{25d}$  is uncorrelated across outcomes despite the fact that the same student sample is used; the unadjusted correlations make no adjustments for estimation error. Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results.

**Table 10**  
Adjusted and unadjusted correlations of  $\bar{O}_{25d}$  across cohorts for each outcome, on average across states. Grade-8 Test.

	Adjusted Correlations					Unadjusted Correlations			
	2006	2007	2008	2009		2006	2007	2008	2009
2006	1				2006	1.00			
2007	0.67	1			2007	0.57	1.00		
2008	0.59	0.69	1		2008	0.49	0.60	1.00	
2009	0.49	0.63	0.73	1	2009	0.40	0.53	0.61	1.00
<b>HS Test</b>									
	Adjusted Correlations					Unadjusted Correlations			
	2006	2007	2008	2009		2006	2007	2008	2009
2006	1				2006	1.00			
2007	0.70	1			2007	0.52	1.00		
2008	0.67	0.70	1		2008	0.50	0.58	1.00	
2009	0.61	0.63	0.72	1	2009	0.45	0.50	0.58	1.00
<b>Graduation</b>									
	Adjusted Correlations					Unadjusted Correlations			
	2006	2007	2008	2009		2006	2007	2008	2009
2006	1.00				2006	1.00			
2007	0.59	1.00			2007	0.44	1.00		
2008	0.48	0.59	1.00		2008	0.37	0.44	1.00	
2009	0.56	0.62	0.68	1.00	2009	0.44	0.48	0.50	1.00
<b>Graduation + 1</b>									
	Adjusted Correlations					Unadjusted Correlations			
	2006	2007	2008	2009		2006	2007	2008	2009
2006	1.00				2006	1.00			
2007	0.59	1.00			2007	0.43	1.00		
2008	0.42	0.61	1.00		2008	0.31	0.44	1.00	
2009	0.56	0.64	0.66	1.00	2009	0.43	0.48	0.48	1.00

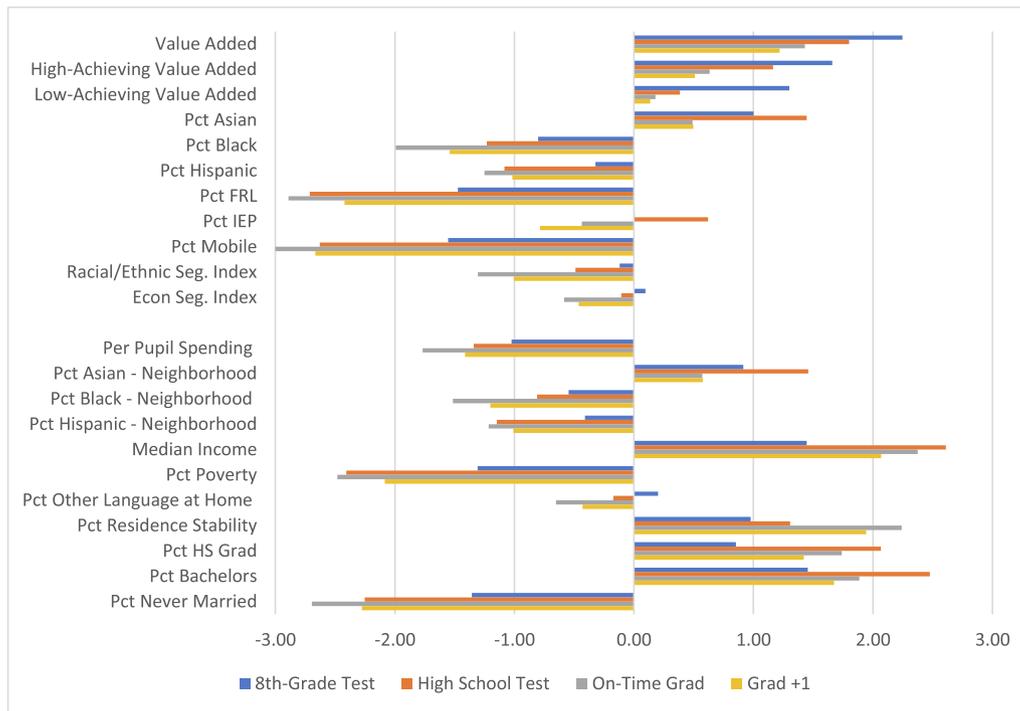
Notes: Correlation matrix entries are simple, cross-state averages of the correlations of the district-level mobility parameters within outcomes and across cohorts. State-by-state correlation matrices are reported in [Appendix Table A3](#). The adjusted correlations are best interpreted as upper bounds because they assume estimation error in the estimates of  $\bar{O}_{25d}$  is uncorrelated across cohorts; the unadjusted correlations make no adjustments for estimation error. Not all states contribute to the averages in all cells in this table—which states contribute depends on the year-cohorts available in each state (see [Table 1](#)). Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results.

For example, the student-level differences in academic mobility by race/ethnicity and FRL status in [Tables 5 and 6](#) are reflected in the district-level relationships in [Fig. 4](#). More broadly, [Fig. 4](#) shows absolute upward mobility is higher in socioeconomically advantaged areas. Some of the strongest predictors of high absolute upward mobility are: lower shares of underrepresented minorities in districts and their local areas, higher local-area incomes and education levels, less school mobility and greater residential stability, fewer never-married parents, and value-added to student achievement. The value-added associations are somewhat stronger in the regressions of test-based academic mobility, which

is not surprising, but are also positive for the graduation-based mobility metrics.<sup>25</sup>

Among the more interesting correlates for which we do not identify a positive relationship with academic mobility is district per-pupil spending. In fact, in most states and for most outcomes, higher per-pupil spending is associated with lower academic mobility (see [Appendix Table A4](#)). The relationships in [Fig. 4](#) are not causal and there are many potential explanations. In this instance, possibilities include redistributive spending that targets disadvantaged children such as Title I, compensating differentials for educators to account for more

<sup>25</sup> The average coefficients on value added in the graduation-based  $\bar{O}_{25}$  models are buoyed by particularly large estimates in Michigan. Still, the coefficients in all states are positive and many are statistically significant. See [Appendix B](#).



**Fig. 4.** Correlates of  $\bar{O}_{25d}$  for each outcome, on average across states. Notes: The bars represent cross-state averages of coefficient estimates from univariate regressions of  $\bar{O}_{25d}$  for each outcome on a wide range of district and local-area attributes. All independent variables are standardized so that the interpretation of each coefficient is in standard-deviation units. Simple average values of the coefficients across states are reported. The top vertical panel shows correlates constructed using the state administrative education datasets. The bottom vertical panel shows correlates taken from the NCES (either from the EDGE program based on data for parents with school-aged children, or the District Finance Survey, as described in the text). The state-by-state results underlying this figure are reported in [Appendix Table A4](#), which includes information on the statistical significance of individual coefficients in individual states.

challenging working conditions, and inefficient resource use.

4.2. Extensions

We conduct three extensions of the analysis of correlates. First, we replicate the correlational analysis but use as dependent variables student-type-specific estimates of absolute upward mobility, estimated separately in each district for students who are Black, Hispanic, and FRL-enrolled. This is to assess whether the predictors of higher academic mobility overall also predict higher academic mobility among students in these at-risk groups. A limitation is that the subgroup specific estimates of academic mobility are less precisely estimated—sometimes by a considerable margin depending on the composition of a district. This should not cause bias because  $\bar{O}_{25d}$  is the dependent variable, but it does reduce efficiency and lower the precision with which we can identify some relationships. Noting this limitation, we generally find the attributes that predict higher academic mobility overall also predict higher mobility for at-risk students. Figures analogous to [Fig. 4](#) for each student

type are in [Appendix Figures A3-A5](#).

Second, we correlate district value-added separately with  $\alpha_d$  and  $\beta_d$ , instead of  $\bar{O}_{25d}$ , to assess whether districts with high value added promote greater convergence in student outcomes. [Table 11](#) shows these results. We do not find consistent evidence of a relationship between district value-added and  $\beta_d$  (the association is small and inconsistently signed across states and outcomes). In contrast, the associations between district value-added and  $\alpha_d$  are overwhelmingly positive. This suggests the correlations between value added and  $\bar{O}_{25d}$  in [Fig. 4](#) are driven primarily by variation across districts in  $\alpha_d$ , not  $\beta_d$ , further reinforcing our findings on the relative importance of slopes and intercepts in driving variation in absolute upward mobility across districts.

Third, we aggregate our estimates of academic mobility up to the commuting-zone and county levels in order to correlate them to external estimates of intergenerational economic mobility from CHKS (2014) and [Chetty and Hendren \(2018\)](#), respectively. Details for this portion of our analysis—including discussion of a number of challenges associated

**Table 11**  
Correlations of district-level value added to student achievement with  $\alpha_d$  and  $\beta_d$ .

	Grade-8 Test		HS Test		Grad		Grad + 1	
	$\alpha$	$\beta$	$\alpha$	$\beta$	$\alpha$	$\beta$	$\alpha$	$\beta$
All (Avg)	0.35	0.05	0.21	0.02	0.14	-0.10	0.14	-0.10
GA	0.29	-0.14	0.15	-0.18	0.00	0.05	-0.01	0.06
MA	0.28	0.21	0.23	0.17	0.17	-0.14	0.17	-0.15
MI	0.40	0.15	0.33	0.18	0.22	-0.20	0.21	-0.20
MO	0.34	0.05	0.19	0.01	0.14	-0.16	0.18	-0.20
OR	0.47	-0.04	Not Applicable		0.13	-0.02	0.08	0.03
TX	0.32	-0.03	0.23	-0.10	0.12	-0.07	0.16	-0.12
WA	0.32	0.16	0.11	0.06	0.21	-0.15	0.18	-0.13

Notes: See [Appendix D](#) for information about our procedure for estimating value added for each district. Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results.

with linking our estimates of intragenerational academic mobility to their estimates of intergenerational economic mobility—are provided in [Appendix E](#). A high-level summary of the results is as follows. There is insufficient cross-commuting-zone variance in academic mobility to account for observed variance in economic mobility at this same level of geography (also see [Rothstein, 2019](#)).<sup>26</sup> A key factor contributing to this result is that most of the variance in academic mobility across school districts occurs within, and not between, commuting zones.<sup>27</sup> Between counties there is more variation in academic mobility because counties typically cover much smaller geographic areas. In [Appendix E](#), we show our estimates of academic mobility are positively correlated with Chetty and Hendren's economic mobility estimates at the county level. While we are hesitant to draw strong conclusions from the correlations, they at least allow for the possibility of a substantive link between academic and economic mobility.

## 5. Conclusion

We introduce the concept of “academic mobility” and use it to study the distributional stickiness of student performance during K-12 schooling. On the whole, we find that academic mobility in the education system is limited—students' ranks in the academic performance distribution in the third grade are highly predictive of their ranks in higher grades. However, we also estimate statistically significant and educationally meaningful differences in academic mobility across school districts. Initially low-performing students who attend districts one standard deviation higher in the academic mobility distribution perform about 5 percentile points higher on tests in the eighth grade and high school relative to their peers who attend districts with average mobility. They are also 5–6 percentage points more likely to graduate from high school.

Our analysis of academic mobility across student groups divided by race-ethnicity, eligibility for free and reduced-price lunch, and district urbanicity produces patterns that are largely as expected based on existing research. Still, some results stand out. One is the large and consistent upward mobility advantage among Asian students relative to all other racial/ethnic groups. Another is that initially low-performing students in rural districts have broadly similar upward mobility to their suburban peers, which is at odds with the prevailing theme of the “rural schools problem” in education research ([Burton et al., 2013](#)).

When we decompose total academic mobility into its components and examine cross-district heterogeneity, we find cross-district differences in baseline mobility are the primary driver of cross-district variance in total academic mobility. This suggests low-performing students experience the largest performance gains when attending districts where students generally excel. It also casts doubt on the narrative that districts vary substantially in the degree to which they narrow within-district achievement gaps as students progress through school.

We correlate absolute upward mobility with a wide array of district and local-area characteristics. We find that absolute upward mobility is largest in socioeconomically advantaged areas as measured along a variety of dimensions. We also show districts with high value added to student test scores have significantly higher upward mobility (as

<sup>26</sup> [Rothstein \(2019\)](#) shows differences in the relationship between parental income and children's human capital across commuting zones—a metric akin to academic mobility, although Rothstein's data and methods differ substantially from our own—can account for only a small fraction of the cross-commuting-zone variance in economic mobility documented by CHKS.

<sup>27</sup> This result is in line with recent, related place-based work by [Schoefer and Ziv \(2021\)](#), who show that most of the measured variance in productivity across cities is driven by plant-level productivity differences. In the education context, [Laliberte \(2021\)](#) finds differences in schools are an important driver of place-based effects on students' educational attainment using narrowly-defined geographic areas.

measured by test and non-test outcomes).

Finally, we briefly consider the potential for differences in academic mobility to explain geographic variation in economic mobility across commuting zones and counties. Variation in academic mobility cannot explain a meaningful fraction of the variance in economic mobility across commuting zones documented by CHKS (2014), corroborating related findings from [Rothstein \(2019\)](#). There is much more variation in academic mobility at the county level, and we find that county-level estimates of academic and economic mobility are positively correlated.<sup>28</sup>

It bears repeating that our academic mobility metrics do not carry a causal interpretation. We do not know if our estimates reflect the true impacts of the local areas we define by school districts, or something else like the selection of families ([Bruhn, 2020; Chyn and Katz, 2021](#)). Moreover, if we overcome this hurdle and can recover causal estimates of these areas—an objective we intend to pursue in future research—it will still be difficult to assess what it is about them that drives the findings (inclusive of factors inside and outside of schools). These are problems endemic to the burgeoning field of place-based research ([Chetty et al., 2020; Harding et al., 2021; Kaestner, 2020](#)). Noting this important caveat, our findings illuminate broad patterns in academic mobility and suggest directions for future research that will create a body of evidence to guide the development of policies that support academic mobility.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The authors do not have permission to share data.

## Acknowledgements

We gratefully acknowledge financial support from CALDER, which is funded by a consortium of foundations (for more information about CALDER funders, see [www.caldercenter.org/about-calder](http://www.caldercenter.org/about-calder)). All opinions expressed in this paper are those of the authors and do not necessarily reflect the views of our funders, data providers, or the institutions to which the author(s) are affiliated, and all errors are our own. This work would not have been possible without the support of our state partners. We make the following acknowledgments and disclaimers regarding the provision of data:

Georgia: The contents of this report were developed using data provided by Georgia's Academic and Workforce Analysis and Research Data System (GA•AWARDS). However, those contents do not necessarily represent the policy of GA•AWARDS or any of its participating organizations, and you should not assume endorsement by GA•AWARDS or any of its participating organizations. We acknowledge valuable research assistance from Henry Woodward.

Massachusetts: The authors wish to thank partners at the Massachusetts Department of Elementary and Secondary Education for the provision of data to support this work.

Michigan: This research result used data structured and maintained by the MERI-Michigan Education Data Center (MEDC). MEDC data is modified for analysis purposes using rules governed by MEDC and are not identical to those data collected and maintained by the Michigan

<sup>28</sup> As shown by [Biasi \(2023\)](#), even if variation in school quality explains little of the variation in economic mobility across geographies, school-based policies can still impact economic mobility (in particular, Biasi shows that state school finance equalization policies promote intergenerational economic mobility).

Department of Education (MDE) and/or Michigan's Center for Educational Performance and Information (CEPI). Results, information and opinions solely represent the analysis, information and opinions of the author(s) and are not endorsed by, or reflect the views or positions of, grantors, MDE and CEPI or any employee thereof. We acknowledge valuable research assistance from Dongming Yu, Katie Bollman, and Amy Auletto.

Missouri: The authors wish to thank partners at the Missouri Department of Elementary and Secondary Education for the provision of data to support this work. We acknowledge valuable research assistance from Kyung Seong-Jeon.

Oregon: The authors wish to thank partners at the Oregon Department of Education for the provision of data to support this work.

Texas: The authors wish to thank the Texas Schools Project and UT Dallas Education Research Center. The conclusions of this research do not necessarily reflect the opinions or official position of the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission or the State of Texas.

Washington: The research presented here would not have been possible without the administrative data provided by the Washington Office of Superintendent of Public Instruction through data-sharing agreement 2015DE-030.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jpubeco.2023.105016>.

## References

- Aaronson, D., Barrow, L., Sander, W., 2007. Teachers and student achievement in the Chicago public high schools. *J. Labor Econ.* 25 (1), 95–135.
- Ashenfelter, O., Krueger, A., 1994. Estimates of the economic return to schooling from a new sample of twins. *Am. Econ. Rev.* 84 (5), 1157–1173.
- Biasi, B., 2023. School finance equalization increases intergenerational mobility. *J. Labor Econ.* 41 (1), 1–38.
- Black, S.E., Devereux, P.J., 2011. Recent developments in intergenerational mobility. In: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*, Vol. 4. Elsevier, Amsterdam, pp. 1487–1541.
- Boyd, D., Lankford, H., Loeb, S., Wyckoff, J., 2013. Measuring Test Measurement Error: A General Approach. *J. Educ. Behav. Stat.* 38 (6), 629–663.
- Bruhn, J., 2020. The consequences of sorting for understanding school quality. Unpublished manuscript. Brown University.
- Burton, M., Brown, K., Johnson, A., 2013. Storylines about rural teachers in the United States: A narrative analysis of the literature. *J. Res. Rural. Educ.* 28 (12), 1–12.
- Chetty, R., Hendren, N., Kline, P., Saez, E., 2014c. Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Q. J. Econ.* 129 (4), 1553–1623.
- Chetty, R., Hendren, N., Jones, M.R., Porter, S.R., 2018. Race and economic opportunity in the United States: An intergenerational perspective. National Bureau of Economic Research Working Paper No. 24441.
- Chetty, R., Hendren, N., 2018. The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *Q. J. Econ.* 133 (3), 1107–1162.
- Chetty, R., Hendren, N., Katz, L.F., 2020. Response to “The moving to opportunity experiment: What do heterogeneous estimates of the effect of moving imply about causes?”. *Econ Journal Watch* 17 (2), 299–304.
- Chetty, R., Jackson, M.O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R.B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barbera, P., Bhole, M., Wernerfelt, N., 2022. Social capital I: Measurement and associations with economic mobility. *Nature* 608 (4), 108–121.
- Chyn, E., Katz, L.F., 2021. Neighborhoods matter: Assessing the evidence for place effects. *J. Econ. Perspect.* 35 (4), 197–222.
- Cicala, S., Fryer, R.G., Spenkuch, J.L., 2018. Comparative advantage in social interactions. *J. Eur. Econ. Assoc.* 16 (4), 983–1020.
- Clotfelter, C.T., Ladd, H.F., Vigdor, J.L., 2009. The Academic Achievement Gap in Grades 3 To 8. *Rev. Econ. Stat.* 91 (2), 398–419.
- Clotfelter, C.T., Ladd, H.F., Vigdor, J.L., 2015. The aftermath of accelerated algebra: Evidence from district policy initiatives. *J. Hum. Resour.* 50 (1), 159–188.
- Costrell, R.M., 1994. A simple model of educational standards. *Am. Econ. Rev.* 84 (4), 956–971.
- de Grujter, D.N.M., 1997. On Information of Percentile Ranks. *J. Educ. Meas.* 34 (2), 177–178.
- Denning, J.T., Murphy, R., Weinhardt, F., 2020. Class rank and long-run outcomes. National Bureau of Economic Research Working Paper no. 27468.
- Domina, T., McEachin, A., Penner, A., Penner, E., 2015. Aiming high and falling short: California's eighth-grade algebra-for-all effort. *Educ. Eval. Policy Anal.* 37 (3), 275–295.
- Elsner, B., Isphording, I.E., 2017a. A big fish in a small pond: Ability rank and human capital investment. *J. Labor Econ.* 35 (3), 787–828.
- Elsner, B., Isphording, I.E., 2017b. Rank, sex, drugs and crime. *J. Hum. Resour.* 53 (2), 356–381.
- Elsner, B., Isphording, I.E., Zoelitz, U., 2019. Achievement Rank Affects Performance and Major Choices in College. University of Zurich working paper.
- Fuller, W.A., 1987. *Measurement Error Models*. Wiley, New York.
- Goldhaber, D., Özek, U., 2019. How much should we reply on student test achievement as a measure of success? *Educ. Res.* 48 (7), 479–483.
- Hanushek, E.A., Rivkin, S.G., 2009. Harming the best: How schools affect the black-white achievement gap. *J. Policy Anal. Manage.* 28 (3), 366–393.
- Harding, D.J., Sanbonmatsu, L., Duncan, G.J., Genetian, L.A., Katz, L.F., Kessler, R.C., Kling, J.R., Sciandra, M., and Ludwig, J. (2021). Evaluating contradictory experimental and non-experimental estimates of neighborhood effects on economic outcomes for adults. NBER Working Paper No. 28454. Cambridge, MA: National Bureau of Economic Research.
- Haskins, R., Rouse, C., 2005. *Closing Achievement Gaps*. Policy Brief. Princeton-Brookings, Washington DC.
- Jang, H., Reardon, S.F., 2019. States as sites of educational (in)equality: State contexts and the socioeconomic achievement gradient. *AERA Open* 5 (3), 1–22.
- Kaestner, R., 2020. The moving to opportunity experiment: What do heterogeneous estimates of the effect of moving imply about causes? *Econ Journal Watch* 17 (2), 282–298.
- Koedel, C., Parsons, E., 2021. The effect of the community eligibility provision on the ability of free and reduced-price meal data to identify disadvantaged students. *Educ. Eval. Policy Anal.* 43 (1), 3–31.
- Kraft, M.A. (2017). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, vol. 54, no. 1, 2017, pp. 1-36.
- Laliberte, J.W., 2021. Long-term contextual effects in education: Schools and neighborhoods. *Am. Econ. J. Econ. Pol.* 13 (2), 336–377.
- Lefgren, L., Sims, D.P., 2012. Using subject test scores efficiently to predict teacher value-added. *Educ. Eval. Policy Anal.* 34 (1), 109–121.
- Lockwood, J.R., McCaffrey, D.F., 2014. Correcting for Test Score Measurement Error in ANCOVA Models for Estimating Treatment Effects. *J. Educ. Behav. Stat.* 39 (1), 22–52.
- May, K., Nicewander, W.A., 1994. Reliability and Information Functions for Percentile Ranks. *J. Educ. Meas.* 31 (4), 313–325.
- McDonough, I.K., 2015. Dynamics of the black-white gap in academic achievement. *Econ. Educ. Rev.* 47, 17–33.
- Murphy, R., Weinhardt, F., 2020. Top of the class: The importance of ordinal rank. *Rev. Econ. Stud.* 87, 2777–2826.
- National Center for Education Statistics. (2022). Public Charter School Enrollment. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved 10.15.2022 from <https://nces.ed.gov/programs/coe/indicator/cgb>.
- Parsons, E., Koedel, C., Podgursky, M., Ehlert, M., Xiang, P.B., 2015. Incorporating End-of-Course Exam Timing into Educational Performance Evaluations. *J. Res. Educ. Effect.* 8 (1), 130–147.
- Parsons, E., Koedel, C., Tan, L., 2019. Accounting for Student Disadvantage in Value-Added Models. *J. Educ. Behav. Stat.* 44 (2), 144–179.
- Reardon, S.F., 2011. The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In: Duncan, G.J., Murnane, R.J. (Eds.), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. Russell Sage Foundation, New York, NY, pp. 91–115.
- Reardon, S.F., Galindo, C., 2009. The Hispanic-White Achievement Gap in Math and Reading in the Elementary Grades. *Am. Educ. Res. J.* 46 (3), 853–891.
- Rothstein, J., 2019. Inequality of educational opportunity? Schools as mediators of the intergenerational transmission of income. *J. Labor Econ.* 37 (S1), S85–S123.
- Schoefer, B., Ziv, O., 2021. Productivity, place, and plants: Revisiting the measurement. Unpublished manuscript. UC Berkeley Department of Economics.
- Solon, G., 1999. Intergenerational mobility in the labor market. In: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*, Vol. 3. Elsevier, Amsterdam, pp. 1761–1800.
- Spearman, C., 1904. The proof and measurement of association between two things. *Am. J. Psychol.* 15, 72–101.
- Tavakol, M., and Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education* 2, 53-55.
- Todd, P.E., Wolpin, K.I., 2007. The production of cognitive achievement in children: Home, school, and racial test score gaps. *J. Hum. Cap.* 1 (1), 91–136.
- Wang, M., Stanley, J., 1970. Differential weighting: A review of methods and empirical studies. *Rev. Educ. Res.* 40, 663–705.

## Further reading

- Castex, G., Kogan Dechter, E., 2014. The changing roles of education and ability in wage determination. *J. Labor Econ.* 32 (4), 685–710.
- Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., Yagan, D., 2011. How does your kindergarten classroom affect your earnings? Evidence from Project Star. *Q. J. Econ.* 126 (4), 1593–1660.
- Chetty, R., Friedman, J.N., Rockoff, J.E., 2014a. Measuring the impacts of teacher I: Evaluating bias in teacher value-added estimates. *Am. Econ. Rev.* 104 (9), 2593–2632.

Chetty, R., Friedman, J.N., Rockoff, J.E., 2014b. Measuring the impacts of teacher II: Teacher value-added and student outcomes in adulthood. *Am. Econ. Rev.* 104 (9), 2633–2679.

Clark, D., Martorell, P., 2014. The signaling value of a high school diploma. *J. Polit. Econ.* 122 (2), 282–318.

Ferrer, A., Riddell, W., 2008. Education, credentials, and immigrant earnings. *Can. J. Econ.* 41 (1), 186–216.

Lazear, E.P., 2003. Teacher incentives. *Swedish. Econ. Policy Rev.* 10 (3), 179–214.

Mulligan (1999). Galton versus the human capital approach to inheritance. *Journal of Political Economy* 107, S184–S224.