Academic Mobility in U.S. Public Schools: Evidence from Nearly 3 Million Students

Wes Austin
David Figlio
Dan Goldhaber
Eric Hanushek
Tara Kilbride
Cory Koedel
Jaeseok Sean Lee
Jin Luo
Umut Ozek
Eric Parsons
Steven Rivkin
Tim Sass
Katharine Strunk

March 2023

WORKING PAPER No. 227-0323-3

CALDER
National Center for Analysis of Longitudinal Data in Education Research

AIR®

Academic Mobility in U.S. Public Schools: Evidence from Nearly 3 Million Students

Wes Austin
Georgia State University

David Figlio
Northwestern University / CALDER

Dan Goldhaber
American Institutes for Research / CALDER
University of Washington

Eric Hanushek
Stanford University / CALDER

Tara Kilbride
Michigan State University / EPIC

Cory Koedel
University of Missouri / CALDER

Jaeseok Sean Lee
University of Missouri

Jin Luo
University of Texas at Dallas

Umut Özek
RAND

Eric Parsons
University of Missouri

Steven Rivkin
University of Illinois at Chicago / CALDER

Tim Sass
Georgia State University / CALDER

Katharine Strunk
Michigan State University / EPIC / CALDER

Contents

Contents i
Acknowledgments ii
Abstract iii
1. Introduction 1
2. Data and Measurement of Academic Mobility 4
3. Findings 17
4. Correlates of Academic Mobility 26
5. Conclusion 30
References 33
Figures and Tables 38
Appendix A: Supplementary Figures & Tables 56
Appendix B: Replication of Results Using Alternative Measurement Error Corrections 75
Appendix C: Imputation Procedure 98
Appendix D: Estimating District Value-Added 101
Appendix E: Connecting Academic and Economic Mobility 103

Acknowledgments

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). All opinions expressed in this paper are those of the authors and do not necessarily reflect the views of our funders, data providers and partners, 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 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.

CALDER • American Institutes for Research

1400 Crystal Drive 10th Floor, Arlington, VA 22202

202-403-5796 • www.caldercenter.org

Academic Mobility in U.S. Public Schools: Evidence from Nearly 3 Million Students
Wes Austin, David Figlio, Dan Goldhaber, Eric Hanushek, Tara Kilbride, Cory Koedel, Jaeseok Sean Lee, Jin Lou, Umut Ozek, Eric Parsons, Steven G. Rivkin, Tim Sass, and Katharine Strunk
CALDER Working Paper No. 227-0323-3
March 2023

Abstract

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. On average, we show that student ranks are highly persistent during elementary and secondary education—that is, academic mobility is limited in U.S. schools as a 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 in principle provide a pathway to success. However, research suggests 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 generally 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 that the 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 over the course of 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 in the distribution of academic performance are less predictive of their later-grade ranks, and vice-versa for a system with low academic mobility. We estimate academic mobility using administrative panel data from seven states in the U.S. covering almost 3 million students.

Our estimation procedures for academic mobility 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).[1] 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 as students progress through the K-12 education system: eighth-grade test performance, high-school test performance, on-time high school graduation, and high school graduation within

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

one year of on-time. For most of our analysis we focus on upward mobility among initially low-achieving students. Specifically, and 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 who starts 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 performance rank, students from more advantaged backgrounds generally have greater upward mobility than their peers from less advantaged backgrounds. These results buttress existing research on the persistence, and even widening, of academic achievement gaps during K-12 schooling.

In addition to providing system-wide estimates of academic mobility, we also leverage the detailed microdata to explore variation in academic mobility across school districts. We focus primarily on understanding the extent of variation across districts in absolute upward mobility. Despite finding that academic mobility is low on average in U.S. public schools, we document statistically and economically significant variance in upward mobility across districts. We decompose the variance into two components, which we call "baseline mobility" and "relative mobility." Districts with high baseline mobility promote gains throughout the performance distribution; initially low achieving students are caught in a rising tide that lifts all boats. Alternatively, in districts with high relative mobility, initially low achieving students gain on their higher-achieving peers as they progress through school—i.e., these districts narrow their internal achievement gaps over time.[2]

[2] 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 Ozek, 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 Insphording (2017a, 2017b), Elsner et al. (2019), and Murphy and Weinhardt (2020).

Variation across districts in both baseline and relative mobility contributes to the total variance in absolute upward mobility. However, we find that most of the variance in absolute mobility is driven by district differences in baseline mobility. While our district-level mobility metrics are descriptive and should not be interpreted causally, these results are informative about the ways in which districts are most likely to help low-achieving students improve. The low variation in relative mobility is consistent with limited differences in success across districts at reducing internal achievement gaps.[3]

We also explore the correlates of academic mobility at the district level. We find that absolute upward mobility is largest in districts serving students from more socioeconomically advantaged backgrounds, measured along a variety of dimensions. 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 consistently 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 by race-ethnicity and income, we generally find that the same factors predict upward academic mobility for all students.

Finally, it is not surprising that differences in school quality have been postulated as a driver of the considerable geographic heterogeneity in economic mobility documented in recent research. Motivated by the potential connection between schools and economic mobility, we explore the link between our estimates of *intragenerational academic mobility* and external estimates of *intergenerational economic mobility*. To facilitate this analysis, we aggregate our estimates to the commuting-zone and county levels to match the levels of aggregation in recent studies of economic mobility by Chetty, Hendren, Kline, and Saez (2014) and Chetty and

[3] It may be that some districts are particularly effective in this way, as has been argued in several small-scale studies (Leithwood, 2010; Rorrer, Skrla, and Scheurich, 2008), but our results cast doubt on the notion that there are large differences across districts along this dimension more broadly.

Hendren (2018). We find that differences in academic mobility across commuting zones cannot account for a meaningful fraction of the observed variance in economic mobility at this level of geography (also see Rothstein, 2019).[4] A key reason is that most of the variance in academic mobility across school districts occurs within rather than between commuting zones. Across counties there is more variance in academic mobility, and we show that academic and economic mobility are positively correlated at the county level.

2. Data and Measurement of Academic Mobility

2.1 Data

The 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 who have standardized test scores in math and English language arts (ELA) in the third grade—the initial statewide testing grade in most states—and follow them through high school. Academic mobility is assessed as cohorts progress through school.

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 2005-06 and 2008-09 (hereafter, including in Table 1, we identify school years by the spring year; e.g., 2006 for "2005-06"). 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.

Our analysis includes about 2.9 million students, and 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

[4] Rothstein (2019) shows that 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.

an Individualized Education Program (IEP), and who are geographically mobile.[5] The structure of the education system also differs across states in terms of the shares of schools located in urban, suburban, and rural areas; and the numbers districts and schools, both in absolute and per-capita terms. While our sample is not designed to be representative of the United States as a whole, the seven states are diverse along many dimensions and provide substantively different evaluation contexts.[6]

Under the No Child Left Behind and Every Student Succeeds Acts, all students in public schools are required to be tested in math and ELA/reading in grades 3-8, and at least once in high school. Our analysis of academic mobility between grades 3 and 8 is fairly uniform across states due to federal testing requirements (although each state administers its own tests). At the high school level, however, the flexibility of testing requirements means the grades in which test outcomes are observed, and in which subjects, vary across states. 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.[7] With the exception of Michigan, which has a universal ACT/SAT policy, the common-grade requirement is such that the subject of the selected test is ELA-based. This is because the English curriculum in high schools is more rigidly structured than in other subjects. Table 2 shows that the focal high school tests are administered mostly in the tenth and eleventh grades (the exception is Georgia, where the test is administered in ninth grade), have very high coverage rates, and are overwhelmingly taken in a common grade. In Oregon there is no test given overwhelmingly in the same grade in high school, so we omit Oregon from the high-school achievement portion of our analysis.

---

[5] 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).

[6] Appendix Table A1 further shows that enrollment shares in charter schools in our third-grade cohorts is small in all states, ranging from 0-7.7 percent, with a median value of 1.8 percent.

[7] 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).

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 mobility. Focusing first on our test-based mobility metrics, they are constructed based on 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, which permits a more parsimonious presentation, is that the rank-rank relationship between intergenerational economic outcomes is functionally linear. This also turns out to be 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 ranks in each state. Figure 1 shows scatterplots from Georgia as an example, and 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 tests 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 and for all tests (we discuss the scatterplots for graduation outcomes below). Given the linear relationships, the mapping between students' early- (grade-3) and late-career outcomes can be summarized by equation (1):

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

In the equation, O_i is a late-career outcome rank for student i and R_i is student i's initial rank assessed in the third grade.

Figure 2 shows two stylized (extreme) mobility scenarios within the percentiles-to-percentiles framework. The first graph in the figure shows a case where α=0 and β=1. This is a scenario with no academic mobility, as the average entry and outcome ranks are the same at every percentile. At the other extreme, the second graph where α=50 and β=0 indicates perfect academic mobility; here, the average outcome rank is at the median regardless of the student's entry percentile. These examples illustrate the interdependence of α and β when the rank-rank relationship is estimated on the entire population, which in our context is the population of students in a state. Because the estimated regression line for an entire state must pass through the mean of the data and the model regresses percentiles on percentiles, then by construction it must pass through (50, 50). As a result, the mobility relationship is fully captured by the slope coefficient, β, which also defines the y-intercept, α, given by α=50-50β.

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 α and β are separately identified and provide unique information about baseline and relative mobility, respectively. This is because while the rank-rank lines pass through the point (50, 50) for the sample in each state as a whole, they need not pass through this point for each subpopulation. To illustrate, consider the following modified 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 measured in percentiles—continue to be based on 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. Analogously to CHKS, we focus on the mobility of students at the 25th percentile of the initial performance distribution to produce measures of absolute upward mobility for initially low-achieving students throughout our analysis, denoted by $\bar{O}_{25}$. From equation (4), $\bar{O}_{25}$ for students in district *d* is estimated by $\hat{\alpha}_d + \hat{\beta}_d * 25$.

Figure 3 plots linear mobility functions for two hypothetical districts corresponding to equation (3). In each of the three panels, the slopes of the solid and dashed lines are held constant (i.e., neither β_{Solid} nor β_{dashed} change across panels), with the solid line exhibiting more relative mobility because $\beta_{solid}<\beta_{dashed}$. Hence, two students in the solid-line district with a given performance gap at panel entry would be expected to have a larger outcome gap than two students in the dashed line district with the same initial gap at panel entry. In each graph the intercept is also larger for the solid district ($\alpha_{solid}>\alpha_{dashed}$)—i.e., the lowest-performing students from the solid district perform better on the outcome measure than the lowest-performing students from the dashed district.

Figure 3 shows that α has uniform implications throughout the initial performance distribution—i.e., if α is high for district *d*, it increases the ranks of all students in the statewide distribution of later outcomes. In contrast, changes to β have different implications depending on students' positions in the initial performance distribution. Because of this, and as noted by CHKS, it is straightforward to interpret a higher value of baseline mobility (α) as a positive attribute, but the same is not true of relative mobility (β). This point is exemplified by the comparison in the third panel of Figure 3 where initially low-achieving students perform similarly in both districts and the higher relative mobility of the solid line reflects the underperformance of initial high achievers. Many researchers and education systems use the district achievement gap as a measure of performance, but Figure 3 highlights the insufficiency of this measure due to the potential tradeoff between inequality and the outcome level for initially low achievers. For instance, 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. The interpretation of the intercept and slope parameters described thus far applies to their estimation on percentile-based outcomes, but graduation is a binary outcome. Noting this difference, the academic mobility parameters are conceptually similar in the graduation models (CHKS, 2014). For example, $\bar{O}_{25d}$ for on-time graduation indicates the likelihood of high-school graduation for a student in 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 Figure 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 attributable to a combination of strong floor effects on graduation likelihoods 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' observed third grade test scores. These high-stakes state tests meet the highest standards of test publishers in terms of their reliability, but they are not error-free.[8] Measurement error in student scores derives 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).

The measurement error in students' initial test scores will induce mean reversion in our models that track their ranks over time. If left unaccounted for, this will lead us to overstate academic mobility. To illustrate, consider an extreme scenario where initial test scores are comprised entirely of measurement error. Under the standard assumption that the error is uncorrelated with the outcome, the expected value of β in our mobility regressions would be zero, implying perfect academic mobility. More generally, measurement error in students' initial scores will attenuate our estimates of β and correspondingly inflate our estimates of α.[9]

Within the context of a latent-ability framework, we use two different approaches to address measurement error in students' initial ranks. Both approaches 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 can write:

$$S_i^M = S_i + M_i + e_i^M \quad (5)$$

$$S_i^E = S_i + E_i + e_i^E \quad (6)$$

[8] The test reliability estimates for the tests 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).
[9] 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).

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 ranks in math and ELA in the third grade to set the initial rank. By averaging the two noisy measures, the error variance is reduced. Then, we make an additional correction for test measurement error in the average initial ranks. The additional correction is designed to correct for error deriving from the testing instruments themselves—i.e., due to sampling variance in the items that appear on the tests. We make the correction using test reliability ratios reported by test publishers that capture error from this source, which we incorporate into our models using a standard errors-in-variables (EIV) regression framework. The EIV models disattenuate β by the expected value of the attenuation bias caused by the measurement error (and correspondingly shrink α). Procedurally, the error variance is subtracted out of the total variance in the initial ranks in calculating the error-adjusted parameter β (Fuller, 1987; Lockwood and McCaffrey, 2014).

An additional technical complication is that we use the average of the entry ranks to set the initial rank variable, but the reliability ratios from the test publishers are for the individual tests. Define r_m and r_e as the reliability ratios for the third grade math and ELA standardized tests individually, and $\theta_{m,e}$ as the correlation of performance on the two tests. Following Wang and Stanley (1970), the reliability of performance as measured by the 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 in each state.[10,11]

The above approach to adjust for measurement error has strengths and weaknesses. Its greatest strength is that it is efficient, which is an especially important property when we estimate mobility parameters for each district individually (as in equation (3) above). But it is not a comprehensive correction and has two notable limitations. First, it is unlikely to fully correct for measurement error. The averaging strategy is conceptually appealing for reducing the error variance, but we average over just two scores, leaving room for ample error to remain. 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 β to be fully disattenuated; rather, they provide lower-bound estimates of β. This means that the EIV specifications will overstate relative academic mobility to some degree (recall that a lower value of β corresponds to more academic mobility).

The second limitation is that our 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 individual 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 our other comparisons, including across individual districts.

We assess the severity of these concerns by replicating our entire analysis using an alternative measurement error correction where we set the initial rank based on the third grade

[10] 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.

[11] Note that 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.

math score, then instrument for the math rank with the ELA rank. The IV approach addresses the two main limitations of our preferred approach. First, while in principle the averaging approach could match the IV approach in terms of the strength of the error correction if the number of initial tests was large; in practice, with the availability of just two tests, the IV approach will make a stronger correction (Ashenfelter and Krueger, 1994). Moreover, the correction via instrumental variables 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 error correction.

But while these theoretical benefits make the IV approach appealing, it also has limitations. Although the subject-specific skills are constructed to be orthogonal to general skill in third grade, they may be related to the outcome directly and not just through the third-grade math score, in which case the IV estimates will not be consistent. Such a failure of the exclusion restriction would likely inflate the estimates of β because the subject-specific skill in third grade would be positively related to skill in the future. This would lead to upper-bound estimates of β that understate relative academic mobility. In addition, the IV models are substantially less efficient, which is problematic for our district-level analysis in particular.[12]

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 β 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 β, 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

[12] When we estimate the district-level IV models, we must drop small districts from the analysis. See Appendix B for details.

academic mobility across school districts, are similar using either approach. This suggests that while conceptually concerning, in practice the potential for measurement-error heterogeneity to confound these comparisons is limited.

Finally, in addition to the EIV and IV results, we also estimate our models using a third approach where we set the initial rank using the average of the third-grade math and ELA ranks, but without the additional EIV correction. As expected, given the weaker correction for measurement error, our estimates of β are consistently smallest using this approach. But again, our comparative findings are substantively similar, further reinforcing the idea that there is a limited scope for measurement error to confound the comparisons. Results using this third approach are also shown in Appendix B in tandem with the IV results.

2.2.2.2 Geographic Mobility

We do not observe later-grade outcomes for all students in the intial entry cohorts because some students exit the public school system and/or leave their home states.[13] Table 3 shows that system exiters are negatively selected—i.e., the average entry percentiles of students whose later-grade outcomes are unobserved are almost always below those of students with observed outcomes.[14] These results are aligned with outside evidence on negative selection into student mobility across schools and districts, on average—e.g., see Goldhaber et al. (2022), Grigg (2012), and Hanushek, Kain, and Rivkin (2004).

The sample attrition raises two concerns. The first is reference bias and applies to our analyses of the eighth grade and high school test percentiles, which are normed against the population of test takers. Because state leavers are negatively selected, their departure from the data, if left unaddressed, would lead us to understate upward mobility in our sample even if there is no unobserved selection associated with system exits (note that the reference bias issue is not

[13] For the test outcomes, we also lose students who do not take the tests, although the tests we use are meant to be given to all students, minimizing sample attrition for this reason.

[14] 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, which produces lumpiness of percentiles that can fall above or below the median.

relevant to our analysis of graduation outcomes because these outcomes are not normed in the distribution). The second concern is related to unobserved heterogeneity: our findings will be biased by sample composition changes if state exiters are different from state stayers in unobserved ways, conditional on their initial ranks. This concern applies to both our test-based and graduation-based mobility metrics and has the potential to be especially problematic for subgroup analyses. For example, consider a scenario where 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 the outcome variance between districts.

We address both of these concerns 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 high school graduation outcomes.[15] The imputed values allow us to preserve the full entry-cohort distributions in each state, mitigating the concern about reference bias.[16]

In terms of unobserved selection bias, we cannot directly estimate the effect of unobserved selection. Instead we address this issue by building hypothetical selection scenarios into the imputation framework. The baseline selection scenario, which is the scenario 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."

---

[15] 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.

[16] Noting that system exiters and system entrants after the third grade are both negatively selected (for the same reasons), an alternative but less comprehensive strategy to combat reference bias is to replace state exiters with state entrants in the outcome distributions. We have pursued this strategy as well and our results are qualitatively similar throughout (results omitted for brevity); we prefer our imputation-based approach because it is more comprehensive and tractable.

Using this scenario as an anchor, we consider the sensitivity of our findings to four scenarios where the degree of selection among students with missing outcomes is re-parameterized. The re-paramaterizations of selection relative to baseline are as follows: (1) 25 percent more negative, (2) 10 percent more negative, (3) 10 percent less negative, and (4) 25 percent less negative. With the selection-adjusted imputed values in hand, we can 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 missing outcomes, above-and-beyond selection into district mobility within the public school system of a state. Full details regarding our imputation procedure are provided 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 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 also not extreme, and (3) most students in our sample states are not missing outcomes (Table 3), which limits the scope for attrition to impact our findings.

Finally, we turn to the issue of geographic mobility within the state public school systems. We assign students to districts based on the third grade, which means that our estimates of cross-district variability take on an interpretation akin to "intent-to-treat" parameters. They reflect the set of schooling or other experiences associated with students' third-grade districts. Although most students remain in the same district during grades 3-12, many change districts as well. Some of these 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.[17] Disentangling the reasons for student mobility across districts, and the implications, is a substantial undertaking and a natural extension of this work, but here we focus on understanding differences in student academic mobility across districts as defined by the district attended in the third grade.

3. Findings

For presentational convenience we focus the discussion of our findings primarily on simple averages of the state-level results.[18] That said, we conduct our entire analysis separately for each state and present many of the state-by-state results alongside the state averages in the main text. State-by-state results that are suppressed in the main text are available in Appendix A.

3.1 Broad Patterns of Academic Mobility at the State Level

Table 4 reports estimates of β and $\bar{O}_{25}$ from equation (1) for each state and on average across the states in our sample (recall that α is redundant in the statewide models). Consistent with prior evidence that early measures of achievement are highly predictive of later outcomes, a student's position in the test distribution in the third grade is highly predictive of both eighth grade and high school test rankings. The cross-state average estimates of β in the eighth-grade and high-school test models are 0.84 and 0.82, respectively, and the estimates are quite similar across states, ranging from 0.80-0.87. As high as these are, these estimates likely understate β because of the incomplete correction for measurement error. As expected, the alternative IV estimates exceed the EIV estimates in each of the seven states (see Appendix Table B2a), bounding β from above. Put plainly, in all the states we examine, 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 β are much lower and more variable across states. The simple-average values of β for on-time and lagged graduation are 0.35 and

[17] 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.
[18] 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.

0.27, respectively, reflecting a much weaker gradient between initial percentile ranks and the likelihood of graduating from high school. The weaker gradient is visually apparent in the scatterlplots in Figure 1 and Appendix Figure A1, and is driven by the fact that graduation rates are high throughout most of the entry-rank distribution. Put another way, because high school graduation is a fairly indiscriminate outcome, early-career performance ranks are weaker predictors of success.

Like with the β's, the $\bar{O}_{25}$ values for the test outcomes are similar across states and tests, ranging from 28.2-33.2, with average values around 30 for each test. The IV analogs in Appendix B are lower—ranging from 26.0-28.2. 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 estimates 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-based mobility parameters are differences in statewide graduation rates. Given that most students graduate, the graduation-rate differences are particularly salient for students in the lower end of the performance distribution. Unsurprisingly, states with higher graduation rates have higher graduation-based $\bar{O}_{25}$ values. This finding highlights an important source of ambiguity in interpreting the mobility findings with respect to graduation. One interpretation of a high $\bar{O}_{25}$ value is that it reflects a state's success in pushing initially low-performing students through high school. But an alternative interpretation is that a high graduation rate for initially low-performing students reflects low standards for receiving a high school diploma (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 that districts' test-based and graduation-based mobility metrics are positively correlated (ρ ≈ 0.2-0.3). 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 entering and outcome percentile ranks are not group-specific, but rather remain normed against the full state distribution. This allows for the separate identification of α_s and β_s, with the tradeoff that it may also overstate the academic mobility of higher performing subgroups relative to lower performing subgroups in the presence of uncorrected test measurement error (Hanushek 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 β, but we are focusing on the differences between subgroups.

Table 5 and Figure 4 show results by race/ethnicity where we compare mobility for Asian, Black, Hispanic, and White students.[19] Focusing on $\bar{O}_{25s}$, which is reported directly in Table 5 and marked by a vertical line at the 25th percentile of the entry distribution in each graph in Figure 4, we find that initially low-performing Asian students have much higher upward academic mobility than all other racial-ethnic groups: the $\bar{O}_{25s}$average value for the eighth grade test equals 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. Figure 4 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 α_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

[19] 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.

approaches 1.0 for students with high entry percentiles.

A comparison of the estimates in Table 5 with those based on the IV measurement error correction reported in Appendix Table B3a suggests that for the test outcomes, measurement error may account for almost half of the Asian-other race gaps in $\bar{O}_{25s}$ and most of the differences among Blacks, Hispanics, and Whites. For example, the average values of $\bar{O}_{25s}$for the eighth grade test based on the IV correction are 33.9 for Asians, 26.7 for Blacks, 27.6 for Hispanics, and 27.1 for Whites. Importantly, substantial gaps remain between Asian and other students despite the more comprehensive measurement-error correction, indicating much higher upward mobility for Asian students in terms of test performance. For graduation outcomes, the racial-ethnic gaps are similar regardless of which error correction we use and like with the test-based gaps, the graduation gaps strongly favor Asian students.

The results in Table 5 (and Appendix B3a) also show the outsized influence of baseline mobility (α) in driving variation in absolute upward mobility ($\bar{O}_{25}$) across the racial-ethnic groups. The differences in relative mobility (β) are modest in comparison. The importance of baseline mobility can be demostrated by decomposing the total change in $\bar{O}_{25}$ between groups. For instance, consider 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 α between Asian and Black students, with only a small fraction of the gap remaining to be explained the gap in β (which is multiplied by a factor of 25 to map to $\bar{O}_{25}$). The value of α is mechanically overstated relative to β by focusing at a point in the distribution below the 50th percentile; still, even evaluated at $\bar{O}_{50}$, α 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 (most consistently for graduation outcomes) 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 contribute to mixed findings in the literature. For example, Clotfelter, Ladd, and Vigdor (2009) find that the Hispanic-White achievement gap narrows during grades 3-8 in North Carolina. Alternatively, Reardon and Galindo (2009) find that the Hispanic-White gap is flat from grades 1-5 using a nationally representative sample, and Todd and Wolpin (2007) find it remains flat or widens modestly.[20]

Next, Table 6 and Appendix Table B4a follow the structure of Table 5 but show splits by FRL status instead of race-ethnicity. Compared to FRL students, non-FRL students have greater absolute upward mobility on average across states. For test scores in eighth grade and high school, the average $\bar{O}_{25}$ gaps in Table 6 are 4.5 and 6.7 percentage points, respectively, between FRL and non-FRL students. The average on-time and lagged graduation gaps are 12.5 and 11.0 percentage points. Again, the more comprehensive treatment of measurement error in Appendix Table B4a reduces the magnitude of the test-based $\bar{O}_{25}$ gaps to 2.5 and 4.0 percentage points on the eighth-grade and high-school tests, respectively. Like with the racial-ethnic comparsions, the more comprehensive treatment of measurement error has little effect on the graduation-based mobility gaps.

The last subgroup comparison is by school urbanicity in the third grade, shown in Table 7. Here there is much less heterogeneity across groups. For the eighth-grade test, there are only small differences in absolute upward mobility between the urbanicity subgroups on average across states. The gaps widen modestly for the high school test, but the most notable differences in Table 7 are in terms of graduation outcomes. Graduation rates for initially low-performing students who attend urban schools are significantly lower than graduate rates for their peers who

[20] A more nuanced explanation of Reardon and Galindo's (2009) findings is as follows: point estimates imply a modest shrinking of the gap in math and a modest increase in reading. Although we do not perform formal tests, based on their reported standard errors it seems likely that their confidence intervals would include our estimates if the analytic approaches were otherwise aligned.

attend suburban and rural schools (who have similar graduation rates to each other). These results are 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 α, β, and $\bar{O}_{25}$ as estimated by equations (3) and (4). These estimates capture 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 very small number of charters are intergated 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).[21]

The raw variances of $\hat{\alpha}_d$, $\hat{\beta}_d$, and $\hat{\bar{O}}_{25d}$ will overstate the true variances due to sampling variance. We net out the sampling variance using a randomized inference procedure in which we randomly assign students to districts, then estimate "null distributions" of $\hat{\alpha}_d$, $\hat{\beta}_d$, and $\hat{\bar{O}}_{25d}$ that entirely reflect sampling variance. We repeat this procedure 300 times and use the average variance across the 300 null distributions as an estimate of the sampling variance, which we subtract from each raw-variance estimate.[22] The randomized inference procedure maintains all aspects of the data structure in each state (e.g., district sizes and the relationships between the entry ranks and outcomes for individual students).

To illustrate, define $\sigma^2_{\hat{\alpha}}$ as the unadjusted variance of $\hat{\alpha}_d$ estimated using the real data in a given state, and $\bar{\sigma}^2_{\hat{\alpha},null}$ as the average value of the null-distribution variance with random student

[21] Total 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.
[22] We replicate this procedure just 200 times in Texas due to complications associated with computing demands (and noting that the Texas database is especially large).

assignments to districts. The standard deviation of the parateter of interest, α_d, net of sampling variance, is estimated as:

$$\sigma_\alpha = \sqrt{\sigma_{\hat{\alpha}}^2 - \bar{\sigma}_{\hat{\alpha},null}^2} \tag{5}$$

We apply a similar procedure to obtain error-variance-corrected estimates of σ_β and σ_{O25}.[23] The null distributions from the randomized inference procedure also allow for empirical tests of statistical significance of the variances of these parameters. We say that the variance of a given parameter across districts is statistically signficant in a state if the variance estimate using the observed data falls outside of the 95-percent confidence interval of the null-distribution values.

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 signficant. On average across states, the EIV results in Table 8 indicate that one standard deviation in the distribution of absolute upward mobility ($\bar{O}_{25}$) corresponds to a change in student rank on the eighth-grade and high-school tests of 4.8-4.9 percentile points. The estimated standard deviations for the IV-based estimates are about 10 percent larger. For on-time and delayed graduation, the analogous average standard deviations of our EIV estimates are 5.6 and 4.8 percentage points, respectively. Similar to the patterns for race-ethnicity, income, and urbanicity differences, the graduation-based estimates are far less sensitive to the treatment of measurement error.

These cross-state averages indicate modest but non-trivial variance in academic mobility across districts. Adding context from Table 4, the estimates in Table 6 indicate that a third-grade entrant at the 25th percentile who attends a district with academic mobility that is one-standard-deviation above average on the high school test would be expected to score at the 35.3rd percentile in the state distribution, compared to the 30.4th percentile at the average district. In

[23] Our randomized inference procedure yields estimates of the error variance that can be approximated using the average of the squared standard errors on α_d and β_d from the individual district regressions. In results omitted for brevity, we confirm that the "squared standard errors" approximation yields similar results.

terms of on-time graduation, a similar comparison at the 25th percentile of the entry distribution would yeild a graduation likelihood at the high-mobility district of 81.4 percent, versus 75.8 percent at the average district.

Figure 5 shows the distributions of $\bar{O}_{25d}$ in two example states, Missouri and Washington, for all outcomes. The distributions exhibit differences consistent with the results in Table 8 (e.g., see the especially tight distribution of $\bar{O}_{25d}$ for the late graduation outcome in Missouri). In terms of their properties, the distributions are consistently unimodal and smooth, ruling out odd patterns of heterogeneity across districts (for instance, a pattern where most districts do not differ in terms of academic mobility but a small handful exhibit exceptional differences).

It is also of interest to compare the importance of α_d and β_d in driving upward mobility, but the comparison is complicated because this varies by the initial rank—i.e., at low initial ranks, variation in α_d will be a more important driver of upward mobility but as the initial rank increases, β_d becomes more important. This dynamic is illustrated in Figure 6, which we construct based on the cross-state averages of our estimates of the variance of α_d and β_d for the high school test from Table 8. The solid, parallel lines are at plus-and-minus one standard deviation of α_d and are fixed throughout the distribution. The dashed, parting lines are equal to the percentile on the horizontal axis multiplied by plus-and-minus one standard devation of β_d.

The figure shows the relative importance of variation in α_d and β_d across districts in driving differences in total academic mobility over the support of initial ranks. At a given entry percentile, if the gap between the solid lines is larger, then variation in α_d is a larger driver of variance in total upward mobility at that percentile, and vice versa if the gap between the dashed lines is larger. The key takeaway from Figure 6 is that over most of the distribution of initial ranks, and certainly at lower-valued ranks, variation in α is by far the 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 73rd percentile, where $\sigma_\beta * p \approx \sigma_\alpha$.

Although separable inference is challenging because α_d and β_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.

Figure 6 also provides another illustration of the consistent theme in our results that there is more variation in academic-mobility intercepts than slopes, 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 α_d and β_d is suggestive of how districts are likely to affect the trajectories of low performers absent reforms to current practice.

Finally, in Tables 9 and 10 we report correlations between district estimates of absolute upward mobility across outcomes and cohorts. We adjust the correlation between any two sets of estimates by first estimating the ratio of the true variance to the total variance for each set of estimates (where the true variance is estimated using the randomized-inference procedure described above). Then we multiply the 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 bounds on the correlations because it assumes all estimation error is uncorrelated. We also show unadjusted correlations that provide complementary lower bounds, which are smaller but of the same sign. For ease or 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.[24]

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

[24] The state-by-state correlations are directionally aligned and share broad patterns, but there is substantial heterogeneity in their magnitudes. We were unable to identify any insightful patterns in the cross-state heterogeneity worthy of discussion.

high—for test outcomes the average correlation across states is 0.84, and for graduation outcomes it is 0.98. The adjusted correlations across outcome modes are positive but lower, ranging from 0.24-0.29.

Table 10 shows analogous correlations within districts and outcomes, but across cohorts. Note the states that contribute to the average in each cell depend on which year cohorts are included in the state samples (per Table 1). The contiguous-cohort, adjusted correlations are between 0.54 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 none of the correlations across any cohort for any outcome is below 0.41.

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-ethnicty (measured by the segregation of underrepresented minority students, who we define as Black and Hispanic), and in addition, we construct a parallel segregation index based on economic status (measured by FRL enrollment) motivated by recent research on economic connectedness (Chetty et al., 2022).[25] All these variables are constructed for school districts using data from students in our cohorts in the third grade.

[25] 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.

An additional district attribute we construct based on 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 conditional on student characteristics. 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 also construct the value-added estimates so they are 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 additional 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. We pull these variables 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. We use NCES data from the 2010-14 period to construct all of these variables.[26]

In Figure 7 we show 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 of one—therefore, the

[26] 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.

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.[27] The detailed state-by-state regression output underlying Figure 7, including information about statistical significance for individual coefficients, is available in Appendix Table A4. The broad patterns in the results are similar if we use IV-based estimates of $\bar{O}_{25d}$ as dependent variables in place of the EIV-based estimates (see Appendix Table B9a).[28]

The preceding analysis offers some predictions about the directions of the coefficients. 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 Figure 7. More broadly, Figure 7 shows that absolute upward mobility is highest 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.[29]

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). Given that the relationships in Figure 7 are not causal there are many

---

[27] A one-standard-deviation change with respect to the value-added measure is based on the raw data. Given that the value-added measures are shrunken using the approach of Lefgren and Sims (2012), 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).

[28] 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 a subsample of larger districts, and correspondingly, the IV-based correlates are estimated for just this subsample. See the discussion in Appendix B for more information.

[29] 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 positively signed and many are statistically significant. See Appendix Table A4.

potential explanations. These include redistributive spending that targets disadvantaged children such as Title I and special education programs, compensating differentials for educators to account for more 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. In other words, we re-estimate the district-level mobility regressions three additional times for each district – once including only Black students, once including only Hispanic students, and once including only FRL-enrolled students. 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 these subgroup specific estimates of academic mobility are less precisely estimated—sometimes by a considerable margin depending on the student composition of a district. This should not cause bias in our regressions because $\bar{O}_{25d}$ is the dependent variable, but it does reduce efficiency, weakening the precision with which we can identify some relationships in the data. That said, we generally find that the attributes that predict higher academic mobility overall also predict higher academic mobility for these at-risk student groups. Figures analogous to Figure 7, but for academic mobility measured for each student type, are provided in Appendix Figures A2-A4.

Second, we correlate district value-added separately with α_d and β_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 β_d—the association is small and inconsistently signed across states and outcomes. In contrast, the associations between district value-added and α_d are overwhelmingly positive. This suggests the correlations between value added and $\bar{O}_{25d}$ documented in Figure 7 are driven primarily by variation across districts in α_d, not β_d, and further reinforces 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 intergenrational economic mobiltiy from CHKS (2014) and Chetty and Hendren (2018), respectively. Details for this portion of our analysis are provided in Appendix E. A high-level summary of the results is as follows. First, there is insufficient cross-commuting-zone variance in academic mobility to account for observed variance in economic mobility at this same level of geography. 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.[30] Between counties there is more variation in academic mobility because counties typically cover much smaller geographic areas. In Appendix E, we show that our estimates of academic mobiltiy 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. Our analysis is based on administrative panel data from seven states covering nearly 3 million students. 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

[30] 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 that differences in schools are an important driver of place-based effects on students' educational attainment using narrowly-defined geographic areas.

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 throughout the distribution of initial performance ranks. 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, Brown, and Johnson, 2013).

When we decompose total academic mobility into its components and examine cross-district heterogeneity, we find differences across districts 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 as students progress through their schooling careers, at least given current educational policies and practices.

We correlate absolute upward mobility with a wide array of district and local-area characteristics. A general theme of this portion of our analysis is that absolute upward mobility is largest in socioeconomically advantaged areas as measured along a variety of dimensions. We also show that districts with high value added to student test scores have significantly higher upward mobility (as measured by test and non-test outcomes).

Finally, we use our estimates of academic mobility to gain insight into the scope for differences in school quality to explain geographic variation in economic mobility across commuting zones and counties. We find that 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.[31]

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, Hendren, and Katz, 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.

[31] 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).

References

Aaronson, Daniel, Lisa Barrow and William Sander. 2007. Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics* 25(1), 95-135.

Ashenfelter, O., and Krueger, A. (1994). Estimates of the economic return to schooling from a new sample of twins. *American Economic Review* 84(5), 1157-73.

Biasi, B. (2023). School finance equalization increases intergenerational mobility. *Journal of Labor Economics* 41(1), 1-38.

Black, S.E., & Devereux, P.J. (2011). "Recent developments in intergenerational mobility," in *Handbook of Labor Economics Vol. 4* (eds. O. Ashenfelter and D. Card), 1487-1541. Amsterdam: Elsevier.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2013). Measuring Test Measurement Error: A General Approach. *Journal of Educational and Behavioral Statistics* 38(6), 629-663.

Bruhn, J. (2020). The consequences of sorting for understanding school quality. Unpublished manuscript. Brown University.

Burton, M., Brown, K., and Johnson, A. (2013). Storylines about rural teachers in the United States: A narrative analysis of the literature. *Journal of Research in Rural Education* 28(12), 1-12.

Castex, G., & Kogan Dechter, E. (2014). The changing roles of education and ability in wage determination. *Journal of Labor Economics* 32(4), 685-710.

Chetty, R., Friedman, J.N., Hilger, N., Saez, E., Schanzenbach, D.W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project Star. *Quarterly Journal of Economics* 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. *American Economic Review* 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. *American Economic Review* 104(9), 2633-2679.

Chetty, R., Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *Quarterly Journal of Economics* 133(3), 1107-1162.

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., and 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., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics* 129(4), 1553-1623.

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., and Katz, L.F. (2021). Neighborhoods matter: Assessing the evidence for place effects. *Journal of Economic Perspectives* 35(4), 197-222.

Cicala, S., Fryer, R. G., and Spenkuch, J. L. (2018). Comparative advantage in social interactions. *Journal of the European Economic Association* 16(4), 983-1020.

Clark, D., & Martorell, P. (2014). The signaling value of a high school diploma. *Journal of Political Economy* 122(2), 282-318.

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2015). The aftermath of accelerated algebra: Evidence from district policy initiatives. *Journal of Human Resources* 50(1), 159-188.

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2009). The Academic Achievement Gap in Grades 3 To 8. *Review of Economics and Statistics* 91(2), 398-419.

Costrell, R.M. (1994). A simple model of educational standards. *American Economic Review* 84(4), 956-971.

de Gruijter, D.N.M. (1997). On Information of Percentile Ranks. *Journal of Educational Measurement* 34(2), 177-78.

Denning, J. T., Murphy, R., and 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. *Educational Evaluation and Policy Analysis* 37(3), 275-295.

Elsner, B., and Isphording, I. E. (2017a). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics* 35(3), 787-828

Elsner, B., and Isphording, I. E. (2017b). Rank, sex, drugs and crime. *Journal of Human Resources* 53(2), 356-381

Elsner, B., Isphording, I.E., and Zoelitz, U. (2019). Achievement Rank Affects Performance and Major Choices in College. University of Zurich working paper.

Ferrer, A., & Riddell, W. (2008). Education, credentials, and immigrant earnings. *The Canadian Journal of Economics* 41(1), 186-216.

Fuller, W. A. (1987). *Measurement Error Models*. New York: Wiley

Goldhaber, D., & Özek, U. (2019). How much should we reply on student test achievement as a measure of success? *Educational Researcher* 48(7), 479-483.

Goldhaber, D., Koedel, C., Özek, U., & Parsons, E. (2022). Using Longitudinal Student Mobility to Identify At-Risk Students. *AERA Open*, 8.

Grigg, J. (2012). School enrollment changes and student achievement growth: A case study in educational disruption and continuity. *Sociology of Education* 85(4), 388-404.

Hanushek, E.A., Kain, J.F., and Rivkin, S.G. (2004). Disruption versus Tiebout improvement: the costs and benefits of switching schools. *Journal of Public Economics* 88, 1721-1746.

Hanushek, E.A., and Rivkin, S.G. (2009). Harming the best: How schools affect the black-white achievement gap. *Journal of Policy Analysis and Management* 28(3), 366-393.

Harding, D.J., Sanbonmatsu, L., Duncan, G.J., Gennetian, 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., and Rouse, C. (2005). Closing Achievement Gaps. Policy Brief. Washington DC: Princeton-Brookings.

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., and Parsons, E. (2021). The effect of the community eligibility provision on the ability of free and reduced-price meal data to identify disadvantaged students. *Educational Evaluation and Policy Analysis* 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. *American Economic Journal: Economic Policy* 13(2), 336-377.

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

Lefgren, L., & Sims, D.P. (2012). Using subject test scores efficiently to predict teacher value-added. *Educational Evaluation and Policy Analysis* 34(1), 109-121.

Leithwood, K. (2010). Characteristics of school districts that are exceptionally effective in closing the achievement gap. *Leadership Policy Schools* 9, 245–291.

Lockwood, J.R. and McCaffrey, D.F. (2014). Correcting for Test Score Measurement Error in ANCOVA Models for Estimating Treatment Effects. *Journal of Educational and Behavioral Statistics* 39(1), 22-52.

May, K., & Nicewander, W.A. (1994). Reliability and Information Functions for Percentile Ranks. *Journal of Educational Measurement* 31(4), 313-25.

McDonough, I.K. (2015). Dynamics of the black-white gap in academic achievement. *Economics of Education Review* 47, 17-33.

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

Murphy, R., and Weinhardt, F. (2020). Top of the class: The importance of ordinal rank. *Review of Economic Studies* 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. *Journal of Research on Educational Effectiveness* 8(1): 130-147

Parsons, E., Koedel, C., & Tan, L. (2019). Accounting for Student Disadvantage in Value-Added Models. *Journal of Educational and Behavioral Statistics* 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 *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances* (eds. G.J. Duncan & R.J. Murnane), 91-115. New York, NY: Russell Sage Foundation.

Reardon, S.F., & Galindo, C. (2009). The Hispanic-White Achievement Gap in Math and Reading in the Elementary Grades. *American Educational Research Journal* 46(3), 853-891.

Rorrer, A.K., Skrla, L., & Scheurich, J.J. (2008). Districts as institutional actors in educational reform. *Educational Administration Quarterly* 44(3), 307-358.

Rothstein. J. (2019). Inequality of educational opportunity? Schools as mediators of the intergenerational transmission of income. *Journal of Labor Economics* 37(S1), S85-S123.

Schoefer, B., and 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 *Handbook of Labor Economics Vol. 3* (eds. O. Ashenfelter and D. Card), 1761-1800. Amsterdam: Elsevier.

Spearman, Charles (1904). The proof and measurement of association between two things. *The American Journal of Psychology* 15, 72-101.

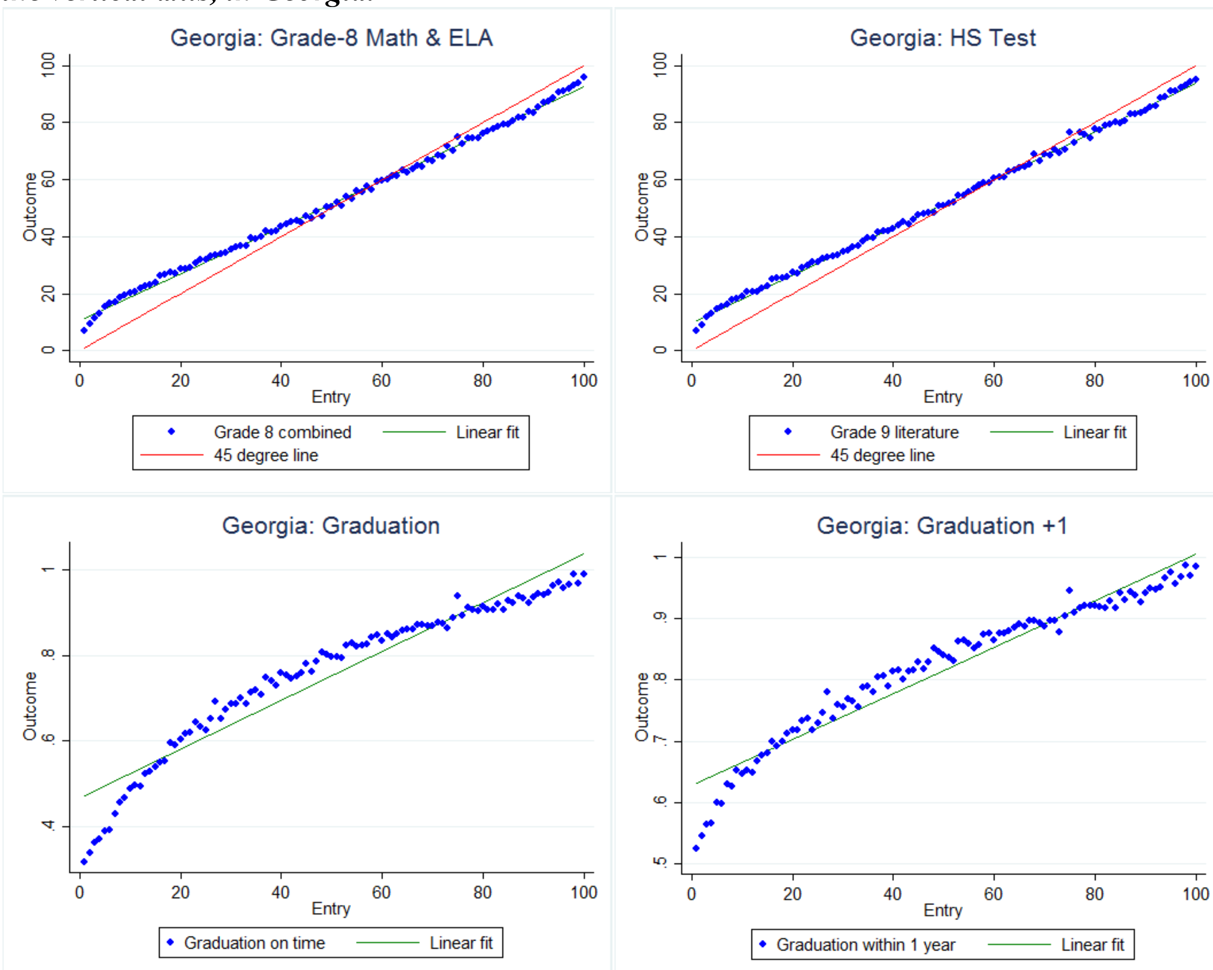
Tavakol, M., and Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education* 2, 53-55.

Todd, P.E., & Wolpin, K.I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital* 1(1), 91-136.

Wang, M. & Stanley, J. (1970). Differential weighting: A review of methods and empirical studies. *Review of Educational Research* 40, 663-705.

Figures and Tables

Figure 1. Binned scatter plots with percentile ranks on the 3rd grade test on the horizonal 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.

Figure 2. Hypothetical illustrations of the linear rank-rank relationship. No mobility (left) versus perfect mobility (right).

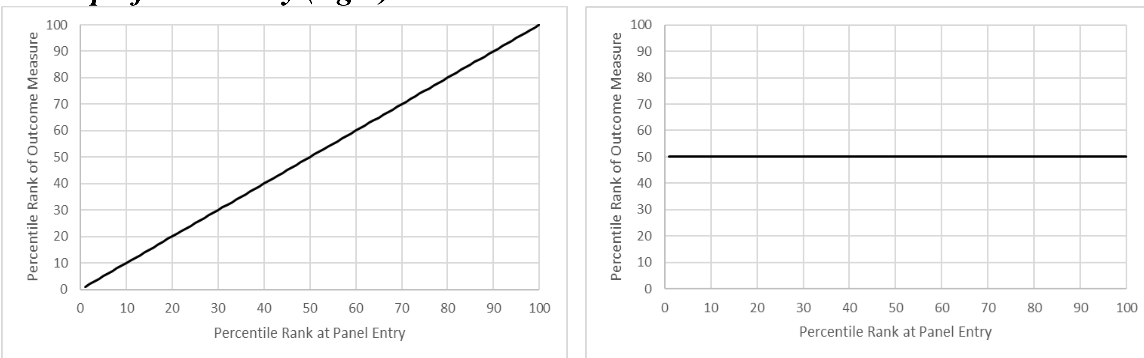
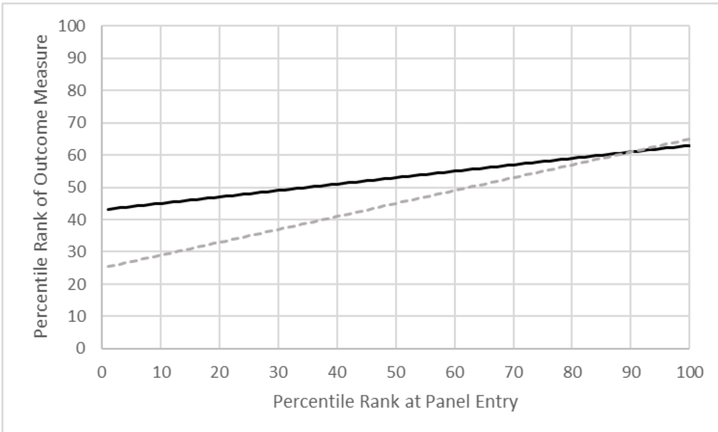
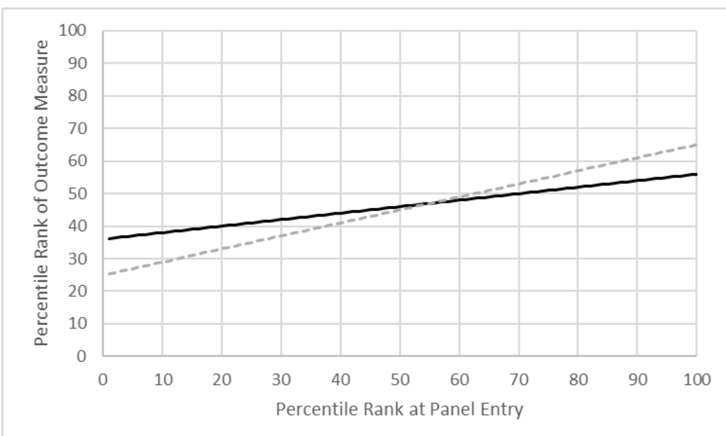


Figure 3. Comparison of two hypothetical districts, one with higher relative mobility (solid lines) and one with lower relative mobility (dashed lines), with differing gaps in baseline mobility.

(a) Large gap in baseline mobility:



(b) Medium gap in baseline mobility:



(c) Small gap in baseline mobility:

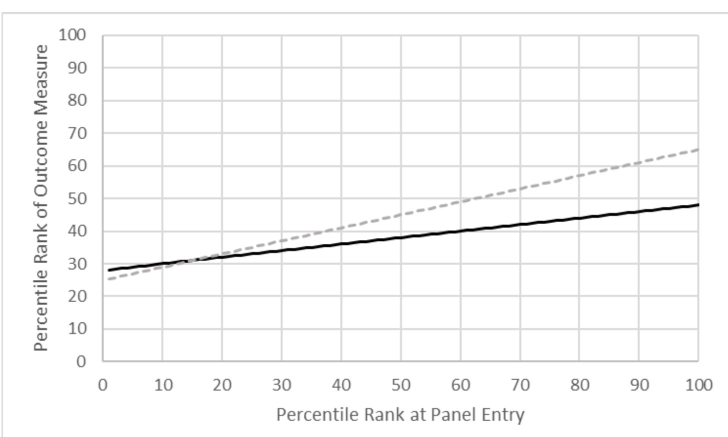
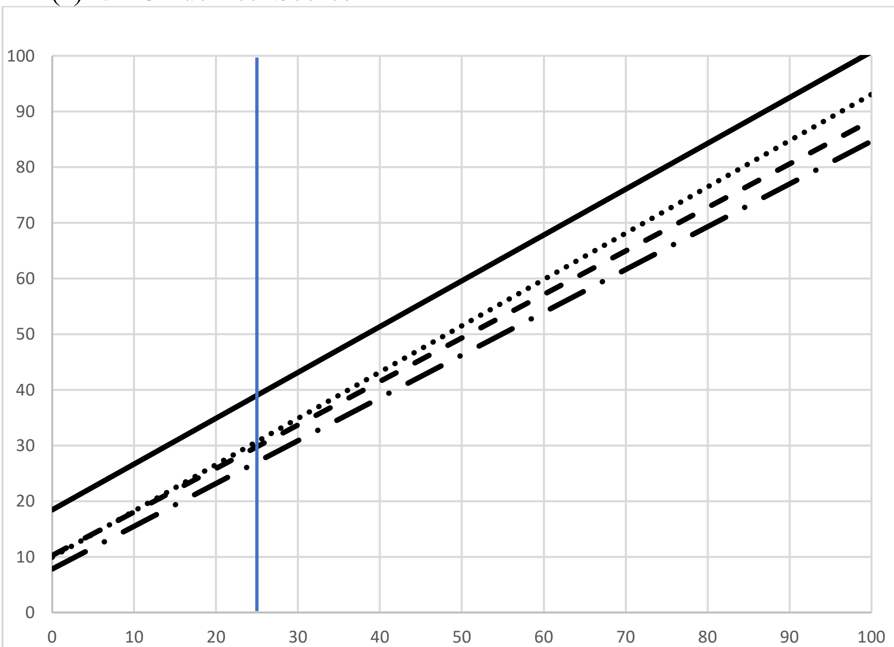
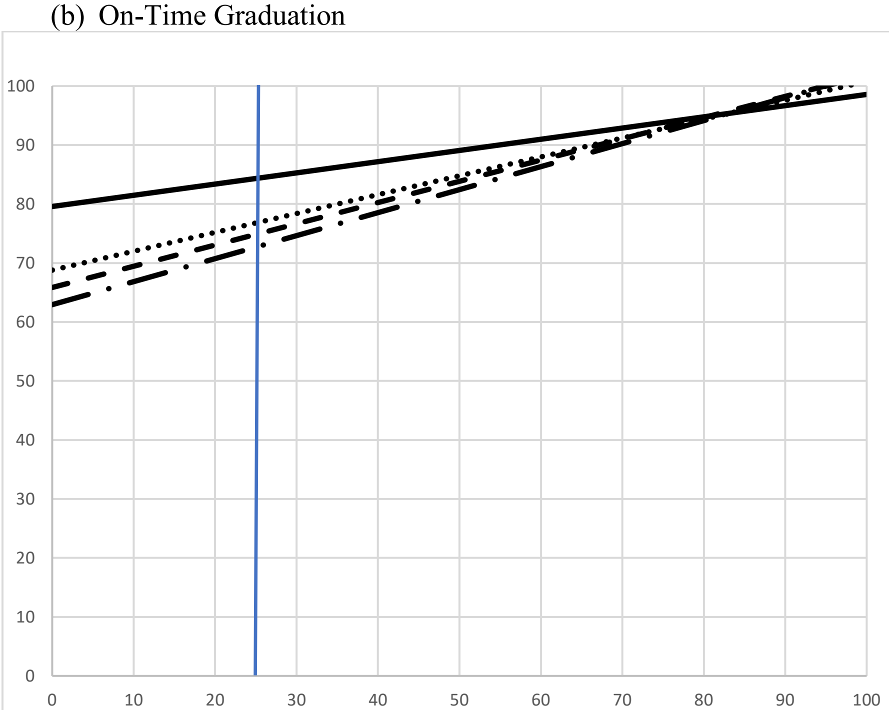


Figure 4. Illustrations of the linearly-estimated rank-rank relationships for 8th-grade test score and on-time graduation outcomes by race-ethnicity, corresponding to the results in

(a) 8th-Grade Test Scores

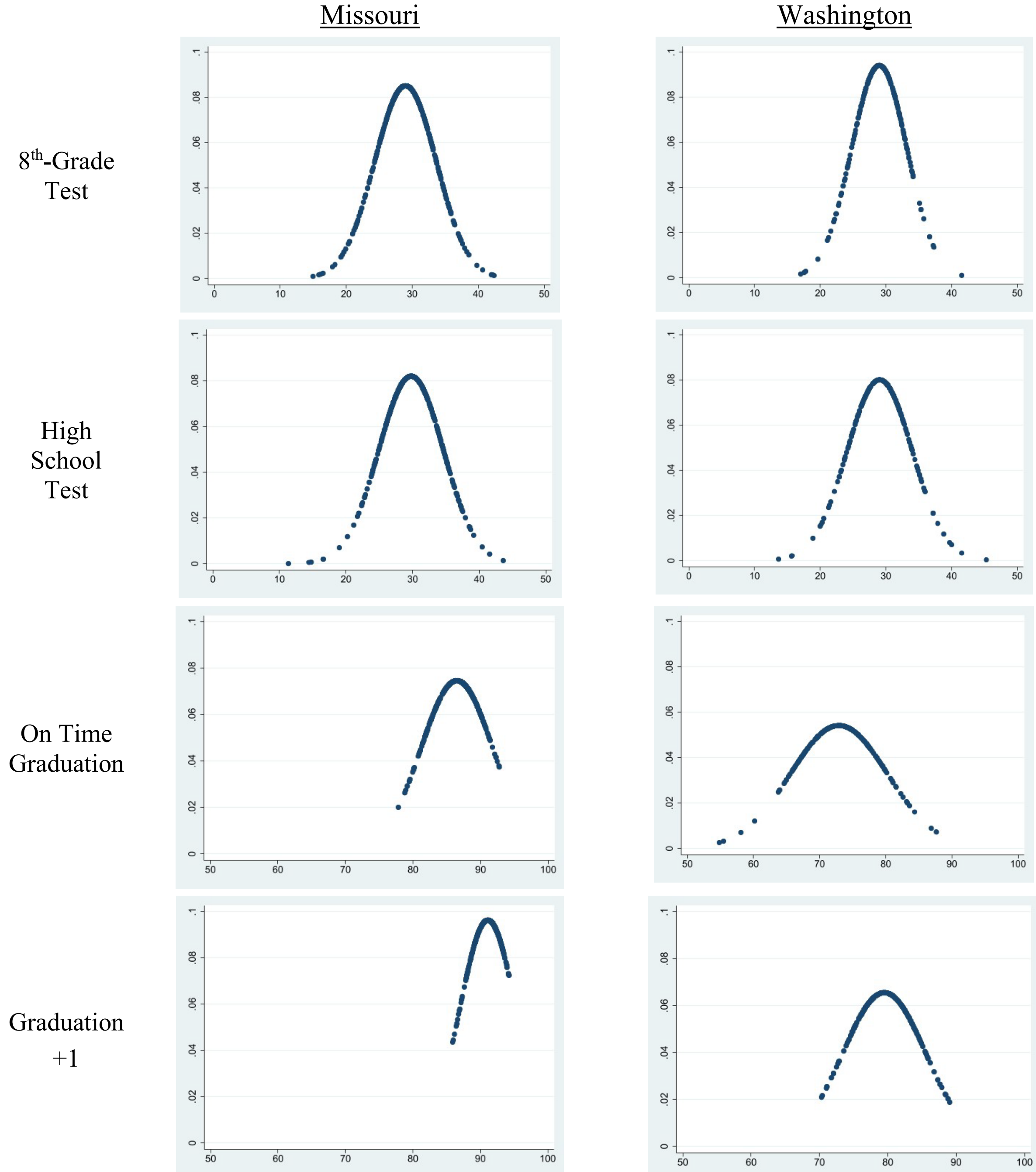


(b) On-Time Graduation



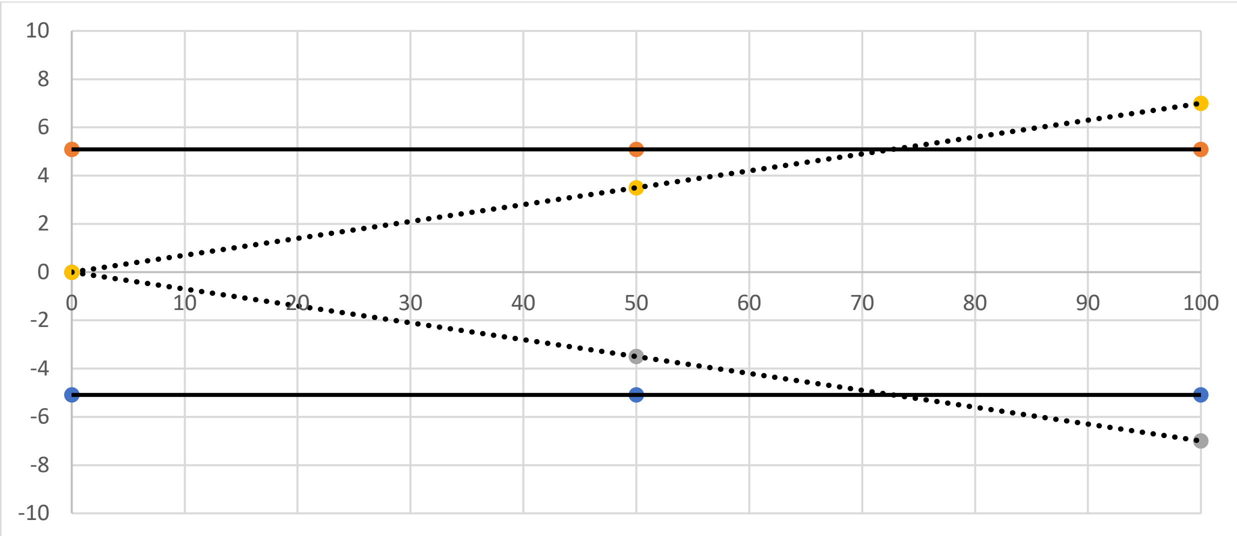
Notes: These graphs illustrate the linearly-estimated relationships by race-ethnicity between the 3rd grade test rank and (a) the 8th-grade test rank and (b) on-time high school graduation. The linear-model parameters for each race-ethnicity and outcome are shown in Table 5. Solid lines: Asian students; Dash-dotted lines: Black Students; Dashed lines: Hispanic Students; Dotted lines: White Students.

Figure 5. Distributions of   in two states, Missouri and Washington, for all outcomes. These distributions are visual complements to the results in Table 8.



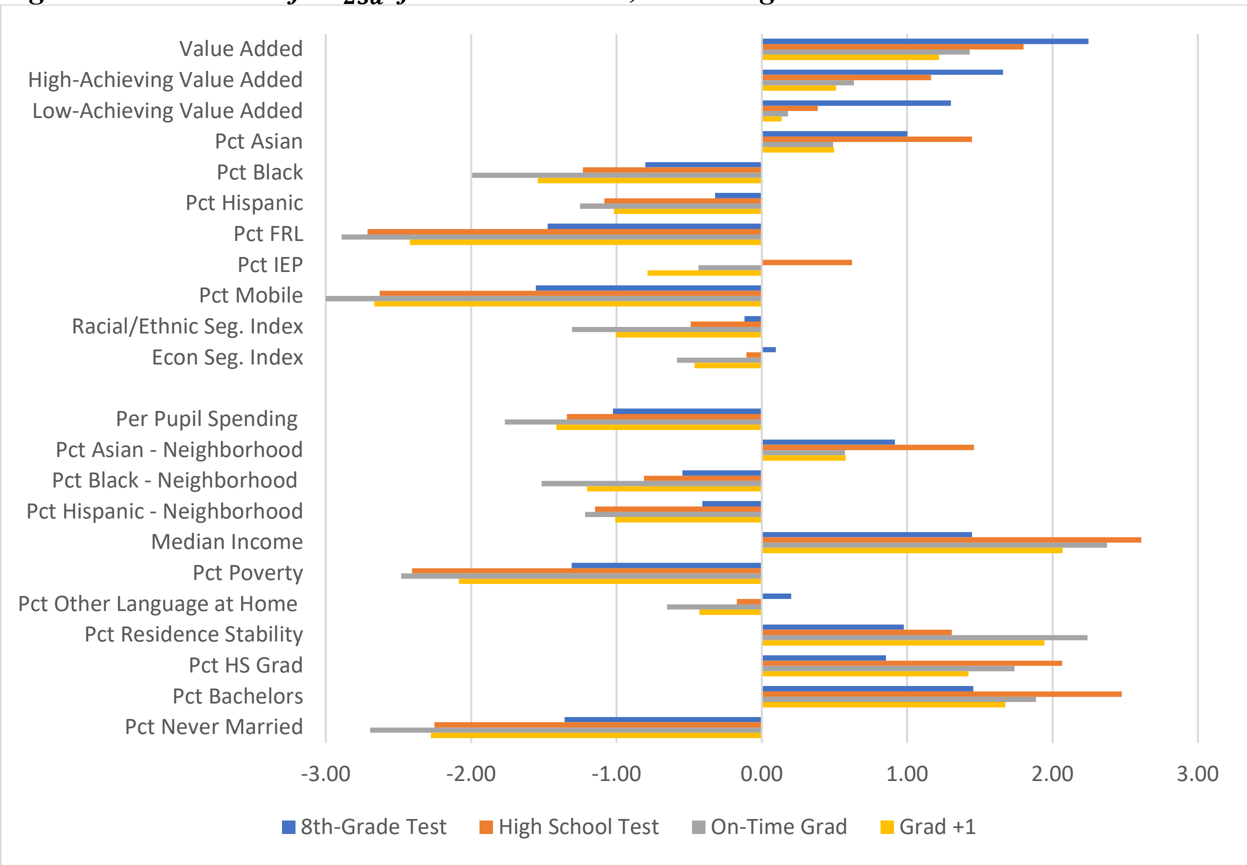
Notes: We shrink O25 to the mean to remove excess variation due to estimation error. The shrinkage formula is $\pi * \bar{O}_{25d}$ + (1-π)*$\bar{O}_{25}$, where $\bar{O}_{25d}$ is the district's estimated O25 value and $\bar{O}_{25}$ is the state average value. π is a common shrinkage factor applied to all estimates and calculated as the ratio of the true variance of $\bar{O}_{25d}$ to the raw variance. The numerator of π is estimated using the randomized inference procedure.

Figure 6. Illustration of the variance in α_d and β_d across the support of entering ranks, based on the simple average standard deviations across states for the high school test in Table 8.



Notes: This figure illustrates the variation in academic mobility across the support of entry ranks as documented in Table 8 for the high school test. The vertical and horizonal axes are in percentile-point units. The solid lines bound +/- one standard deviation of α. The dotted lines bound +/- one standard deviation of β multiplied by the entry percentile. The relative importance of variance in α_d and β_d across the support of entry ranks can be inferred by comparing the gaps between the solid and dashed lines. If the solid-line gap is larger, variance inα_d contributes more to explaining variance in academic mobility at the given percentile, and if the dotted-line gap is larger, variance in β_d contributes more. The crossing point (where the two gaps are the same size) occurs at approximately the 73rd entry percentile.

Figure 7. 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.

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	--	--	--

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.

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 β and $\bar{O}_{25}$ for each outcome.

	Grade-8 Test		HS Test		Grad		Grad +1	
	β	O25	β	O25	β	O25	β	O25
All (Avg)	**0.84**	**29.66**	**0.82**	**30.44**	**0.35**	**75.76**	**0.27**	**80.59**
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: In these statewide regressions corresponding to equation (1), α and β are not separately identified. O25 is equal to α + 25*β. Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results. All β coefficients are statistically significant; standard errors and statistical significance information suppressed for brevity.

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

	Grade-8 Test			HS Test			Grad			Grad +1		
Student Group: Asian												
	α	β	O25	α	β	O25	α	β	O25	α	β	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	39.5	20.85	0.8	40.81	91.43	0.1	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	0.2	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
Student Group: Black												
	α	β	O25	α	β	O25	α	β	O25	α	β	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	80.7	0.21	85.89	81.92	0.19	86.7
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	5.4	0.76	24.49	Not Applicable			57.6	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
Student Group: Hispanic												
	α	β	O25	α	β	O25	α	β	O25	α	β	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	9.1	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	70.5	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	28.3	62.07	0.38	71.61	71.93	0.27	78.66
Student Group: White												
	α	β	O25	α	β	O25	α	β	O25	α	β	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	88.4	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	0.8	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	30.2	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 β coefficients are statistically significant; standard errors and statistical significance information suppressed for brevity.

Table 6. Statewide academic mobility estimates by FRL status.

	Grade-8 Test			HS Test			Grad			Grad +1		
Student Group: FRL												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*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	26.4	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	66.3	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
Student Group: non-FRL												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*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	31.9	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	0.8	32.5	Not Applicable			73.42	0.24	79.47	76.5	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	0.19	84.7

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 β 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.

	Grade-8 Test			HS Test			Grad			Grad +1		
Student Group: Urban												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*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	7.9	0.85	29.15	57.60	0.44	68.51	67.40	0.33	75.61

Student Group: Suburban												
	α	*β*	*O25*	*α*	*β*	*O25*	*Α*	*β*	*O25*	*α*	*β*	*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	0.3	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

Student Group: Rural												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*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	91.6	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	63.28	0.36	72.39	72.4	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 β coefficients are statistically significant; standard errors and statistical significance information suppressed for brevity.

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

	Grade-8 Test			HS Test			Grad			Grad +1		
	Standard Deviations											
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**4.94**	**0.06**	**4.81**	**5.09**	**0.07**	**4.88**	**7.02**	**0.10**	**5.56**	**5.96**	**0.08**	**4.81**
GA	3.41	0.04	3.51	3.87	0.05	4.00	6.62	0.10	6.98	5.93	0.08	6.19
MA	6.58	0.06	6.03	6.55	0.06	6.00	4.98	0.06	3.61	4.72	0.06	3.44
MI	4.11	0.06	4.06	5.04	0.08	4.93	7.66	0.10	6.24	7.00	0.09	5.54
MO	4.24	0.07	4.37	4.45	0.07	4.31	4.95	0.06	3.70	3.05	0.03	2.49
OR	7.33	0.07	6.65	Not Applicable			9.66	0.12	7.27	9.30	0.11	7.03
TX	5.60	0.06	5.15	5.85	0.06	5.47	7.55	0.13	5.08	5.77	0.09	4.27
WA	3.28	0.07	3.88	4.75	0.07	4.57	7.74	0.10	6.05	5.98	0.09	4.73

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 randomized inference procedure described in the text. In results suppressed for brevity, we obtain similar adjusted estimates if we make the adjustment 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.

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

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.84	1.00		
Grad	0.29	0.26	1.00	
Grad +1	0.29	0.24	0.98	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.74	1.00		
Grad	0.25	0.23	1.00	
Grad +1	0.24	0.20	0.91	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				
	2006	2007	2008	2009
2006	1.00			
2007	0.67	1.00		
2008	0.58	0.68	1.00	
2009	0.49	0.63	0.73	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.57	1.00		
2008	0.49	0.60	1.00	
2009	0.40	0.53	0.61	1.00

HS Test

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.68	1.00		
2008	0.65	0.69	1.00	
2009	0.59	0.52	0.71	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.56	1.00		
2008	0.57	0.61	1.00	
2009	0.52	0.57	0.68	1.00

Graduation

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.55	1.00		
2008	0.48	0.55	1.00	
2009	0.54	0.61	0.65	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.44	1.00		
2008	0.38	0.44	1.00	
2009	0.44	0.48	0.50	1.00

Graduation +1

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.54	1.00		
2008	0.41	0.56	1.00	
2009	0.53	0.62	0.63	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.43	1.00		
2008	0.32	0.44	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.

Table 11. Correlations of district-level value added to student achievement with α_d and β_d.

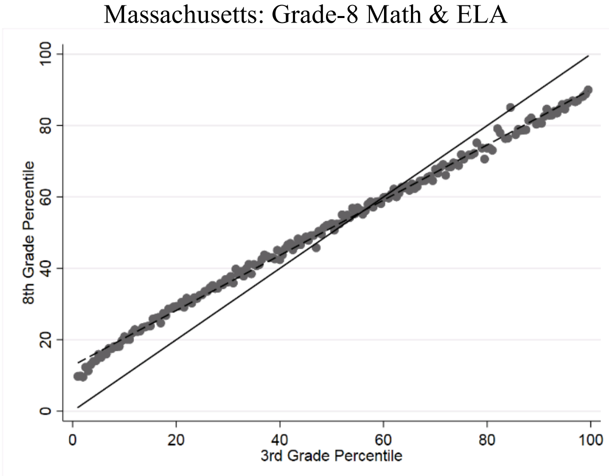
	Grade-8 Test		HS Test		Grad		Grad +1	
	α	β	α	β	α	β	α	β
All (Avg)	**0.35**	**0.05**	**0.21**	**0.02**	**0.14**	**-0.10**	**0.14**	**-0.10**
GA	0.29	-0.14	0.15	-0.18	0	0.05	-0.01	0.06
MA	0.28	0.21	0.23	0.17	0.17	-0.14	0.17	-0.15
MI	0.40	0.15	0.33	0.18	0.22	-0.20	0.21	-0.20
MO	0.34	0.05	0.19	0.01	0.14	-0.16	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.

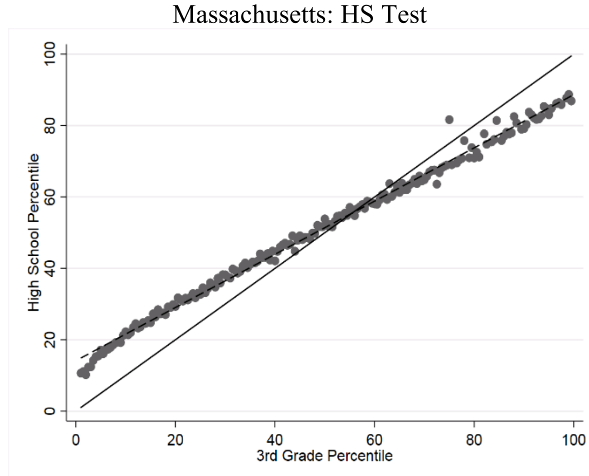
Appendix A: Supplementary Figures & Tables

Appendix Figure A1. Binned scatter plots with entry percentiles on the horizonal axis and either test-outcome percentiles or graduation rates on the vertical axis. Scatterplots are for all outcomes and all states except Georgia, for which the scatterplots are shown in Figure 1 of the main text.

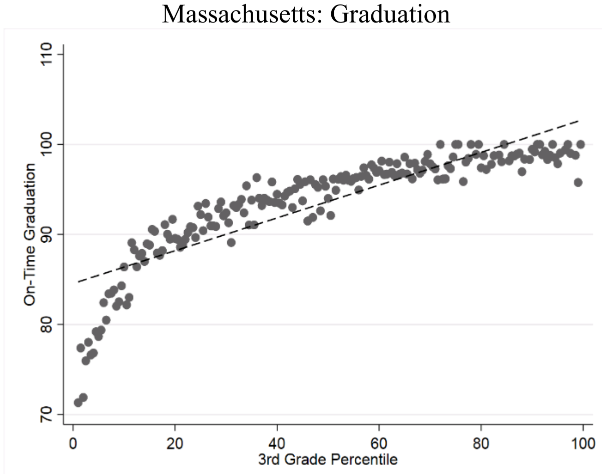
Massachusetts: Grade-8 Math & ELA



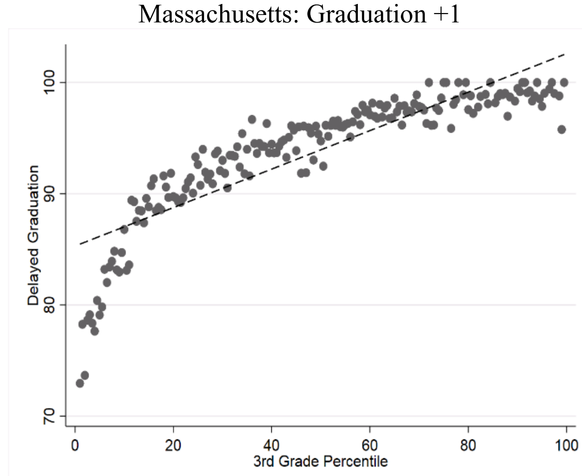
Massachusetts: HS Test



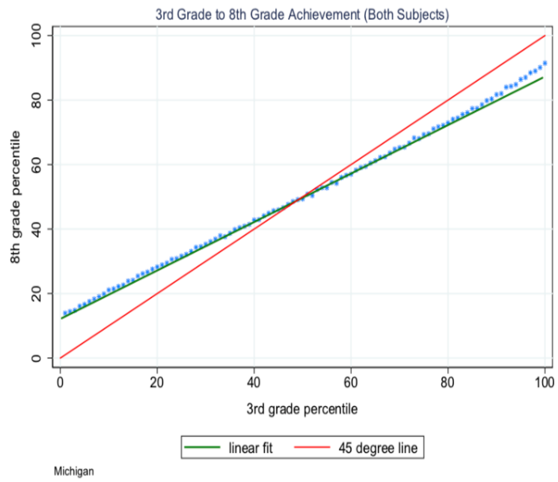
Massachusetts: Graduation



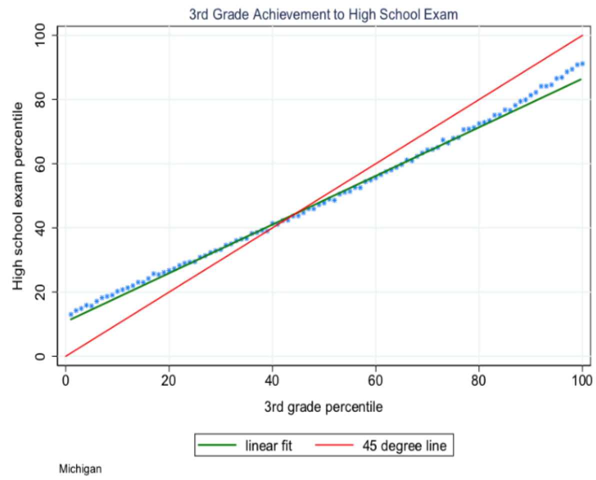
Massachusetts: Graduation +1



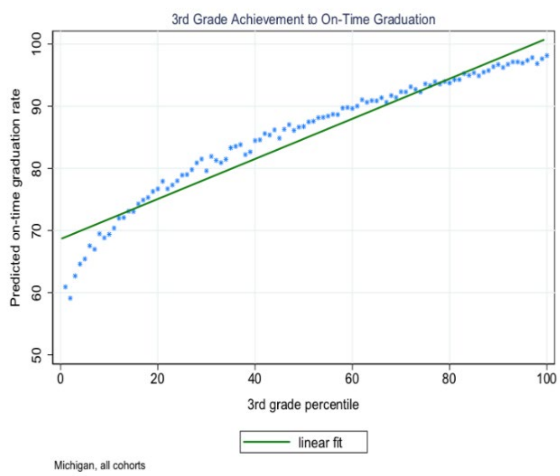
Michigan: Grade-8 Math & ELA



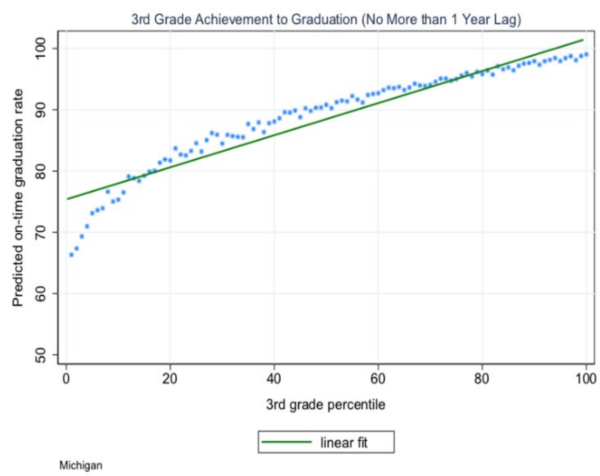
Michigan: HS Test



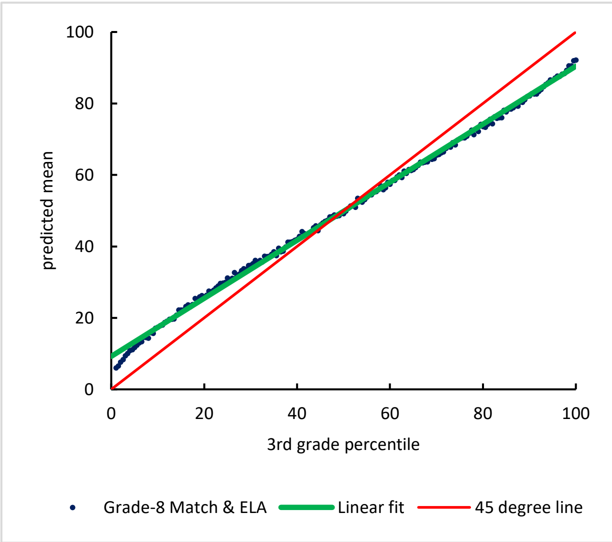
Michigan: Graduation



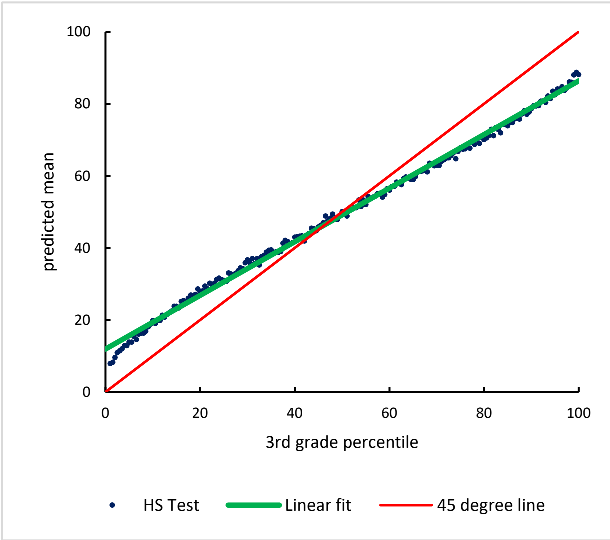
Michigan: Graduation +1



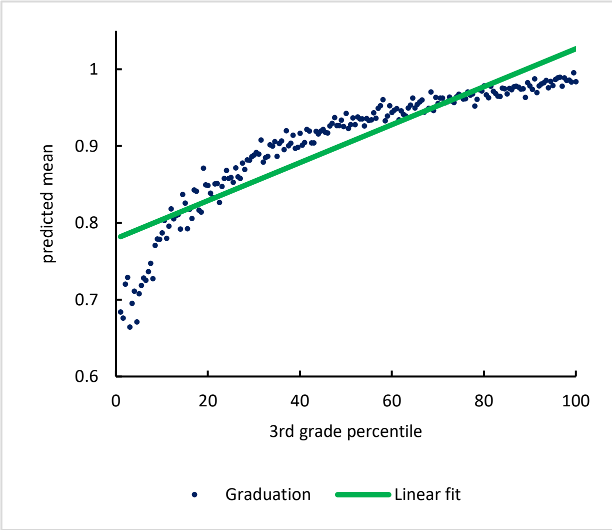
Missouri: Grade-8 Math & ELA



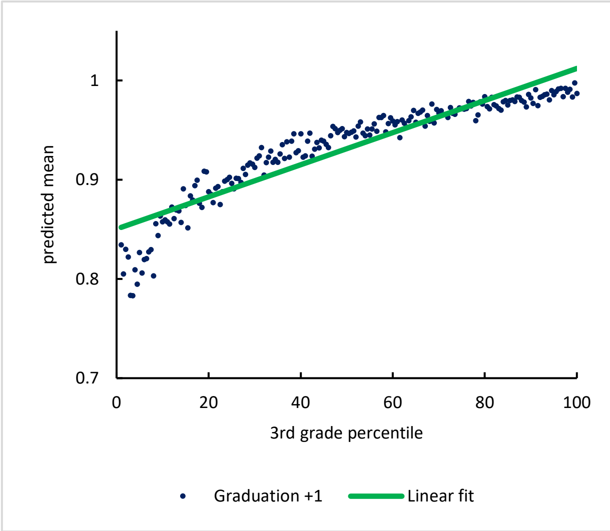
Missouri: HS Test



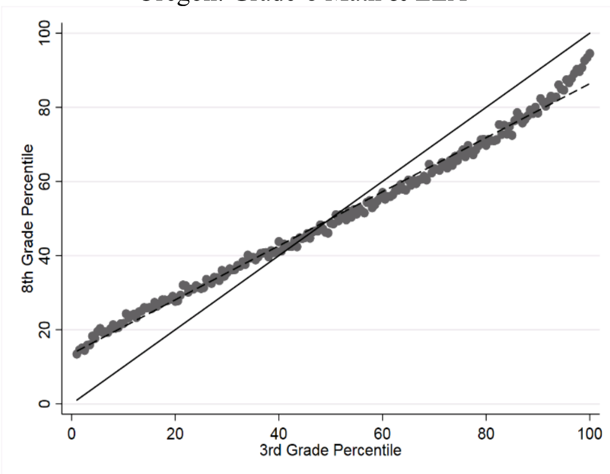
Missouri: Graduation



Missouri: Graduation +1



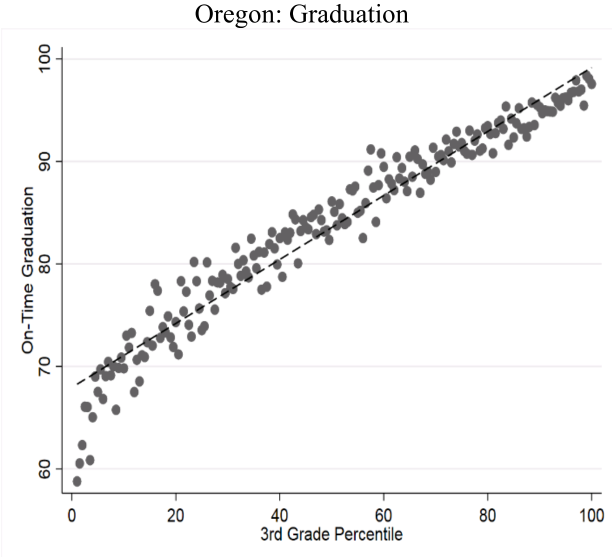
Oregon: Grade-8 Math & ELA



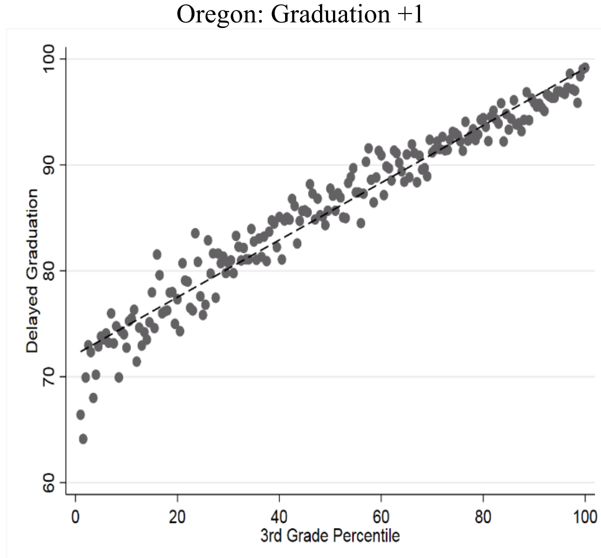
Oregon: HS Test

(omitted)

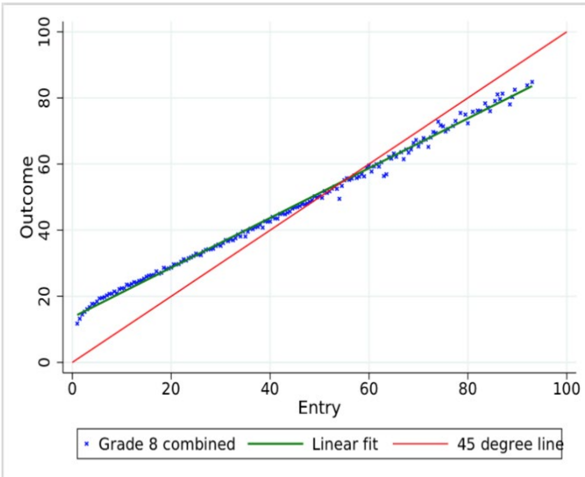
Oregon: Graduation



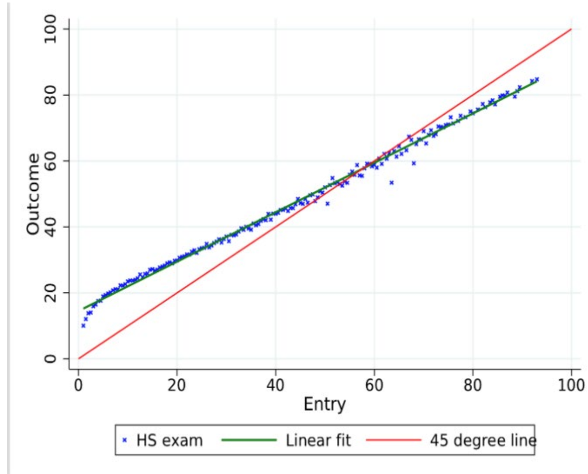
Oregon: Graduation +1



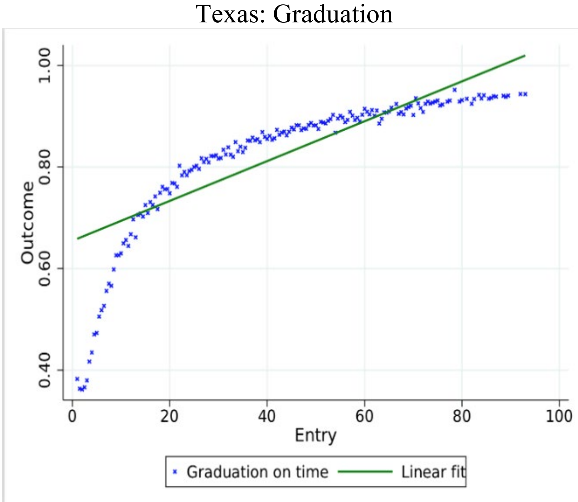
Texas: Grade-8 Math & ELA



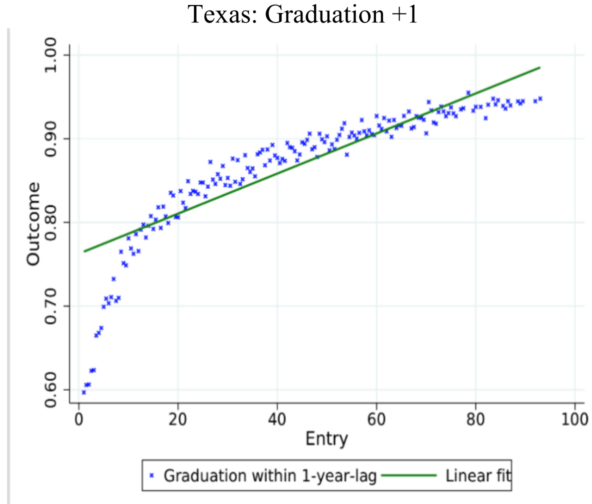
Texas: HS Test



Texas: Graduation

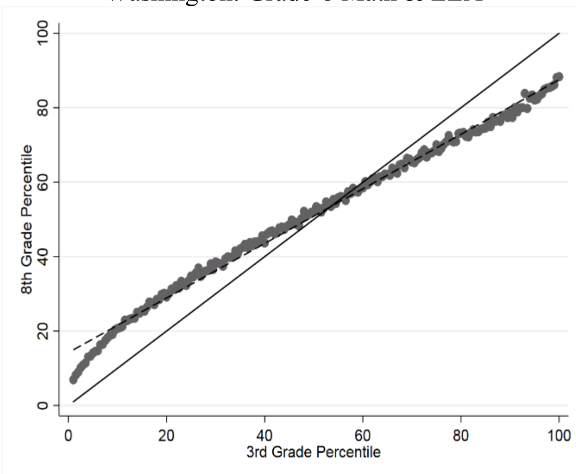


Texas: Graduation +1

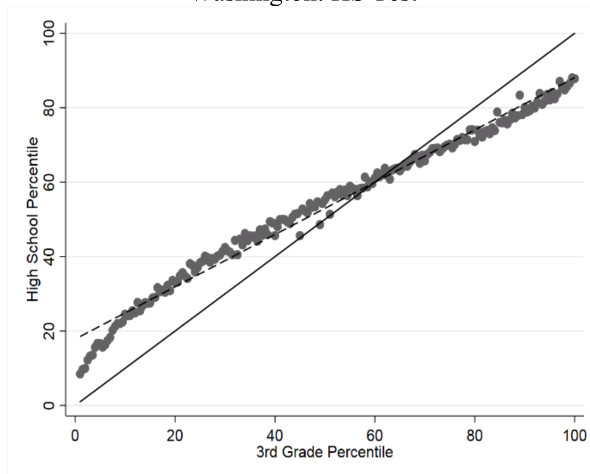


Note: Vertical and horizontal axes are scaled from 0-100 in percentiles.

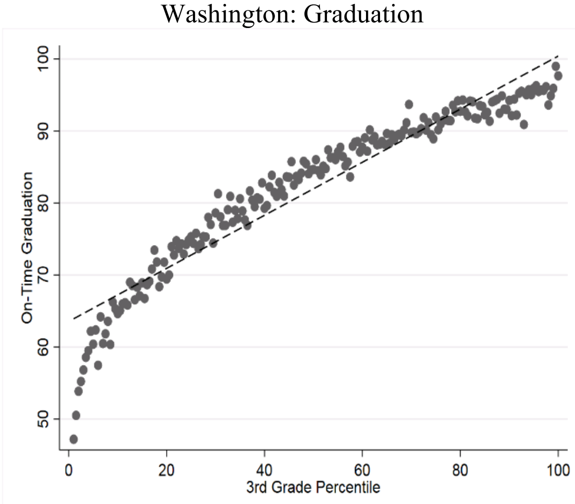
Washington: Grade-8 Math & ELA



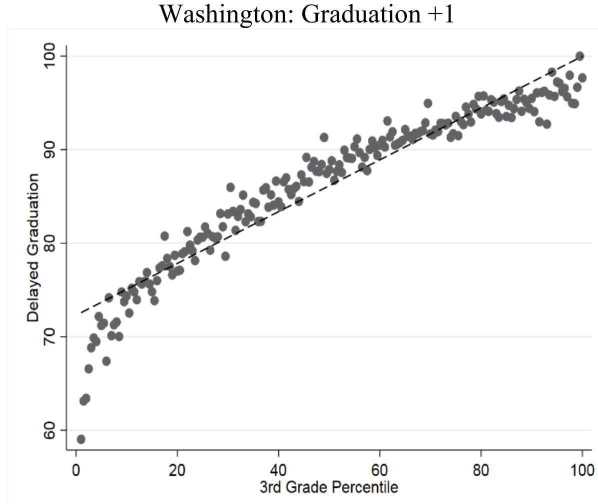
Washington: HS Test



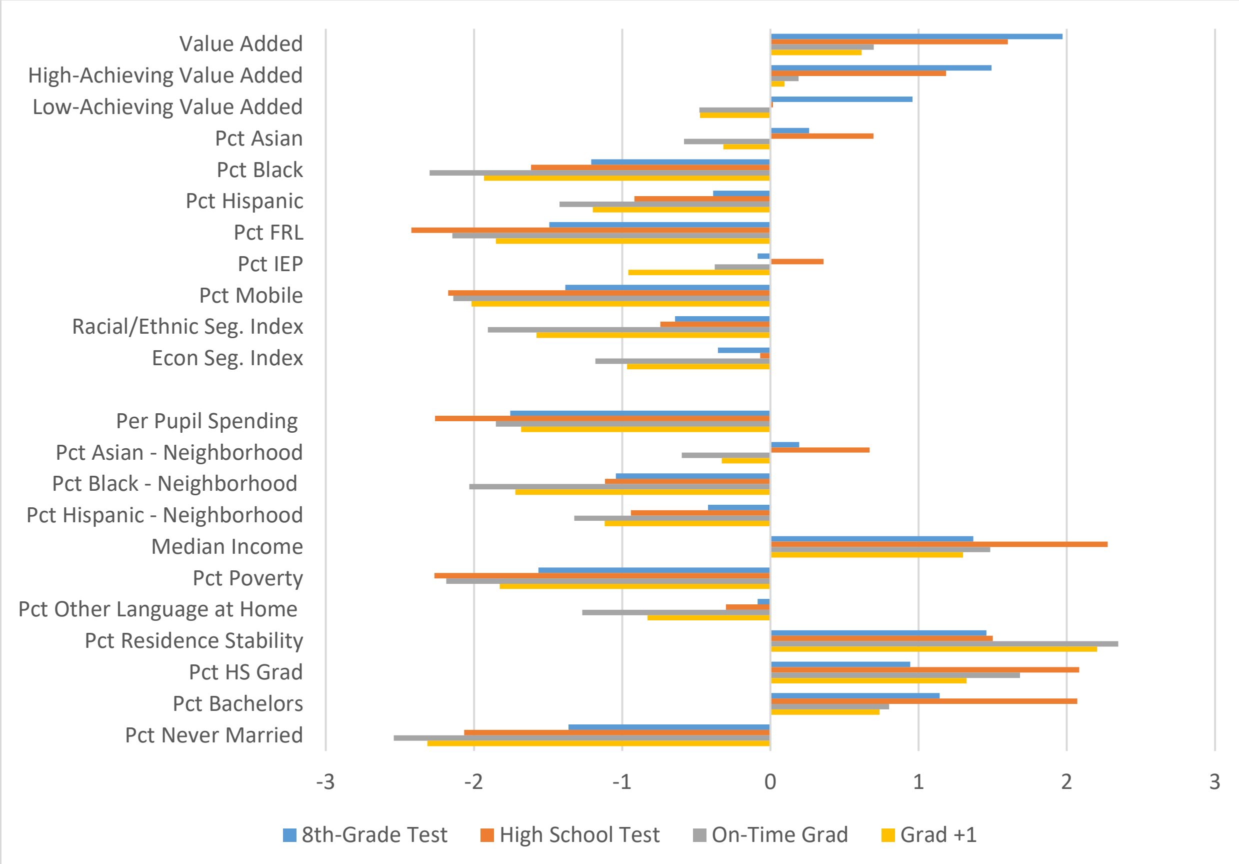
Washington: Graduation



Washington: Graduation +1

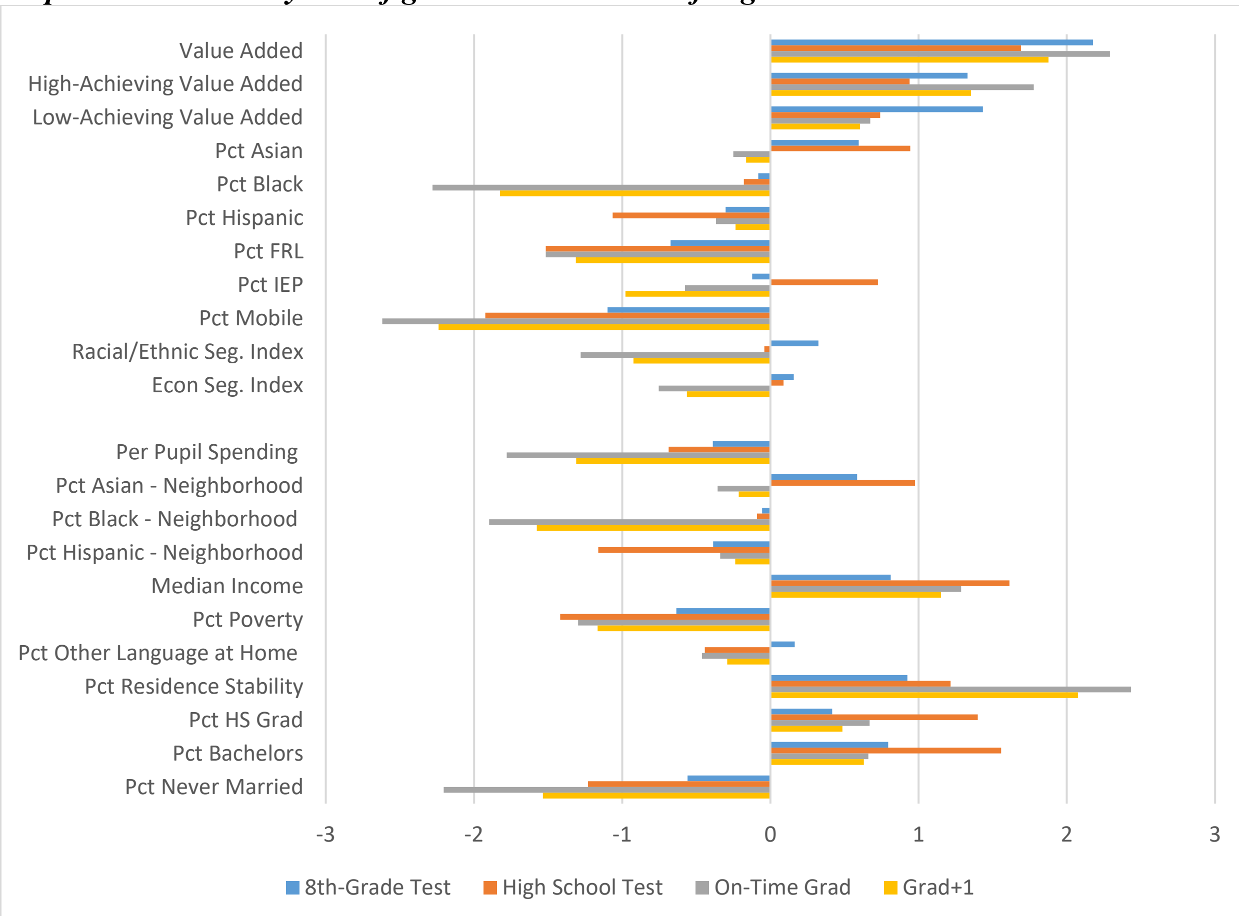


Appendix Figure A2. Correlates of $\bar{O}_{25d}$ for each outcome, where $\bar{O}_{25d}$ is estimated for Black students only. This figure is an extension of Figure 7 in the main text.



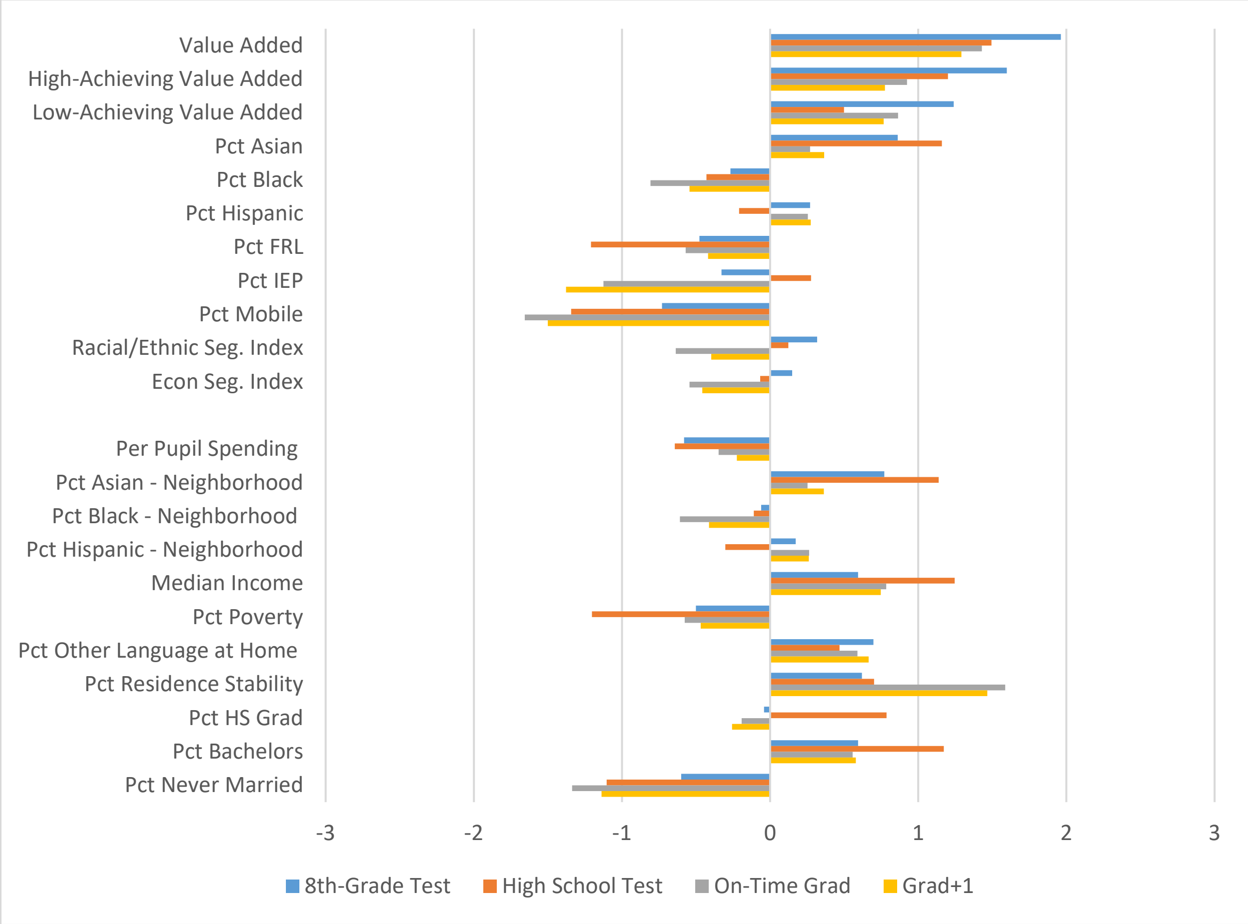
Notes: The notes to Figure 7 apply. We omit the state-by-state results underlying this figure for brevity.

Appendix Figure A3. Correlates of $\bar{O}_{25d}$ for each outcome, where $\bar{O}_{25d}$ is estimated for Hispanic students only. This figure is an extension of Figure 7 in the main text.



Notes: The notes to Figure 7 apply. We omit the state-by-state results underlying this figure for brevity.

Appendix Figure A4. Correlates of $\bar{O}_{25d}$ for each outcome, where $\bar{O}_{25d}$ is estimated for FRL students only. This figure is an extension of Figure 7 in the main text.



Notes: The notes to Figure 7 apply. We omit the state-by-state results underlying this figure for brevity.

Appendix Table A1. Percent of entry cohorts that attend a charter school at panel entry (Grade-3).

	Cohort Years	Percent Charter Enrollment (Grade-3)
Georgia	2007-2009	2.0
Massachusetts	2007-2008	2.2
Michigan	2006-2009	7.7
Missouri	2006-2009	1.3
Oregon	2006-2008	1.3
Texas	2006-2009	1.8
Washington	2006-2008	0.0

Notes: There were no charter schools in WA during the panel-entry years.

Appendix Table A2. State-by-state adjusted and unadjusted correlations of $\bar{O}_{25d}$ across outcomes. These state-by-state results undergird the summary information in Table 9 in the main text.

Georgia

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.77	1.00		
Grad	0.23	0.25	1.00	
Grad +1	0.19	0.13	1.00	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.72	1.00		
Grad	0.21	0.23	1.00	
Grad +1	0.18	0.11	0.90	1.00

Massachusetts

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.98	1.00		
Grad	0.03	0.03	1.00	
Grad +1	0.04	0.04	1.00	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.90	1.00		
Grad	0.03	0.03	1.00	
Grad +1	0.03	0.03	0.99	1.00

Michigan

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.84	1.00		
Grad	0.50	0.46	1.00	
Grad +1	0.50	0.46	1.00	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.74	1.00		
Grad	0.43	0.39	1.00	
Grad +1	0.43	0.39	0.96	1.00

Missouri

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.78	1.00		
Grad	0.34	0.28	1.00	
Grad +1	0.39	0.32	1.00	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.61	1.00		
Grad	0.25	0.20	1.00	
Grad +1	0.28	0.23	0.80	1.00

Oregon

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test		1.00		
Grad	0.47		1.00	
Grad +1	0.47		1.00	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test		1.00		
Grad	0.40		1.00	
Grad +1	0.40		0.98	1.00

Texas

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.92	1.00		
Grad	0.19	0.23	1.00	
Grad +1	0.22	0.27	0.83	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.81	1.00		
Grad	0.18	0.25	1.00	
Grad +1	0.19	0.25	0.83	1.00

Washington

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.76	1.00		
Grad	0.27	0.33	1.00	
Grad +1	0.21	0.24	1.00	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.67	1.00		
Grad	0.22	0.27	1.00	
Grad +1	0.17	0.19	0.93	1.00

Notes: The adjusted correlations are best interpreted as upper bounds because they assume estimation error in the estimates of $\bar{O}_{25d}$ 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 no HS test results are available in Oregon.

Appendix Table A3. State-by-state adjusted and unadjusted correlations of $\bar{O}_{25d}$ across cohorts for each outcome. These state-by-state results undergird the summary information in Table 10 in the main text.

<u>*Grade-8 Test*</u>

Georgia

Adjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.70	1.00	
2009		0.63	0.75	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.63	1.00	
2009		0.57	0.67	1.00

Massachusetts

Adjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.69	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007		1.00		
2008		0.64	1.00	
2009				

Michigan

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.71	1.00		
2008	0.64	0.80	1.00	
2009	0.56	0.70	0.81	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.58	1.00		
2008	0.51	0.64	1.00	
2009	0.44	0.54	0.63	1.00

Missouri

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.56	1.00		
2008	0.47	0.56	1.00	
2009	0.27	0.46	0.47	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.45	1.00		
2008	0.38	0.44	1.00	
2009	0.21	0.36	0.37	1.00

Oregon

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.67	1.00		
2008	0.45	0.72	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.61	1.00		
2008	0.41	0.66	1.00	
2009				

Texas

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.84	1.00		
2008	0.73	0.79	1.00	
2009	0.63	0.74	0.88	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.71	1.00		
2008	0.63	0.71	1.00	
2009	0.55	0.65	0.76	1.00

Washington

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.56	1.00		
2008	0.59	0.52	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.50	1.00		
2008	0.52	0.46	1.00	
2009				

HS Test

Georgia

Adjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.71	1.00	
2009		0.59	0.71	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.61	1.00	
2009		0.52	0.61	1.00

Massachusetts

Adjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.70	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.65	1.00	
2009				

Michigan

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.71	1.00		
2008	0.75	0.79	1.00	
2009	0.65	0.75	0.88	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.49	1.00		
2008	0.51	0.59	1.00	
2009	0.46	0.56	0.68	1.00

Missouri

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.59	1.00		
2008	0.45	0.57	1.00	
2009	0.45	0.38	0.38	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.41	1.00		
2008	0.32	0.41	1.00	
2009	0.31	0.27	0.27	1.00

Oregon

Adjusted Correlations				
	2006	2007	2008	2009
2006				
2007				
2008				
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006				
2007				
2008				
2009				

Texas

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.83	1.00		
2008	0.77	0.79	1.00	
2009	0.68	0.75	0.89	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.69	1.00		
2008	0.64	0.70	1.00	
2009	0.57	0.64	0.75	1.00

Washington

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.58	1.00		
2008	0.63	0.59	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.51	1.00		
2008	0.55	0.52	1.00	
2009				

Graduation

Georgia

Adjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.62	1.00	
2009		0.51	0.60	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.50	1.00	
2009		0.41	0.48	1.00

Massachusetts

Adjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.57	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.47	1.00	
2009				

Michigan

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.96	1.00		
2008	0.91	0.93	1.00	
2009	0.89	0.90	0.96	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.87	1.00		
2008	0.83	0.84	1.00	
2009	0.81	0.81	0.86	1.00

Missouri

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.33	1.00		
2008	0.41	0.47	1.00	
2009	0.21	0.48	0.41	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.20	1.00		
2008	0.25	0.29	1.00	
2009	0.13	0.30	0.25	1.00

Oregon

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.64	1.00		
2008	0.16	0.17	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.53	1.00		
2008	0.13	0.14	1.00	
2009				

Texas

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.61	1.00		
2008	0.55	0.58	1.00	
2009	0.50	0.56	0.63	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.40	1.00		
2008	0.37	0.39	1.00	
2009	0.37	0.39	0.41	1.00

Washington

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.23	1.00		
2008	0.37	0.53	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.20	1.00		
2008	0.31	0.45	1.00	
2009				

Graduation +1

Georgia

Adjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.72	1.00	
2009		0.63	0.74	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.56	1.00	
2009		0.49	0.58	1.00

Massachusetts

Adjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.55	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006				
2007		1.00		
2008		0.46	1.00	
2009				

Michigan

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.96	1.00		
2008	0.92	0.93	1.00	
2009	0.88	0.89	0.97	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.87	1.00		
2008	0.84	0.84	1.00	
2009	0.80	0.81	0.86	1.00

Missouri

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.31	1.00		
2008	0.31	0.48	1.00	
2009	0.24	0.44	0.38	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.19	1.00		
2008	0.19	0.29	1.00	
2009	0.15	0.27	0.23	1.00

Oregon

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.63	1.00		
2008	0.16	0.19	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.52	1.00		
2008	0.13	0.15	1.00	
2009				

Texas

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.54	1.00		
2008	0.47	0.61	1.00	
2009	0.46	0.51	0.45	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.36	1.00		
2008	0.31	0.39	1.00	
2009	0.33	0.34	0.27	1.00

Washington

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.25	1.00		
2008	0.19	0.45	1.00	
2009				

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.21	1.00		
2008	0.15	0.38	1.00	
2009				

Notes: 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. Oregon did not offer a high school test taken in a (near) universal grade for the years we used in the analysis, so no high school test results are available in Oregon.

Appendix Table A4. State-by-state regression output for the correlates of academic mobility, corresponding to Figure 7. All coefficients are from univariate regressions where the independent variable is standardized to have a mean of zero and variance of one within states. The dependent variables in these regressions are estimates of O25 taken from models in which we account for measurement error using our primary approach of averaging the 3rd grade ranks in math and ELA with the addition of the EIV correction.

	Grade-8 Test							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	1.40*	1.86*	3.83*	2.19*	2.31*	2.98*	1.16*	2.25
High-Achieving VA	0.80*	0.99*	3.10*	1.79*	2.37*	1.70*	0.87*	1.66
Low-Achieving VA	1.21*	2.10*	-0.76	0.11	2.17*	3.74*	0.54*	1.30
Pct Asian	0.45*	1.95*	0.97*	0.76*	0.16	1.79*	0.93*	1.00
Pct Black	-1.55*	-0.71*	-1.74*	-1.05*	0.14	-0.50*	-0.21	-0.80
Pct Hispanic	0.93*	-1.88*	-0.74*	-0.30	1.62*	-1.59*	-0.30	-0.32
Pct FRL	-1.44*	-2.53*	-2.13*	-1.58*	0.67	-2.21*	-1.10*	-1.47
Pct IEP	0.58*	-0.54	-2.02*	0.67	-0.70*	2.19*	-0.14	0.01
Pct Mobile	-0.94*	-2.60*	-2.03*	-1.91*	1.34*	-3.97*	-0.78*	-1.56
School Seg. Index	-0.56*	-0.42	-0.30*	-0.56*	1.06*	-0.58*	0.52*	-0.12
Economic Seg. Index	-0.49*	-1.04*	0.85*	-0.21*	1.39*	0.59*	-0.42*	0.10
NCES Data Elements								
Per Pupil Spending	-1.27*	0.18	-1.60*	-0.53*	-0.45	-2.98*	-0.52*	-1.02
Pct Asian - Neighborhood	0.44*	2.21*	0.75*	0.73*	-0.09	1.18*	1.19*	0.92
Pct Black - Neighborhood	-1.17*	-0.94*	-0.91*	-0.81*	0.32	-0.08*	-0.23	-0.55
Pct Hispanic - Neighborhood	0.90*	-1.88*	-1.02*	-0.37*	1.52*	-1.62*	-0.40*	-0.41
Median Income	0.83*	3.33*	1.82*	1.06*	-0.64	2.28*	1.44*	1.45
Pct Poverty	-1.40*	-2.34*	-1.58*	-1.65*	0.71*	-1.87*	-1.02*	-1.31
Pct Other Language at Home	0.87*	-0.77*	0.54*	0.03	1.39*	-0.95*	0.31	0.20
Pct Residence Stability	1.07*	1.32*	1.89*	1.12*	0.35	0.65*	0.44*	0.98
Pct HS Grad	0.04*	2.36*	1.38*	1.61*	-1.62*	1.71*	0.50*	0.85
Pct Bachelors	0.21*	3.60*	1.87*	0.88*	-0.19	2.26*	1.55*	1.45
Pct Never Married	-1.75*	-2.72*	-1.28*	-1.22*	0.67	-2.09*	-1.10*	-1.36

	HS Test							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	0.64*	1.55*	3.94*	1.58*	Not Applicable	2.32*	0.77*	1.80
High-Achieving VA	0.04*	0.76*	3.70*	1.18*		0.88*	0.43*	1.17
Low-Achieving VA	0.41*	1.75*	-2.30*	-0.91*		3.40*	-0.04	0.39
Pct Asian	0.42*	1.97*	1.01*	1.42*		2.45*	1.41*	1.45
Pct Black	-1.59*	-1.02*	-2.54*	-1.07*		-0.96*	-0.20	-1.23
Pct Hispanic	0.95*	-2.49*	-1.06*	-0.68*		-2.14*	-1.08*	-1.08
Pct FRL	-2.00*	-3.33*	-2.78*	-2.75*		-3.10*	-2.31*	-2.71
Pct IEP	0.28*	-0.95*	-1.01	2.08*		3.55*	-0.23	0.62
Pct Mobile	-0.61*	-2.96*	-2.90*	-2.90*		-5.52	-0.89*	-2.63
School Seg. Index	-0.33*	-0.90*	-0.38*	-0.62*		-0.91*	0.20	-0.49
Economic Seg. Index	-0.30*	-1.38*	0.99*	-0.10		0.93*	-0.77*	-0.11
NCES Data Elements								
Per Pupil Spending	-1.47*	0.04	-1.82*	0.13		-4.10*	-0.83*	-1.34
Pct Asian - Neighborhood	0.45*	2.46*	0.93*	1.47*		1.60*	1.85*	1.46
Pct Black - Neighborhood	-1.16*	-1.25*	-1.14*	-0.81*		-0.33*	-0.17	-0.81
Pct Hispanic - Neighborhood	0.98*	-2.48*	-1.31*	-0.71*		-2.21*	-1.16*	-1.15
Median Income	1.27*	4.24*	2.25*	2.21*		3.09*	2.61*	2.61
Pct Poverty	-1.99*	-3.08*	-1.95*	-2.74*		-2.74*	-1.94*	-2.41
Pct Other Language at Home	0.96*	-1.30*	0.55*	0.24		-1.38*	-0.10	-0.17
Pct Residence Stability	0.60*	1.88*	2.18*	2.01*		0.86*	0.31	1.31
Pct HS Grad	0.50*	3.17*	1.89*	2.91*		2.58*	1.35*	2.07
Pct Bachelors	0.57*	4.46*	2.30*	1.93*		3.13*	2.47*	2.48
Pct Never Married	-1.96*	-3.57*	-1.58*	-1.53*		-2.96*	-1.92*	-2.25

	On-Time Graduation							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	0.19*	0.93*	4.31*	1.16*	0.75	1.53*	1.14*	1.43
High-Achieving VA	0.24*	0.65*	1.78	0.82*	0.11	0.57*	0.27	0.63
Low-Achieving VA	-0.23*	0.58*	-1.85	-1.29*	0.96*	2.39*	0.70*	0.18
Pct Asian	-0.22*	0.46	0.89*	0.58*	0.16	1.52*	0.05	0.49
Pct Black	-1.93*	-3.20*	-1.90	-2.66*	-0.67	-1.75*	-1.84*	-1.99
Pct Hispanic	-0.02*	-4.39*	-1.92*	-1.90*	1.33*	-1.10*	-0.75*	-1.25
Pct FRL	-1.66*	-5.22*	-3.48*	-3.23*	-1.59*	-2.41*	-2.65*	-2.89
Pct IEP	1.40*	-2.25*	-5.93*	2.28*	-1.06*	3.23*	-0.73*	-0.44
Pct Mobile	-1.95*	-3.15*	-4.61*	-4.65*	-1.63*	-6.30*	-0.01	-3.19
School Seg. Index	-1.00*	-3.52*	-0.20	-1.64*	-0.30	-1.56*	-0.92*	-1.31
Economic Seg. Index	-1.08*	-1.08*	1.32*	-0.75*	-0.85*	-0.42*	-1.22*	-0.58
NCES Data Elements								
Per Pupil Spending	-2.42*	-2.09*	-1.32	-2.43*	-0.28	-2.58*	-1.26*	-1.77
Pct Asian - Neighborhood	-0.24*	0.96*	0.69	0.35	0.32	0.97*	0.95*	0.57
Pct Black - Neighborhood	-1.67*	-3.36*	-0.79	-1.78*	-0.57	-0.78*	-1.66*	-1.52
Pct Hispanic - Neighborhood	0.09*	-4.13*	-1.86*	-1.78*	1.02*	-1.05*	-0.80*	-1.22
Median Income	1.01*	4.43*	2.83*	1.76*	1.53*	2.27*	2.79*	2.37
Pct Poverty	-1.38*	-5.07*	-2.13*	-3.17*	-1.46*	-2.02*	-2.14*	-2.48
Pct Other Language at Home	-0.24*	-3.56*	0.83	-1.73*	1.08*	-0.62*	-0.33	-0.65
Pct Residence Stability	1.25*	3.32*	3.36*	2.67*	1.89*	1.81*	1.38*	2.24
Pct HS Grad	-0.07*	4.86*	1.80*	3.08*	-0.40	1.89*	1.01*	1.74
Pct Bachelors	-0.20*	4.49*	2.50*	1.11*	0.96*	2.07*	2.28*	1.89
Pct Never Married	-2.06*	-5.24*	-1.84	-2.62*	-2.20*	-2.59*	-2.32*	-2.70

	Graduation +1							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	0.10*	0.89*	3.58*	1.21*	0.54	1.42*	0.79*	1.22
High-Achieving VA	0.05*	0.60*	1.40	0.80*	-0.05	0.73*	0.04	0.51
Low-Achieving VA	-0.14*	0.55*	-1.69	-0.69*	0.54	1.93*	0.45	0.14
Pct Asian	-0.25*	0.51	0.84*	0.44*	0.54	1.25*	0.15	0.50
Pct Black	-1.47*	-3.02*	-1.66	-2.04*	-0.36	-0.68*	-1.56*	-1.54
Pct Hispanic	-0.08*	-4.21*	-1.72*	-1.54*	1.15*	-0.20*	-0.53*	-1.02
Pct FRL	-1.28*	-5.00*	-3.03*	-2.44*	-1.73*	-1.20*	-2.28*	-2.42
Pct IEP	1.21*	-2.13*	-5.40*	1.82*	-0.89*	0.41*	-0.52*	-0.79
Pct Mobile	-1.86*	-3.06*	-4.00*	-3.73*	-1.60*	-4.47*	0.06	-2.67
School Seg. Index	-0.92*	-3.28*	-0.16	-1.32*	-0.09	-0.66*	-0.60*	-1.00
Economic Seg. Index	-0.85*	-1.03*	1.11*	-0.66*	-0.60	-0.25*	-0.96*	-0.46
NCES Data Elements								
Per Pupil Spending	-2.21*	-1.94*	-1.15	-1.84*	-0.18	-1.46*	-1.11*	-1.41
Pct Asian - Neighborhood	-0.27*	0.97*	0.72*	0.25	0.680	0.83*	0.85*	0.58
Pct Black - Neighborhood	-1.32*	-3.19*	-0.69	-1.36*	-0.26	-0.18*	-1.42*	-1.20
Pct Hispanic - Neighborhood	-0.04*	-3.98*	-1.66*	-1.46*	0.82*	-0.15*	-0.59*	-1.01
Median Income	0.87*	4.27*	2.47*	1.35*	1.85*	1.27*	2.41*	2.07
Pct Poverty	-1.07*	-4.87*	-1.89*	-2.42*	-1.57*	-0.89*	-1.89*	-2.09
Pct Other Language at Home	-0.32*	-3.39*	0.86	-1.43*	1.11*	0.35*	-0.19	-0.43
Pct Residence Stability	0.94*	3.20*	2.96*	2.03*	1.95*	1.47*	1.05*	1.94
Pct HS Grad	0.02*	4.68*	1.56*	2.33*	-0.21	0.74*	0.83*	1.42
Pct Bachelors	-0.19*	4.35*	2.19*	0.81*	1.35*	1.27*	1.94*	1.67
Pct Never Married	-1.57*	-5.02*	-1.62	-2.08*	-2.32*	-1.32*	-2.00*	-2.28

Notes: The notes to Figure 7 apply. Statistical significance at the 5 percent level for each coefficient is denoted by *. Standard errors are suppressed for brevity. The symbol † is to denote that statistical significance is not reported for the "All (avg)" values because they are not directly generated from a regression (they are average values of the state-by-state regression coefficients).

Appendix B: Replication of Results Using Alternative Measurement Error Corrections

In this appendix we replicate our full analysis using the two alternative approaches to correcting for measurement error discussed in the text: (1) the instrumental variables approach where we use the 3rd grade math test to set the initial rank and instrument using the 3rd grade ELA rank, and (2) the test-averaging approach where we use the average of the math and ELA ranks to set the initial rank, but unlike in the main text, we do not make the additional errors-in-variables correction for test measurement error. We refer to this latter approach as the "uncorrected" approach for ease of presentation, but this is technically a misnomer because the averaging of the initial ranks is a partial correction.

Theoretically, the IV models should make the most comprehensive measurement error correction, followed by the EIV models presented in the main text. The uncorrected models should make the weakest correction. The magnitudes of our estimates of β using the different approaches are consistent with this expectation. That is, with some exceptions at low levels of data aggregation (e.g., in individual districts), within a given model, β is largest when estimated by IV, next-largest when estimated by averaging with the EIV correction, and lowest in the uncorrected models. A caveat is that the IV estimates of β may also be inflated if the exclusion restriction is violated, as discussed in the main text—for this reason, we interpret the IV estimates as upper bounds on β.

In addition to the generally larger estimates of β using the IV models (and smaller estimates of β from the uncorrected models), the other main takeaway from the appendix results is that our comparative findings are mostly similar regardless of which error correction we use. An exception noted in the main text is that some of the test-based academic mobility gaps between student subgroups are smaller using the IV approach. However, the IV-based graduation gaps are similar in magnitude, and the gaps are directionally aligned using all three error corrections regardless of the outcome. The general similarity of results across methods suggests that error variance in the initial ranks is not an overwhelming factor driving our comparative findings, and/or that heterogeneity in the error-variance across student groups and districts is not large enough to be consequential for the comparisons.

An additional qualification to the IV results is that we impose a sample restriction on the district-level portion of the IV analysis: we exclude districts with fewer than 100 students summed across entry cohorts in each state. This restriction is based on an *ad hoc* investigation of the feasibility of running the IV models at scale for each of the more than 3000 school districts covered by our analysis (per Table 1). The reductions of the district- and student-level sample sizes for the IV analysis in each state due to this restriction are reported in Appendix Table B1 (the loss of districts is substantial in some states, but because it is only small districts that are excluded, the loss of students is negligible in all states).

The tables that follow replicate our results from the main text using the two alternative measurement error corrections. Table numbers ending with the letter “a” show results using the IV approach and table numbers ending in “b” show results using the uncorrected models. Each appendix table title references the table or figure to which it corresponds in the main text.

Appendix Table B1. Percent of districts and students excluded from the state samples in the district-level IV analysis when we exclude districts with N < 100.

	Percent of Districts Excluded	Percent of Students Excluded
Georgia	2.2%	0.1%
Massachusetts	17.7	2.0
Michigan	13.3	1.5
Missouri	26.6	3.6
Oregon	40.0	2.0
Texas	29.8	1.4
Washington	33.8	1.7

Notes: As indicated in the text, the N ≥ 100 condition applies to the sum of students across all cohorts. Note that the restricted sample is used for the district-level IV models only. All analyses above the district level include all districts regardless of method.

Appendix Table B2a. Statewide estimates of $\boldsymbol{\beta}$ and $\bar{O}_{25}$ for each outcome corresponding to Table 4. We account for measurement error in the 3rd-grade tests by using the 3rd-grade math rank as the initial rank variable and instrumenting for it with the 3rd-grade ELA rank.

	Grade-8 Test		HS Test		Grad		Grad +1	
	β	O25	β	O25	β	O25	β	O25
All (Avg)	**0.94**	**27.27**	**0.97**	**26.94**	**0.40**	**74.59**	**0.30**	**79.72**
GA	0.92	28.21	0.96	27.26	0.57	64.80	0.43	72.19
MA	0.95	27.17	0.93	27.41	0.22	88.23	0.21	88.74
MI	0.93	26.64	0.95	27.02	0.39	74.17	0.33	78.60
MO	0.96	25.98	0.95	26.49	0.28	82.95	0.18	88.14
OR	0.93	27.07	Not Applicable		0.38	69.94	0.33	72.68
TX	1.00	27.69	1.06	26.60	0.51	71.87	0.29	80.43
WA	0.89	28.13	0.95	26.85	0.43	70.19	0.32	77.23

Notes: The notes to Table 4 in the main text apply. Our estimates of β are especially high in Texas, and in the test-based models, a value above 1.0 is technically infeasible (because it implies a negative α). We believe this technical issue may be due to excess noise in the upper tail of the test distribution in Texas combined with the fact that the IV models shrink the range of entry ranks through the first stage and thus must extrapolate the linear relationship to the edges.

Appendix Table B2b. Statewide estimates of β and $\bar{O}_{25}$ for each outcome corresponding to Table 4. We account for measurement error in the 3rd-grade tests by averaging the ELA and math tests but make no further corrections.

	Grade-8 Test		HS Test		Grad		Grad +1	
	β	O25	β	O25	β	O25	β	O25
All (Avg)	**0.79**	**30.99**	**0.79**	**31.69**	**0.33**	**76.16**	**0.25**	**80.88**
GA	0.81	31.05	0.82	30.92	0.49	66.72	0.37	73.64
MA	0.79	30.75	0.78	30.81	0.18	89.01	0.17	89.49
MI	0.78	30.61	0.75	33.12	0.33	75.79	0.28	80.01
MO	0.83	29.31	0.78	31.09	0.24	83.94	0.16	88.80
OR	0.75	31.22	Not Applicable		0.31	71.59	0.27	74.12
TX	0.80	32.70	0.80	33.16	0.41	74.29	0.25	81.70
WA	0.77	31.26	0.78	31.02	0.37	71.78	0.27	78.41

Notes: The notes to Table 4 apply, except the measurement error correction is different.

Appendix Table B3a. Statewide academic mobility estimates by race/ethnicity corresponding to Table 5. We account for measurement error in the 3rd-grade tests by using the 3rd-grade math rank as the initial rank variable and instrumenting for it with the 3rd-grade ELA rank.

	Grade-8 Test			HS Test			Grad			Grad +1		
Student Group: Asian												
	α	β	O25	α	β	O25	α	β	O25	α	β	O25
All (Avg)	**11.27**	**0.91**	**33.93**	**10.14**	**0.93**	**33.37**	**77.63**	**0.23**	**83.48**	**83.16**	**0.17**	**87.39**
GA	13.94	0.87	35.80	8.69	0.92	31.80	63.65	0.40	73.66	72.39	0.30	79.82
MA	12.05	0.91	34.74	14.59	0.87	36.45	90.66	0.11	93.31	91.36	0.1	93.8
MI	10.18	0.91	32.99	6.91	0.98	31.43	79.49	0.20	84.59	84.23	0.16	88.14
MO	8.58	0.95	32.35	9.71	0.92	32.78	85.55	0.15	89.39	90.26	0.10	92.75
OR	9.08	0.89	31.42	Not Applicable			75.06	0.24	81.01	77.88	0.2	82.99
TX	13.23	0.95	36.98	12.45	0.98	36.89	78.87	0.22	84.27	88.25	0.09	90.53
WA	11.84	0.86	33.26	8.51	0.89	30.86	70.11	0.32	78.13	77.78	0.24	83.71
Student Group: Black												
	α	β	O25	α	β	O25	α	β	O25	α	β	O25
All (Avg)	**4.48**	**0.89**	**26.70**	**3.80**	**0.92**	**26.71**	**61.07**	**0.45**	**72.38**	**69.74**	**0.33**	**78.06**
GA	5.81	0.89	27.99	3.78	0.93	27.08	49.81	0.64	65.74	61.79	0.46	73.39
MA	6.27	0.87	28.00	7.08	0.83	27.74	79.8	0.24	85.77	81.09	0.22	86.6
MI	5.44	0.84	26.48	1.03	0.90	23.56	60.72	0.44	71.77	66.67	0.39	76.35
MO	3.57	0.86	25.06	4.35	0.88	26.23	68.95	0.37	78.08	78.27	0.24	84.23
OR	-0.08	0.93	23.14	Not Applicable			55.25	0.39	65.07	61.53	0.32	69.63
TX	5.23	0.96	29.18	3.41	1.02	28.85	57.81	0.62	73.22	73.26	0.34	81.85
WA	5.14	0.87	27.02	3.17	0.94	26.77	55.13	0.47	66.99	65.57	0.35	74.38
Student Group: Hispanic												
	α	β	O25	α	β	O25	α	β	O25	α	β	O25
All (Avg)	**5.45**	**0.89**	**27.58**	**3.28**	**0.92**	**26.38**	**63.36**	**0.41**	**73.59**	**71.36**	**0.30**	**78.98**
GA	6.43	0.90	28.82	2.67	0.96	26.55	47.99	0.59	62.68	59.97	0.43	70.74
MA	5.58	0.83	26.23	5.7	0.79	25.49	75	0.3	82.51	76.03	0.29	83.23
MI	4.45	0.91	27.22	0.29	1.00	25.40	64.09	0.35	72.87	69.89	0.30	77.39
MO	4.81	0.91	27.45	5.92	0.89	28.26	74.57	0.28	81.63	82.42	0.18	87.02
OR	5.96	0.87	27.65	Not Applicable			65.29	0.32	73.25	68.74	0.28	75.76
TX	3.57	0.93	26.92	0.52	1.00	25.46	56.13	0.60	71.03	71.76	0.35	80.38
WA	7.37	0.86	28.79	4.59	0.9	27.09	60.48	0.43	71.14	70.74	0.3	78.31
Student Group: White												
	α	β	O25	α	β	O25	α	β	O25	α	β	O25
All (Avg)	**3.32**	**0.95**	**27.12**	**1.99**	**0.99**	**26.66**	**66.08**	**0.37**	**75.26**	**72.70**	**0.29**	**79.87**
GA	4.20	0.94	27.75	3.23	0.97	27.52	49.96	0.56	63.86	59.61	0.44	70.52
MA	2.46	0.97	26.65	3.93	0.94	27.31	87.51	0.15	91.28	88	0.14	91.62
MI	2.75	0.94	26.21	-0.89	1.03	24.79	67.08	0.35	75.88	72.79	0.30	80.23
MO	1.76	0.97	25.92	2.07	0.96	26.10	79.67	0.23	85.40	86.43	0.15	90.08
OR	3.55	0.93	26.91	Not Applicable			58.57	0.41	68.75	62.35	0.36	71.44
TX	3.51	1.01	28.85	1.14	1.07	27.86	60.21	0.45	71.45	70.27	0.31	77.94
WA	5.01	0.9	27.52	2.47	0.96	26.39	59.55	0.43	70.21	69.44	0.31	77.29

Notes: The notes to Table 5 apply, except the measurement error correction is different. There are several instances of estimates of β at or above 1.0 in MI and TX.

Appendix Table B3b. Statewide academic mobility estimates by race/ethnicity corresponding to Table 5. We account for measurement error in the 3rd-grade tests by averaging the ELA and math tests but make no further corrections.

	Grade-8 Test			HS Test			Grad			Grad +1		
Student Group: Asian												
	α	β	O25	α	β	O25	α	β	O25	α	β	O25
All (Avg)	**21.57**	**0.77**	**40.87**	**22.74**	**0.76**	**41.76**	**80.09**	**0.20**	**85.15**	**84.77**	**0.15**	**88.49**
GA	22.87	0.76	41.87	19.92	0.78	39.34	67.97	0.34	76.58	75.53	0.26	81.96
MA	21.7	0.77	40.99	23.49	0.75	42.26	91.75	0.09	94.02	92.38	0.08	94.46
MI	21.52	0.78	41.10	21.23	0.81	41.41	82.02	0.18	86.41	85.97	0.14	89.41
MO	18.25	0.83	38.94	21.26	0.77	40.48	86.87	0.14	90.31	90.94	0.09	93.24
OR	17.25	0.78	36.73	Not Applicable			76.76	0.22	82.16	79.27	0.19	83.94
TX	30.65	0.73	48.78	32.82	0.71	50.52	82.29	0.17	86.63	89.44	0.08	91.38
WA	18.78	0.76	37.71	17.71	0.75	36.56	72.94	0.28	79.93	79.89	0.21	85.04
Student Group: Black												
	α	β	O25	α	β	O25	α	β	O25	α	β	O25
All (Avg)	**9.55**	**0.73**	**27.66**	**9.87**	**0.72**	**27.93**	**63.77**	**0.37**	**72.92**	**71.73**	**0.27**	**78.45**
GA	9.89	0.76	28.77	9.35	0.76	28.38	53.08	0.53	66.43	64.17	0.39	73.89
MA	11.32	0.7	28.84	11.77	0.67	28.52	81.11	0.19	85.99	82.3	0.18	86.79
MI	9.93	0.67	26.67	6.98	0.68	24.05	63.21	0.35	71.91	68.91	0.30	76.49
MO	7.01	0.74	25.49	9.49	0.71	27.25	70.68	0.31	78.35	79.39	0.20	84.40
OR	7.56	0.71	25.29	Not Applicable			58.51	0.3	65.98	64.28	0.25	70.41
TX	11.90	0.76	30.78	12.53	0.75	31.32	62.26	0.48	74.30	75.51	0.27	82.38
WA	9.27	0.74	27.78	9.08	0.76	28.06	57.57	0.4	67.47	67.52	0.29	74.77
Student Group: Hispanic												
	α	β	O25	α	β	O25	α	β	O25	α	β	O25
All (Avg)	**12.09**	**0.73**	**30.42**	**11.58**	**0.74**	**29.99**	**66.54**	**0.34**	**75.00**	**73.68**	**0.25**	**79.98**
GA	13.22	0.77	32.49	11.74	0.78	31.26	52.82	0.50	65.25	63.54	0.36	72.63
MA	10.53	0.69	27.88	10.33	0.67	27.05	76.69	0.26	83.09	77.69	0.24	83.79
MI	11.34	0.75	30.05	9.54	0.78	29.09	66.57	0.29	73.90	71.99	0.25	78.26
MO	11.00	0.78	30.51	13.88	0.73	32.05	76.62	0.24	82.63	83.83	0.15	87.70
OR	13.33	0.67	30.2	Not Applicable			67.9	0.25	74.17	71.08	0.22	76.58
TX	13.22	0.74	31.74	13.41	0.73	31.70	62.25	0.47	74.08	75.05	0.28	82.06
WA	11.96	0.73	30.1	10.59	0.73	28.78	62.95	0.36	71.85	72.55	0.25	78.82
Student Group: White												
	α	β	O25	α	β	O25	α	β	O25	α	β	O25
All (Avg)	**12.83**	**0.78**	**32.30**	**13.86**	**0.78**	**33.21**	**69.74**	**0.30**	**77.28**	**75.53**	**0.24**	**81.43**
GA	12.92	0.80	32.99	13.96	0.80	33.87	55.07	0.47	66.93	63.62	0.37	72.93
MA	11.82	0.78	31.39	12.75	0.76	31.79	88.84	0.12	91.95	89.27	0.12	92.27
MI	13.15	0.76	32.19	13.04	0.79	32.68	70.90	0.29	78.08	76.08	0.24	82.12
MO	10.13	0.82	30.53	13.09	0.76	32.19	81.70	0.19	86.51	87.76	0.12	90.82
OR	13.16	0.75	31.82	Not Applicable			62.77	0.32	70.89	66.07	0.29	73.33
TX	15.59	0.79	35.41	17.23	0.78	36.69	65.37	0.36	74.25	73.53	0.25	79.71
WA	13.04	0.75	31.77	13.09	0.76	32.03	63.55	0.35	72.35	72.38	0.26	78.85

Notes: The notes to Table 5 apply, except the measurement error correction is different.

Appendix Table B4a. Statewide academic mobility estimates by FRL status corresponding to Table 6. We account for measurement error in the 3rd-grade tests by using the 3rd-grade math rank as the initial rank variable and instrumenting for it with the 3rd-grade ELA rank.

	Grade-8 Test			HS Test			Grad			Grad +1		
Student Group: FRL												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**4.77**	**0.87**	**26.44**	**2.59**	**0.91**	**25.19**	**60.29**	**0.40**	**70.23**	**68.53**	**0.29**	**75.84**
GA	5.60	0.87	27.41	3.12	0.92	26.03	46.15	0.61	61.37	58.09	0.44	69.20
MA	5.35	0.82	25.81	5.85	0.78	25.23	76.44	0.24	82.51	77.42	0.23	83.21
MI	4.50	0.86	25.94	-0.29	0.96	23.69	58.58	0.38	67.96	64.63	0.33	72.93
MO	2.85	0.90	25.35	2.72	0.90	25.18	71.59	0.30	79.06	80.17	0.19	84.99
OR	4.96	0.85	26.26	Not Applicable			58.75	0.28	65.72	62.7	0.24	68.61
TX	4.03	0.92	27	0.99	0.98	25.43	55	0.58	69.55	70.78	0.32	78.87
WA	6.13	0.85	27.28	3.12	0.9	25.59	55.52	0.4	65.45	65.94	0.28	73.05

Student Group: non-FRL												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**5.54**	**0.94**	**28.96**	**5.27**	**0.96**	**29.19**	**76.10**	**0.26**	**82.52**	**82.10**	**0.18**	**86.67**
GA	7.60	0.92	30.53	7.39	0.94	30.77	64.91	0.39	74.73	73.81	0.28	80.86
MA	4.64	0.95	28.38	6.68	0.91	29.48	92.42	0.09	94.60	92.86	0.08	94.9
MI	4.33	0.93	27.59	1.15	1.01	26.37	77.45	0.24	83.34	82.27	0.19	86.99
MO	2.95	0.96	27.05	5.21	0.94	28.62	85.84	0.16	89.90	91.41	0.09	93.74
OR	5.79	0.93	28.92	Not Applicable			71.41	0.28	78.39	74.88	0.24	80.85
TX	6.17	0.99	30.85	5.55	1.02	31.08	70.57	0.33	78.85	80.74	0.19	85.38
WA	7.33	0.88	29.41	5.64	0.93	28.79	70.13	0.31	77.82	78.71	0.21	83.97

Notes: The notes to Table 6 apply, except the measurement error correction is different. There are two instances of estimates of β above 1.0 in MI and TX.

Appendix Table B4b. Statewide academic mobility estimates by FRL status corresponding to Table 6. We account for measurement error in the 3rd-grade tests by averaging the ELA and math tests but make no further corrections.

	Grade-8 Test			HS Test			Grad			Grad +1		
Student Group: FRL												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**10.89**	**0.72**	**28.82**	**10.20**	**0.72**	**28.15**	**63.10**	**0.33**	**71.35**	**70.52**	**0.24**	**76.62**
GA	10.55	0.75	29.28	9.73	0.75	28.53	50.19	0.51	62.91	61.11	0.37	70.35
MA	10.43	0.68	27.48	10.49	0.65	26.77	77.61	0.21	82.92	78.56	0.20	83.61
MI	10.59	0.70	28.13	8.16	0.74	26.70	61.39	0.30	68.97	67.17	0.27	73.84
MO	8.17	0.77	27.46	10.11	0.72	28.19	73.55	0.25	79.84	81.48	0.16	85.51
OR	12.62	0.67	29.33	Not Applicable			60.83	0.23	66.57	64.44	0.20	69.32
TX	12.72	0.73	30.96	12.74	0.72	30.65	60.21	0.47	71.96	73.25	0.27	80.05
WA	11.13	0.72	29.12	9.94	0.73	28.05	57.89	0.34	66.31	67.64	0.24	73.66

Student Group: non-FRL												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**15.10**	**0.77**	**34.36**	**17.09**	**0.75**	**35.92**	**78.67**	**0.21**	**83.98**	**83.85**	**0.15**	**87.67**
GA	16.01	0.78	35.56	17.90	0.76	37.03	68.63	0.33	76.95	76.45	0.24	82.43
MA	14.19	0.77	33.43	15.59	0.74	34.19	93.2	0.07	95.01	93.6	0.07	95.29
MI	14.60	0.76	33.66	14.86	0.78	34.38	79.84	0.20	84.76	84.20	0.16	88.14
MO	11.84	0.81	32.17	16.89	0.74	35.40	87.20	0.14	90.68	92.17	0.08	94.18
OR	15.69	0.75	34.32	Not Applicable			74.39	0.23	80.02	77.34	0.19	82.19
TX	17.64	0.78	37.22	20.48	0.76	39.41	74.25	0.27	80.89	82.47	0.15	86.34
WA	15.76	0.74	34.17	16.82	0.73	35.09	73.17	0.25	79.54	80.75	0.17	85.12

Notes: The notes to Table 6 apply, except the measurement error correction is different.

Appendix Table B5a. Statewide academic mobility estimates by urbanicity status corresponding to Table 7. We account for measurement error in the 3rd-grade tests by using the 3rd-grade math rank as the initial rank variable and instrumenting for it with the 3rd-grade ELA rank.

	Grade-8 Test			HS Test			Grad			Grad +1		
Student Group: Urban												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**4.48**	**0.91**	**27.22**	**2.60**	**0.94**	**26.15**	**60.37**	**0.44**	**71.42**	**68.48**	**0.34**	**76.94**
GA	5.10	0.87	26.94	4.29	0.91	27.08	45.95	0.62	61.40	56.69	0.48	68.61
MA	4.71	0.89	26.98	5.04	0.86	26.56	76.19	0.28	83.23	77.34	0.27	84.01
MI	5.78	0.89	27.90	0.09	0.98	24.62	61.32	0.43	72.08	67.31	0.37	76.61
MO	2.23	0.91	24.93	2.97	0.90	25.46	67.73	0.35	76.58	76.74	0.24	82.82
OR	4.18	0.94	27.67	Not Applicable			58.9	0.39	68.7	63.41	0.34	71.93
TX	2.93	0.98	27.56	0.29	1.05	26.55	56.46	0.55	70.3	71.61	0.32	79.63
WA	6.41	0.89	28.54	2.9	0.95	26.61	56.04	0.47	67.67	66.24	0.35	74.99
Student Group: Suburban												
	α	*β*	*O25*	*α*	*β*	*O25*	*A*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**4.03**	**0.94**	**27.56**	**2.97**	**0.97**	**27.32**	**66.61**	**0.37**	**75.95**	**73.94**	**0.28**	**80.98**
GA	5.95	0.93	29.20	4.56	0.97	28.68	49.85	0.58	64.38	60.49	0.45	71.64
MA	3.34	0.96	27.31	4.52	0.93	27.8	85.78	0.18	90.24	86.37	0.17	90.64
MI	4.45	0.90	27.05	0.82	0.99	25.59	68.37	0.35	77.12	74.16	0.29	81.44
MO	1.70	0.97	25.89	4.40	0.94	28.00	77.33	0.26	83.78	84.62	0.17	88.84
OR	4.6	0.92	27.65	Not Applicable			61.72	0.38	71.21	65.41	0.34	73.84
TX	2.4	1.01	27.67	-0.28	1.07	26.51	62.54	0.45	73.89	76.05	0.25	82.4
WA	5.79	0.89	28.16	3.77	0.94	27.32	60.67	0.42	71.06	70.45	0.31	78.08
Student Group: Rural												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**3.70**	**0.94**	**27.16**	**1.78**	**0.98**	**26.30**	**67.34**	**0.36**	**76.33**	**74.32**	**0.27**	**81.08**
GA	4.79	0.93	27.96	2.28	0.97	26.51	52.81	0.54	66.38	63.87	0.40	73.77
MA	2.48	0.96	26.59	3.61	0.94	27.14	87.63	0.15	91.25	88.17	0.14	91.62
MI	4.24	0.91	26.96	-0.49	1.02	24.94	67.20	0.36	76.19	72.64	0.31	80.37
MO	2.50	0.96	26.47	2.15	0.96	26.08	80.09	0.23	85.85	86.96	0.14	90.55
OR	3.64	0.91	26.47	Not Applicable			61.28	0.36	70.32	64.69	0.32	72.75
TX	2.76	1.01	27.94	0.35	1.06	26.73	60.97	0.47	72.77	72.86	0.29	80.09
WA	5.5	0.89	27.73	2.76	0.95	26.4	61.43	0.41	71.57	71.08	0.29	78.39

Notes: The notes to Table 7 apply, except the measurement error correction is different. There are several instances of estimates of β at or above 1.0 in MI and TX.

Appendix Table B5b. Statewide academic mobility estimates by urbanicity status corresponding to Table 7. We account for measurement error in the 3rd-grade tests by averaging the ELA and math tests but make no further corrections.

	Grade-8 Test			HS Test			Grad			Grad +1		
Student Group: Urban												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**10.40**	**0.78**	**29.86**	**9.64**	**0.78**	**29.22**	**63.21**	**0.38**	**72.67**	**70.60**	**0.29**	**77.87**
GA	7.85	0.77	27.22	8.32	0.78	27.86	48.39	0.54	61.81	58.66	0.41	68.96
MA	10.31	0.76	29.22	10.22	0.74	28.64	77.69	0.25	83.84	78.78	0.23	84.59
MI	10.76	0.77	30.01	6.78	0.83	27.42	63.72	0.38	73.09	69.36	0.32	77.48
MO	6.71	0.81	26.97	9.25	0.76	28.36	69.58	0.31	77.42	77.99	0.22	83.38
OR	12.87	0.77	32.06	Not Applicable			62.62	0.32	70.58	66.69	0.28	73.59
TX	12.28	0.8	32.19	12.66	0.8	32.62	61.51	0.45	72.82	74.26	0.27	80.97
WA	12.01	0.77	31.35	10.59	0.79	30.41	58.99	0.41	69.15	68.44	0.31	76.1

Student Group: Suburban												
	α	*β*	*O25*	*α*	*β*	*O25*	*A*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**11.95**	**0.79**	**31.83**	**12.84**	**0.79**	**32.63**	**69.72**	**0.32**	**77.63**	**76.22**	**0.24**	**82.22**
GA	11.52	0.82	32.13	11.82	0.83	32.47	53.84	0.51	66.46	63.51	0.39	73.22
MA	11.52	0.79	31.39	12.3	0.78	31.68	87.3	0.15	90.99	87.83	0.14	91.37
MI	12.55	0.76	31.65	11.74	0.80	31.69	71.42	0.30	78.86	76.73	0.25	82.91
MO	8.73	0.85	29.87	13.92	0.78	33.42	79.09	0.23	84.78	85.78	0.15	89.50
OR	13.02	0.77	32.17	Not Applicable			65	0.32	72.98	68.34	0.28	75.42
TX	13.70	0.80	33.83	14.47	0.80	34.48	67.27	0.37	76.48	78.38	0.21	83.69
WA	12.61	0.77	31.75	12.8	0.77	32.03	64.11	0.35	72.87	73	0.26	79.42

Student Group: Rural												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**11.29**	**0.78**	**30.78**	**11.16**	**0.79**	**30.79**	**70.31**	**0.30**	**77.81**	**76.53**	**0.23**	**82.18**
GA	11.34	0.80	31.36	10.61	0.81	30.78	56.90	0.46	68.49	66.94	0.34	75.36
MA	11.74	0.77	31.11	12.45	0.76	31.45	88.78	0.12	91.81	89.27	0.12	92.15
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	9.89	0.81	30.12	12.14	0.76	31.10	81.95	0.19	86.77	88.15	0.12	91.14
OR	12.29	0.72	30.32	Not Applicable			64.49	0.29	71.75	67.47	0.26	73.98
TX	12.42	0.80	32.53	13.15	0.79	32.88	65.30	0.38	74.83	75.24	0.24	81.21
WA	11.81	0.75	30.59	11.18	0.76	30.24	64.46	0.34	72.94	73.25	0.24	79.37

Notes: The notes to Table 7 apply, except the measurement error correction is different.

Appendix Table B6a. Estimates of the of the within-state, cross-district standard deviations of $\bar{O}_{25d}$, α_d, and β_d, corresponding to Table 8. We account for measurement error in the 3rd-grade tests by using the 3rd-grade math rank as the initial rank variable and instrumenting for it with the 3rd-grade ELA rank.

	Grade-8 Test			HS Test			Grad			Grad +1		
Standard Deviations												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**6.26**	**0.09**	**5.64**	**6.11**	**0.09**	**5.44**	**7.51**	**0.11**	**5.37**	**6.53**	**0.09**	**4.77**
GA	4.02	0.06	3.77	4.24	0.06	3.99	7.30	0.10	5.34	6.70	0.08	4.97
MA	6.39	0.07	5.85	6.52	0.07	6.02	6.36	0.08	4.43	6.04	0.08	4.24
MI	5.40	0.09	5.03	7.16	0.11	6.28	8.31	0.12	6.22	7.80	0.10	5.73
MO	4.41	0.09	4.60	4.71	0.10	4.51	5.58	0.08	4.00	3.78	0.05	2.83
OR	11.20	0.13	9.55	Not Applicable			8.51	0.10	6.55	8.35	0.09	6.39
TX	8.13	0.10	6.66	9.23	0.12	7.44	8.96	0.17	5.48	6.50	0.11	4.31
WA	4.26	0.06	4.05	4.80	0.06	4.39	7.52	0.10	5.59	6.56	0.09	4.94

Notes: The notes to Table 8 apply, except the measurement error correction is different. The standard deviations in Oregon are exacerbated by a number of outlying districts (our error variance correction does not fully offset their effects on the standard deviations).

Appendix Table B6b. Estimates of the of the within-state, cross-district standard deviations of $\bar{O}_{25d}$, α_d, and β_d, corresponding to Table 8. We account for measurement error in the 3rd-grade tests by averaging the ELA and math tests but make no further corrections.

	Grade-8 Test			HS Test			Grad			Grad +1		
Standard Deviations												
	α	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*	*α*	*β*	*O25*
All (Avg)	**4.82**	**0.05**	**4.74**	**5.02**	**0.06**	**4.87**	**6.78**	**0.09**	**5.19**	**5.73**	**0.07**	**4.50**
GA	3.35	0.03	3.40	3.79	0.04	3.79	6.51	0.09	4.90	5.75	0.07	4.48
MA	6.54	0.05	6.07	6.50	0.05	6.04	4.88	0.06	3.61	4.64	0.05	3.44
MI	4.13	0.06	4.16	4.95	0.07	5.01	7.51	0.09	6.24	6.85	0.08	5.54
MO	4.14	0.06	4.28	4.37	0.06	4.24	4.87	0.06	3.68	3.04	0.03	2.50
OR	6.78	0.06	6.22	Not Applicable			9.00	0.10	6.96	8.67	0.10	6.72
TX	5.51	0.06	5.11	5.80	0.06	5.46	7.35	0.12	5.02	5.62	0.08	4.22
WA	3.31	0.06	3.96	4.72	0.06	4.65	7.31	0.09	5.93	5.56	0.08	4.62

Notes: The notes to Table 8 apply, except the measurement error correction is different. The standard deviations in Oregon are exacerbated by a number of outlying districts (our error variance correction does not fully offset their effects on the standard deviations).

Appendix Table B7a. Adjusted and unadjusted correlations of $\bar{O}_{25d}$ across outcomes, on average across states, corresponding to Table 9. We account for measurement error in the 3rd-grade tests by using the 3rd-grade math rank as the initial rank variable and instrumenting for it with the 3rd-grade ELA rank.

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.93	1.00		
Grad	0.34	0.33	1.00	
Grad +1	0.32	0.32	1.00	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.82	1.00		
Grad	0.29	0.29	1.00	
Grad +1	0.27	0.27	0.92	1.00

Notes: The notes to Table 9 apply, except the measurement error correction is different. We do not show state-by-state correlations for brevity, but they are available upon request (also, for our primary EIV method, state-by-state correlations are reported in Appendix A).

Appendix Table B7b. Adjusted and unadjusted correlations of $\bar{O}_{25d}$ across outcomes, on average across states, corresponding to Table 9. We account for measurement error in the 3rd-grade tests by averaging the ELA and math tests but make no further corrections.

Adjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.87	1.00		
Grad	0.38	0.40	1.00	
Grad +1	0.36	0.36	1.00	1.00

Unadjusted Correlations				
	Grade-8 test	HS test	Grad	Grad +1
Grade-8 test	1.00			
HS test	0.77	1.00		
Grad	0.33	0.33	1.00	
Grad +1	0.32	0.30	0.92	1.00

Notes: The notes to Table 9 apply, except the measurement error correction is different. We do not show state-by-state correlations for brevity, but they are available upon request (also, for our primary EIV method, state-by-state correlations are reported in Appendix A).

Appendix Table B8a. Adjusted and unadjusted correlations of $\bar{O}_{25d}$ across cohorts for each outcome, on average across states, corresponding to Table 10. We account for measurement error in the 3rd-grade tests by using the 3rd-grade math rank as the initial rank variable and instrumenting for it with the 3rd-grade ELA rank.

Grade-8 Test

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.71	1.00		
2008	0.36	0.63	1.00	
2009	0.40	0.57	0.69	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.34	1.00		
2008	0.29	0.35	1.00	
2009	0.32	0.47	0.55	1.00

HS Test

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.57	1.00		
2008	0.53	0.63	1.00	
2009	0.42	0.58	0.69	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.33	1.00		
2008	0.38	0.44	1.00	
2009	0.32	0.44	0.51	1.00

Graduation

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.68	1.00		
2008	0.23	0.44	1.00	
2009	0.56	0.54	0.64	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.43	1.00		
2008	0.32	0.33	1.00	
2009	0.44	0.43	0.50	1.00

Graduation +1

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.67	1.00		
2008	0.18	0.40	1.00	
2009	0.50	0.51	0.60	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.42	1.00		
2008	0.25	0.32	1.00	
2009	0.39	0.40	0.47	1.00

Notes: The notes to Table 10 apply, except the measurement error correction is different. We do not show state-by-state correlations for brevity, but they are available upon request (also, for our primary EIV method, state-by-state correlations are reported in Appendix A). Note there are two instances where the "lower bound" correlation exceeds the "upper bound" correlation in the grad and grad +1 models (between 2006 and 2008). This because in one state (OR), the unadjusted correlation is negative and the error-adjustment is large, resulting in a very large and negative adjusted correlation that overwhelms the other, positive error-variance adjustments.

Appendix Table B8b. Adjusted and unadjusted correlations of $\bar{O}_{25d}$ across cohorts for each outcome, on average across states, corresponding to Table 10. We account for measurement error in the 3rd-grade tests by averaging the ELA and math tests but make no further corrections.

Grade-8 Test

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.70	1.00		
2008	0.61	0.72	1.00	
2009	0.53	0.68	0.77	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.59	1.00		
2008	0.51	0.62	1.00	
2009	0.43	0.56	0.63	1.00

HS Test

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.74	1.00		
2008	0.71	0.75	1.00	
2009	0.67	0.69	0.77	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.55	1.00		
2008	0.53	0.62	1.00	
2009	0.48	0.54	0.61	1.00

Graduation

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.54	1.00		
2008	0.46	0.54	1.00	
2009	0.52	0.59	0.64	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.45	1.00		
2008	0.38	0.45	1.00	
2009	0.44	0.49	0.51	1.00

Graduation +1

Adjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.53	1.00		
2008	0.40	0.55	1.00	
2009	0.52	0.60	0.63	1.00

Unadjusted Correlations				
	2006	2007	2008	2009
2006	1.00			
2007	0.44	1.00		
2008	0.33	0.44	1.00	
2009	0.43	0.48	0.49	1.00

Notes: The notes to Table 10 apply, except the measurement error correction is different. We do not show state-by-state correlations for brevity, but they are available upon request (also, for our primary EIV method, state-by-state correlations are reported in Appendix A).

Appendix Table B9a. Regression output for the correlates of academic mobility corresponding to Figure 7 (and Appendix Table A4). All coefficients are from univariate regressions where the independent variable is standardized to have a mean of zero and variance of one within states. The dependent variables in these regressions are estimates of O25 taken from models in which we account for measurement error in the 3rd-grade tests by using the 3rd-grade math rank as the initial rank variable and instrumenting for it with the 3rd-grade ELA rank.

	Grade-8 Test							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	1.56*	2.07*	3.62*	2.23*	3.39*	3.35*	1.35*	2.51
High-Achieving VA	1.00*	1.25*	2.16*	1.96*	3.77*	2.33*	1.31*	1.97
Low-Achieving VA	1.35*	2.49*	1.54*	0.50	3.30*	3.67*	0.78*	1.95
Pct Asian	0.48*	2.19*	0.43	0.68*	-0.35	1.42*	0.79*	0.81
Pct Black	-0.78*	-0.08	-0.01	-0.24	0.18	0.06*	0.06	-0.12
Pct Hispanic	0.75*	-1.19*	-0.06	-0.02	2.38*	-1.34*	0.32	0.12
Pct FRL	-1.11*	-1.63*	-0.56	-0.65*	1.86*	-1.62*	-0.36*	-0.58
Pct IEP	0.23*	0.11	-2.15*	0.18	-1.04	1.31*	-0.13	-0.21
Pct Mobile	-0.40*	-2.14*	0.08	-1.05*	2.38*	-2.00*	-0.69*	-0.55
School Seg. Index	-0.05*	0.21	0.07	-0.20*	1.44*	-0.31*	0.73*	0.27
Econ Seg. Index	0.02*	-0.88*	0.84*	-0.23*	1.73*	0.59*	-0.44*	0.23
NCES Data Elements								
Per Pupil Spending	-0.72*	0.66*	0.29	0.13	-0.75	-2.61*	-0.34	-0.48
Pct Asian - Neighborhood	0.56*	2.20*	0.30	0.67*	-0.68	0.92*	0.90*	0.70
Pct Black - Neighborhood	-0.55*	-0.26	-0.06	-0.25*	0.35	0.16*	0.08	-0.08
Pct Hispanic - Neighborhood	0.76*	-1.30*	-0.07	-0.07	2.27*	-1.35*	0.20	0.06
Median Income	0.88*	2.53*	0.69*	0.53*	-1.82*	1.54*	0.72*	0.72
Pct Poverty	-1.16*	-1.43*	-0.35	-0.69*	1.66*	-1.49*	-0.41*	-0.55
Pct Other Language at Home	0.91*	-0.200	0.50*	0.46	1.80*	-0.84*	0.75*	0.48
Pct Residence Stability	0.58*	0.66*	0.67*	0.38	0.12	0.24*	0.40*	0.44
Pct HS Grad	0.29*	1.66*	0.11	0.77*	-2.75*	1.42*	-0.15	0.19
Pct Bachelors	0.57*	2.86*	0.75*	0.57*	-1.10	1.63*	0.82*	0.87
Pct Never Married	-1.05*	-1.88*	-0.22	-0.47*	1.64*	-1.48*	-0.51*	-0.57

	HS Test							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	0.84*	1.72*	3.16*	1.58*	Not Applicable	2.74*	0.98*	1.84
High-Achieving VA	0.31*	0.93*	2.17	1.37*		1.72*	0.97*	1.25
Low-Achieving VA	0.64*	2.09*	1.21*	-0.38		3.20*	0.26	1.17
Pct Asian	0.44*	2.20*	0.44	1.29*		1.97*	1.18*	1.25
Pct Black	-0.61*	-0.46	-0.21	-0.02		-0.22*	0.13	-0.23
Pct Hispanic	0.75*	-1.88*	-0.12	-0.27		-1.63*	-0.22	-0.56
Pct FRL	-1.46*	-2.54*	-0.84*	-1.54*		-2.19*	-1.26*	-1.64
Pct IEP	-0.18*	-0.34	-0.85	1.43*		2.04*	-0.16	0.32
Pct Mobile	0.09*	-2.57*	-0.09	-1.84*		-3.03*	-0.77*	-1.37
School Seg. Index	0.27*	-0.35	0.09*	-0.14		-0.47*	0.44*	-0.03
Econ Seg Index	0.27*	-1.22*	1.01*	-0.10		0.93*	-0.78*	0.02
NCES Data Elements								
Per Pupil Spending	-0.77*	0.51	0.80	1.05*		-3.54*	-0.58*	-0.42
Pct Asian - Neighborhood	0.56*	2.47*	0.46	1.38*		1.27*	1.42*	1.26
Pct Black - Neighborhood	-0.35*	-0.66*	-0.03	-0.09		-0.01*	0.21	-0.16
Pct Hispanic - Neighborhood	0.82*	-1.97*	-0.12	-0.25		-1.67*	-0.32	-0.59
Median Income	1.20*	3.56*	0.90*	1.52*		2.04*	1.58*	1.80
Pct Poverty	-1.56*	-2.28*	-0.42	-1.52*		-2.06*	-1.07*	-1.49
Pct Other Language at Home	1.00*	-0.79*	0.60*	0.88*		-1.01*	0.51*	0.20
Pct Residence Stability	-0.04*	1.29*	0.56	1.11*		0.41*	0.24	0.60
Pct HS Grad	0.69*	2.56*	0.28	1.81*		2.03*	0.43*	1.30
Pct Bachelors	0.90*	3.82*	0.96*	1.51*		2.27*	1.45*	1.82
Pct Never Married	-1.03*	-2.84*	-0.23	-0.55*		-2.07*	-1.07*	-1.30

	Grad							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	0.32*	1.00*	4.54*	1.06*	1.12*	1.81*	1.26*	1.59
High-Achieving VA	0.41*	0.69*	1.58	0.81*	0.66	0.72*	0.47	0.76
Low-Achieving VA	-0.09*	0.65*	-1.22	-1.76*	1.38*	2.72*	0.80*	0.35
Pct Asian	-0.23*	0.57	0.87	0.74*	-0.11	1.59*	0.04	0.50
Pct Black	-1.39*	-3.22*	-1.53	-2.48*	-0.70	-1.71*	-1.76*	-1.83
Pct Hispanic	-0.12*	-4.46*	-1.76*	-1.90*	1.74*	-1.47*	-0.51	-1.21
Pct FRL	-1.34*	-5.31*	-3.25*	-3.24*	-1.08*	-2.68*	-2.46*	-2.77
Pct IEP	1.16*	-2.20*	-6.17*	2.63*	-1.27*	3.87*	-0.80*	-0.40
Pct Mobile	-1.55*	-3.29*	-4.20*	-5.00*	-1.19*	-6.47*	0.11	-3.08
School Seg. Index	-0.66*	-3.51*	-0.11	-1.59*	-0.17	-1.66*	-0.83*	-1.22
Econ Seg Index	-0.75*	-1.09*	1.44*	-0.75*	-0.79	-0.37*	-1.19*	-0.50
NCES Data Elements								
Per Pupil Spending	-1.96*	-2.27*	-0.76	-2.43*	-0.28	-3.44*	-1.21*	-1.76
Pct Asian - Neighborhood	-0.20*	1.06*	0.66	0.49*	0.07	1.00*	0.92*	0.57
Pct Black - Neighborhood	-1.24*	-3.39*	-0.61	-1.66*	-0.60	-0.78*	-1.60*	-1.41
Pct Hispanic - Neighborhood	0	-4.23*	-1.64*	-1.75*	1.43*	-1.40*	-0.56	-1.16
Median Income	0.95*	4.54*	2.72*	1.85*	1.09*	2.42*	2.63*	2.31
Pct Poverty	-1.09*	-5.15*	-1.93*	-3.15*	-1.09*	-2.38*	-1.95*	-2.39
Pct Other Language at Home	-0.23*	-3.63*	0.93*	-1.60*	1.32*	-1.00*	-0.11	-0.62
Pct Residence Stability	0.87*	3.37*	3.18*	2.60*	1.97*	1.71*	1.43*	2.16
Pct HS Grad	0.02*	4.99*	1.53*	3.21*	-0.98*	2.26*	0.77*	1.69
Pct Bachelors	-0.02*	4.62*	2.40*	1.24*	0.50	2.21*	2.12*	1.87
Pct Never Married	-1.53*	-5.36*	-1.64	-2.51*	-1.94*	-2.83*	-2.20*	-2.57

	Grad+1							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	0.22*	0.94*	3.77*	1.16*	0.85*	1.67*	0.86*	1.35
High-Achieving VA	0.21*	0.62*	1.25	0.80*	0.42	0.90*	0.16	0.62
Low-Achieving VA	-0.04*	0.62*	-1.25	-1.05*	0.85*	2.19*	0.51*	0.26
Pct Asian	-0.24*	0.61*	0.83*	0.56*	0.33	1.32*	0.11	0.50
Pct Black	-1.05*	-3.04*	-1.39	-1.91*	-0.41	-0.58*	-1.55*	-1.42
Pct Hispanic	-0.13*	-4.28*	-1.61*	-1.54*	1.52*	-0.40*	-0.36	-0.97
Pct FRL	-1.06*	-5.08*	-2.89*	-2.48*	-1.34*	-1.33*	-2.16*	-2.33
Pct IEP	1.00*	-2.07*	-5.61*	2.10*	-1.11*	0.60*	-0.57*	-0.81
Pct Mobile	-1.53*	-3.20*	-3.69*	-4.04*	-1.24*	-4.54*	0.15	-2.58
School Seg. Index	-0.64*	-3.26*	-0.10	-1.28*	0.01	-0.69*	-0.56*	-0.93
Econ Seg index	-0.57*	-1.04*	1.23*	-0.65	-0.56	-0.23*	-0.92*	-0.39
NCES Data Elements								
Per Pupil Spending	-1.86*	-2.11*	-0.73	-1.86*	-0.22	-2.13*	-1.08*	-1.43
Pct Asian - Neighborhood	-0.22*	1.07*	0.71*	0.36	0.49	0.87*	0.81*	0.58
Pct Black - Neighborhood	-0.98*	-3.21*	-0.55	-1.27*	-0.32	-0.13*	-1.41*	-1.12
Pct Hispanic - Neighborhood	-0.08*	-4.08*	-1.52*	-1.43*	1.19*	-0.34*	-0.42	-0.95
Median Income	0.85*	4.39*	2.42*	1.43*	1.52*	1.34*	2.29*	2.03
Pct Poverty	-0.86*	-4.95*	-1.75*	-2.43*	-1.29*	-1.08*	-1.77*	-2.02
Pct Other Language at Home	-0.28*	-3.46*	0.95*	-1.35*	1.34*	0.16*	-0.04	-0.38
Pct Residence Stability	0.61*	3.25*	2.85*	2.00*	2.05*	1.41*	1.08*	1.89
Pct HS Grad	0.10*	4.81*	1.37*	2.46	-0.70	0.93*	0.66*	1.38
Pct Bachelors	-0.01*	4.48*	2.14*	0.92	0.98*	1.34*	1.81*	1.67
Pct Never Married	-1.17*	-5.13*	-1.47	-2.00	-2.17*	-1.42*	-1.93*	-2.18

Notes: The notes to Appendix Table A4 apply. Statistical significance at the 5 percent level for each coefficient in the state-by-state results is denoted by *. Standard errors are suppressed for brevity. The symbol † is to denote that statistical significance is not reported for the "All (avg)" values because they are not directly generated from a regression (they are average values of the state-by-state regression coefficients). These models are estimated on the subsample of districts that excludes the smallest districts in each state as described in the introductory text to this appendix.

Appendix Table B9b. Regression output for the correlates of academic mobility corresponding to Figure 7 (and Appendix Table A4). All coefficients are from univariate regressions where the independent variable is standardized to have a mean of zero and variance of one within states. The dependent variables in these regressions are estimates of O25 taken from models in which we account for measurement error in the 3rd-grade tests by averaging the ELA and math tests but make no further corrections.

	Grade-8 Test							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	1.40*	1.84*	3.98*	2.25*	2.20*	2.97*	1.14*	2.25
High-Achieving VA	0.55*	0.95*	3.49*	1.81*	2.09*	1.62*	0.77*	1.61
Low-Achieving VA	1.10*	2.06*	-1.61*	-0.06	1.97*	3.80*	0.46*	1.10
Pct Asian	0.70*	1.96*	1.19*	0.86*	0.39	1.92*	1.06*	1.15
Pct Black	-2.17*	-1.01*	-2.31*	-1.25*	0.20	-0.60*	-0.24	-1.05
Pct Hispanic	1.08*	-2.27*	-0.96*	-0.42*	1.35*	-1.80*	-0.55*	-0.51
Pct FRL	-2.43*	-2.99*	-2.73*	-1.86*	0.27	-2.45*	-1.42*	-1.94
Pct IEP	0.97*	-0.75*	-2.06*	0.89*	-0.74*	2.60*	-0.13	0.11
Pct Mobile	-1.45*	-2.84*	-2.72*	-2.20*	1.01*	-4.40*	-0.85*	-1.92
School Seg. Index	-0.60*	-0.72*	-0.42*	-0.66*	1.05*	-0.67*	0.47*	-0.22
Econ Seg Index	-0.35*	-1.11*	0.98*	-0.21*	1.40*	0.65*	-0.43*	0.13
NCES Data Elements								
Per Pupil Spending	-1.54*	0.04	-2.20*	-0.64*	-0.45	-3.27*	-0.61*	-1.24
Pct Asian - Neighborhood	0.73*	2.28*	0.94*	0.81*	0.14	1.26*	1.36*	1.07
Pct Black - Neighborhood	-1.71*	-1.27*	-1.18*	-0.95*	0.33	-0.13*	-0.28	-0.74
Pct Hispanic - Neighborhood	0.99*	-2.23*	-1.33*	-0.49*	1.25*	-1.83*	-0.65*	-0.61
Median Income	1.59*	3.74*	2.25*	1.24*	-0.24	2.49*	1.74*	1.83
Pct Poverty	-2.36*	-2.78*	-2.03*	-1.92*	0.37	-2.10*	-1.30*	-1.73
Pct Other Language at Home	1.15*	-1.09*	0.55*	-0.06	1.24*	-1.14*	0.16	0.12
Pct Residence Stability	1.54*	1.58*	2.32*	1.30*	0.41	0.69*	0.48*	1.19
Pct HS Grad	0.64*	2.78*	1.87*	1.90*	-1.25*	1.92*	0.79*	1.24
Pct Bachelors	0.75*	4.02*	2.31*	1.05*	0.22	2.46*	1.85*	1.81
Pct Never Married	-2.54*	-3.16*	-1.64*	-1.42*	0.40	-2.32*	-1.34*	-1.72

	HS Test							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	0.63*	1.53*	4.22*	1.62*		2.30*	0.74*	1.84
High-Achieving VA	-0.22*	0.71*	4.12*	1.20*		0.79*	0.32	1.15
Low-Achieving VA	0.32*	1.72*	-3.18*	-1.06*		3.44*	-0.11	0.19
Pct Asian	0.68*	1.98*	1.20*	1.51*		2.59*	1.55*	1.59
Pct Black	-2.19*	-1.30*	-3.06*	-1.25*		-1.05*	-0.21	-1.51
Pct Hispanic	1.10*	-2.86*	-1.27*	-0.80*		-2.35*	-1.33*	-1.25
Pct FRL	-2.96*	-3.77*	-3.29*	-2.99*		-3.34*	-2.61*	-3.16
Pct IEP	0.64*	-1.16*	-1.12	2.30*		3.95*	-0.22	0.73
Pct Mobile	-1.10*	-3.19*	-3.50*	-3.15*		-5.93*	-0.97*	-2.97
School Seg. Index	-0.36*	-1.19*	-0.48*	-0.71*		-1.00*	0.18	-0.59
Econ Seg Index	-0.16*	-1.45*	1.11*	-0.10		1.00*	-0.76*	-0.06
NCES Data Elements					Not Applicable			
Per Pupil Spending	-1.68*	-0.10	-2.34*	0.03		-4.41*	-0.90*	-1.57
Pct Asian - Neighborhood	0.74*	2.52*	1.09*	1.53*		1.69*	2.01*	1.60
Pct Black - Neighborhood	-1.68*	-1.56*	-1.39*	-0.94*		-0.37*	-0.21	-1.03
Pct Hispanic - Neighborhood	1.08*	-2.81*	-1.57*	-0.82*		-2.43*	-1.41*	-1.33
Median Income	2.00*	4.63*	2.62*	2.36*		3.30*	2.89*	2.97
Pct Poverty	-2.92*	-3.50*	-2.34*	-2.98*		-2.96*	-2.20*	-2.82
Pct Other Language at Home	1.25*	-1.61*	0.55*	0.16		-1.57*	-0.25	-0.25
Pct Residence Stability	1.03*	2.12*	2.56*	2.17*		0.89*	0.35	1.52
Pct HS Grad	1.08*	3.58*	2.31*	3.16*		2.80*	1.63*	2.43
Pct Bachelors	1.11*	4.85*	2.69*	2.07*		3.33*	2.76*	2.80
Pct Never Married	-2.72*	-3.99*	-1.89*	-1.71*		-3.17*	-2.14*	-2.60

	Grad							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	0.22*	0.92*	4.28*	1.15*	0.70	1.50*	1.11*	1.41
High-Achieving VA	0.11*	0.64*	1.83	0.82*	0.01	0.54*	0.22	0.60
Low-Achieving VA	-0.21*	0.56*	-2.05	-1.32*	0.87*	2.36*	0.65*	0.12
Pct Asian	-0.07*	0.45	0.92*	0.59*	0.24	1.52*	0.11	0.54
Pct Black	-2.12*	-3.21*	-1.99*	-2.68*	-0.64	-1.74*	-1.81*	-2.03
Pct Hispanic	0.10*	-4.39*	-1.94*	-1.91*	1.21*	-1.12*	-0.83*	-1.27
Pct FRL	-2.06*	-5.21*	-3.54*	-3.26*	-1.71*	-2.42*	-2.73*	-2.99
Pct IEP	1.51*	-2.26*	-5.88*	2.31*	-1.06*	3.26*	-0.71*	-0.40
Pct Mobile	-2.09*	-3.12*	-4.69*	-4.67*	-1.72*	-6.31*	-0.04	-3.23
School Seg. Index	-0.96*	-3.54*	-0.22	-1.65*	-0.29	-1.55*	-0.91*	-1.30
Econ Seg Index	-0.93*	-1.07*	1.32*	-0.74*	-0.83*	-0.39*	-1.19*	-0.55
NCES Data Elements								
Per Pupil Spending	-2.41*	-2.10*	-1.46	-2.45*	-0.30	-2.62*	-1.29*	-1.80
Pct Asian - Neighborhood	-0.06*	0.94*	0.71	0.35	0.41	0.98*	0.99*	0.62
Pct Black - Neighborhood	-1.83*	-3.37*	-0.84	-1.80*	-0.56	-0.77*	-1.65*	-1.55
Pct Hispanic - Neighborhood	0.16*	-4.14*	-1.90*	-1.80*	0.90*	-1.08*	-0.88*	-1.25
Median Income	1.34*	4.40*	2.85*	1.77*	1.64*	2.28*	2.85*	2.45
Pct Poverty	-1.79*	-5.06*	-2.19*	-3.20*	-1.56*	-2.03*	-2.20*	-2.58
Pct Other Language at Home	-0.05*	-3.56*	0.80	-1.74*	1.01*	-0.64*	-0.38	-0.65
Pct Residence Stability	1.43*	3.31*	3.39*	2.68*	1.89*	1.80*	1.39*	2.27
Pct HS Grad	0.23*	4.86*	1.88*	3.10*	-0.26	1.90*	1.10*	1.83
Pct Bachelors	0.09*	4.47*	2.54*	1.13*	1.09*	2.08*	2.35*	1.96
Pct Never Married	-2.34*	-5.23*	-1.88*	-2.64*	-2.25*	-2.59*	-2.36*	-2.76

	Grad+1							
	GA	MA	MI	MO	OR	TX	WA	All (Avg)†
Administrative Data Elements								
Value Added	0.14*	0.87*	3.55*	1.20*	0.50	1.40*	0.77*	1.20
High-Achieving VA	-0.03*	0.59*	1.44	0.79*	-0.13	0.71*	0.05	0.49
Low-Achieving VA	-0.12*	0.53	-1.83	-0.72*	0.47	1.90*	0.42	0.09
Pct Asian	-0.11*	0.49	0.85*	0.44*	0.61	1.25*	0.19	0.53
Pct Black	-1.62*	-3.03*	-1.73	-2.06*	-0.33	-0.69*	-1.54*	-1.57
Pct Hispanic	0.02*	-4.21*	-1.73*	-1.54*	1.05*	-0.22*	-0.59*	-1.03
Pct FRL	-1.61*	-4.99*	-3.07*	-2.46*	-1.83*	-1.21*	-2.33*	-2.50
Pct IEP	1.29*	-2.14*	-5.33*	1.84*	-0.90*	0.45*	-0.50*	-0.76
Pct Mobile	-1.94*	-3.04*	-4.06*	-3.74*	-1.67*	-4.47*	0.04	-2.70
School Seg. Index	-0.87*	-3.29*	-0.18	-1.32*	-0.09	-0.66*	-0.59*	-1.00
Econ Seg Index	-0.73*	-1.02*	1.11*	-0.65*	-0.58	-0.24*	-0.94*	-0.44
NCES Data Elements								
Per Pupil Spending	-2.17*	-1.95*	-1.26	-1.85*	-0.20	-1.48*	-1.12*	-1.43
Pct Asian - Neighborhood	-0.11*	0.95*	0.73*	0.25	0.74	0.83*	0.88*	0.61
Pct Black - Neighborhood	-1.45*	-3.20*	-0.72	-1.37*	-0.25	-0.19*	-1.41*	-1.23
Pct Hispanic - Neighborhood	0.03*	-3.98*	-1.69*	-1.46*	0.72	-0.17*	-0.65*	-1.03
Median Income	1.13*	4.25*	2.48*	1.36*	1.93*	1.28*	2.44*	2.12
Pct Poverty	-1.40*	-4.87*	-1.93*	-2.44*	-1.65*	-0.91*	-1.94*	-2.16
Pct Other Language at Home	-0.15*	-3.40*	0.84*	-1.44*	1.05*	0.33*	-0.22	-0.43
Pct Residence Stability	1.09*	3.19*	2.98*	2.04*	1.95*	1.46*	1.05*	1.97
Pct HS Grad	0.25*	4.68*	1.62*	2.34*	-0.08	0.76*	0.89*	1.49
Pct Bachelors	0.06*	4.33*	2.21*	0.82*	1.45*	1.28*	1.98*	1.73
Pct Never Married	-1.80*	-5.01*	-1.64	-2.09*	-2.36*	-1.33*	-2.03*	-2.32

Notes: The notes to Appendix Table A4 apply. Statistical significance at the 5 percent level for each coefficient in the state-by-state results is denoted by *. Standard errors are suppressed for brevity. The symbol † is to denote that statistical significance is not reported for the "All (avg)" values because they are not directly generated from a regression (they are average values of the state-by-state regression coefficients).

Table B10a. Correlations of value-added to student achievement with α_d and β_d corresponding to Table 11. We account for measurement error in the 3rd-grade tests by using the 3rd-grade math rank as the initial rank variable and instrumenting for it with the 3rd-grade ELA rank.

	Grade-8 Test		HS Test		Grad		Grad +1	
	α	β	α	β	α	β	α	β
All (Avg)	**0.32**	**0.06**	**0.19**	**0.05**	**0.13**	**-0.08**	**0.13**	**-0.08**
GA	0.33	-0.12	0.18	-0.17	0.01	0.05	0.00	0.06
MA	0.26	0.23	0.20	0.24	0.08	-0.07	0.08	-0.08
MI	0.33	0.14	0.18	0.16	0.23	-0.22	0.22	-0.22
MO	0.27	0.04	0.17	0.01	0.12	-0.14	0.16	-0.17
OR	0.42	-0.03	Not Applicable		0.12	0.04	0.08	0.08
TX	0.30	-0.04	0.22	-0.09	0.12	-0.07	0.17	-0.13
WA	0.33	0.18	0.20	0.13	0.21	-0.14	0.17	-0.11

Notes: The notes to Table 11 apply, except the measurement error correction is different.

Table B10b. Correlations of value-added to student achievement with α_d and β_d corresponding to Table 11. We account for measurement error in the 3rd-grade tests by averaging the ELA and math tests but make no further corrections.

	Grade-8 Test		HS Test		Grad		Grad +1	
	α	β	α	β	α	β	α	β
All (Avg)	**0.34**	**0.06**	**0.21**	**0.06**	**0.13**	**-0.08**	**0.13**	**-0.08**
GA	0.26	-0.14	0.11	-0.12	0.01	0.05	0.00	0.06
MA	0.27	0.25	0.20	0.25	0.12	-0.04	0.14	-0.05
MI	0.41	0.15	0.35	0.18	0.22	-0.20	0.21	-0.20
MO	0.35	0.05	0.20	0.01	0.14	-0.16	0.18	-0.20
OR	0.45	-0.04	Not Applicable		0.09	0.01	0.06	0.04
TX	0.32	-0.03	0.22	-0.10	0.12	-0.07	0.16	-0.12
WA	0.29	0.20	0.15	0.13	0.20	-0.14	0.17	-0.12

Notes: The notes to Table 11 apply, except the measurement error correction is different.

Appendix C: Imputation Procedure

We retain the full entering third-grade cohorts throughout our analysis by imputing missing later-grade outcomes. The imputation is performed on an outcome-by-outcome basis—e.g., for a student with an eighth-grade test score, but no high school test score and no data on high school graduation, we retain the observed eighth-grade score for use in our analysis and impute the latter three outcomes.

Imputed values for each missing outcome are a function of student demographics in the third grade (race-ethnicity, gender) along with information on FRL status, English as a second language (ESL) status, IEP status, and available test scores from grades 3-7. For example, for a student who exits one of our sample states after the fifth grade, we impute the four focal later-grade outcomes using information from her profile from grades 3-5. We do not use student characteristics or test scores after the seventh grade for imputation for any student in order to enforce consistency of the imputation procedure across all later-grade outcomes, the first of which is recorded in the eighth grade.

We begin the imputation process by using data for students with all observed later-grade outcomes to estimate a series of regressions of the following form:

$$O_{iq} = \beta_0 + \boldsymbol{Y_{iq}\beta_1} + \boldsymbol{X_{1i}\beta_2} + \boldsymbol{X_{2iq}\beta_3} + \varepsilon_{iq} \tag{C1}$$

where O_{iq} is a later-grade outcome for student *i* predicted using student characteristics and test-score records through grade *q* (*q*=3, 4, 5, 6, 7). $\boldsymbol{Y_{iq}}$ is a vector of test data for student *i* in math and ELA (with test scores standardized by subject-grade-year), the length of which depends on *q*, $\boldsymbol{X_{1i}}$ is a vector of racial-ethnic and gender designations based on the third-grade record, and $\boldsymbol{X_{2iq}}$ is a vector containing year-by-year student designations for FRL, ELS, and IEP. We estimate versions of equation (C1) for each later-grade outcome and all five values of q, using the samples of students in each state for whom all later-grade outcomes are observed.

The parameters from equation (C1) can be applied to predict later-grade outcomes for students who are missing these outcomes with q-values ranging from 3-7 (inclusive). These predictions form the basis of our imputation procedure, to which we make two additional adjustments and extend for sensitivity testing.

The first adjustment is that we add an indicator for within-state district mobility to equation (C1), which we expand as follows:

$$O_{iq} = \delta_0 + \boldsymbol{Y_{iq}}\boldsymbol{\delta_1} + \boldsymbol{X_{1i}}\boldsymbol{\delta_2} + \boldsymbol{X_{2iq}}\boldsymbol{\delta_3} + Z_i\delta_4 + \eta_{iq} \tag{C2}$$

Like terms in equation (C2) are defined as in equation (C1). The addition to equation (C2), Z_i, is an indicator variable equal to one if student *i* is observed changing districts within the state at least once prior to the time at which the outcome is assessed, and zero otherwise. Therefore, the coefficient δ_4 captures the additional predictive power of cross-district mobility within a state over the outcome. Using this adjusted equation, we impute later-grade outcomes for students who are missing these outcomes with the following predictions, where q indicates the last grade in which student *i* is observed with a test record in the state data through grade-7:

$$\hat{O}_{iq} = \hat{\delta}_0 + \boldsymbol{Y_{iq}}\widehat{\boldsymbol{\delta}}_{\boldsymbol{1}} + \boldsymbol{X_{1i}}\widehat{\boldsymbol{\delta}}_{\boldsymbol{2}} + \boldsymbol{X_{2iq}}\widehat{\boldsymbol{\delta}}_{\boldsymbol{3}} + Z_i\hat{\delta}_4$$

(C3)

Equation (C3) is the imputation equation used in the primary analysis in the paper. If all students with missing later-grade outcomes were state exiters, these imputed values would be accurate under the assumption that within-state district mobility and cross-state mobility are equally predictive of student outcomes. Treating this as a working assumption is approximately accurate because *most* students with missing outcomes are state exiters (although not all; e.g., in practice, some students miss the tests each year).

The second adjustment is needed because shrinkage is inherent in the predictions in equation (C3). If left unaccounted for, the shrinkage would result in compressed distributions of imputed outcomes relative to the distributions observed in the real data. This is problematic for the test-score outcomes because they are ranked and increasing the weight in the middle of the distribution (due to the shrinkage) will have implications for the measurement of outcomes for all students. We address this issue by inflating the variance of the imputed test scores by a factor θ to align the variance of the imputed values with the variance observed in the real outcome data. The variance-inflation adjustment is not necessary for the graduation outcomes because they are not ranked.

Finally, we extend the imputation framework to examine the sensitivity of our findings to the potential for additional selection into state exit along unobserved dimensions. To do this, we parameterize different levels of selection into state exit, above and beyond what is captured by δ_4. We produce four alternative sets of imputed values assuming that the true state-exit mobility parameter is (1) 10 percent larger than δ_4, (2) 25 percent larger than δ_4, (3) 10 percent smaller than δ_4, and (4) 25 percent smaller than δ_4. That is, we allow for varying degrees of positive and negative selection into cross-state mobility, relative to within-state-cross-district mobility. We suppress the results from the sensitivity analysis for brevity because our findings are not meaningfully sensitive to changing the selection conditions.

Appendix D: Estimating District Value-Added

We use the larger state data samples of all students in grades 4-8 to estimate district value added with a two-step model based on Parsons, Koedel, and Tan (2019):

$$Y_{ijdkt} = \gamma_0 + \boldsymbol{Y_{i(t-1)}}\boldsymbol{\gamma_1} + \boldsymbol{X_{it}}\boldsymbol{\gamma_2} + \boldsymbol{S_{kt}}\boldsymbol{\gamma_3} + \boldsymbol{L_{dt}}\boldsymbol{\gamma_4} + \varepsilon_{ijdkt} \quad \text{(D1)}$$

$$\varepsilon_{ijdkt} = \phi_d + \nu_{ijdkt} \quad \text{(D2)}$$

In equation (D1), Y_{ijdkt} is the test score of student *i* in subject *j* taken at district *d* in school *k* at time *t*, which is standardized by subject, grade, and year within each state. $\boldsymbol{Y_{i(t-1)}}$ is a vector of test scores in math and ELA taken by student *i* the previous year. $\boldsymbol{X_{it}}$ is a vector of characteristics of student *i* in time *t* that includes information on the student's FRL status, IEP status, gender, race, English as a second language (ELL) status, and geographic mobility. $\boldsymbol{S_{kt}}$ and $\boldsymbol{L_{dt}}$ contain the variables included in $\boldsymbol{Y_{i(t-1)}}$ and $\boldsymbol{X_{it}}$ aggregated at the school and district levels, respectively, and ε_{ijdkt} is the error term.

In equation (D2), the error term from equation (D1) is regressed on a vector of district indicators to recover district value added, ϕ_d, by subject *j*. We then combine the subject-specific estimates to summarize district value-added to both subjects using the weighting approach of Lefgren and Sims (2012), which also inherently shrinks the value-added estimates toward the mean in a regression-based framework (this is similar to Chetty, Friedman, and Rockoff, 2014).

A desirable feature of the two-step modeling structure described by equations (D1) and (D2) is that variation in achievement attributable to student and district characteristics is partialled out in the first equation. The resulting value-added estimates from the second equation are orthogonal to these characteristics by construction. This is useful when we correlate value-added with our estimates of academic mobility at the district level, as it rules out some explanations for the relationships we find. Parsons, Koedel, and Tan (2019) also show that estimates from a two-step model of this form are less biased than more common "one-step" models under student-teacher sorting conditions that have been shown to be the most prevalent in practice.

Data for students in grades 4-8 from the entire panel period in each state are used to estimate district value added. All students in the analysis cohorts are omitted from the models in order to remove any mechanical correlation between our academic-mobility and value-added metrics. That is, the value-added models are jackknifed around the focal cohorts we use to study academic mobility, but otherwise cover the timeframe of their enrollment.

Appendix E: Connecting Academic and Economic Mobility

E.1 Overview

CHKS (2014) and CH (2018) document geographic heterogeneity in *intergenerational* economic mobility (EM) and suggest that differences in schools may be a contributing factor. In this appendix, we explore this possibility using our measures of *intragenerational* academic mobility (AM). Given well-documented relationships between family income and student achievement (Jang and Reardon, 2019; Reardon, 2011), and student achievement (and achievement-inducing interventions) and earnings (Chetty et al., 2011; Chetty, Friedman, and Rockoff, 2014b; Lazear, 2003; Murnane et al., 2000), it is reasonable to hypothesize that all else equal, areas with higher AM will have higher EM. The relationships we consider between AM and EM are not causal, but nonetheless help us understand how these concepts are related across geographies.

Two technical issues with connecting the AM and EM metrics merit discussion. First, the time frames for the metrics are misaligned. Linking the estimates is still useful if we assume that some aspects of place that contribute to the different types of mobility are fixed, but some divergence should be expected due to time inconsistency (note that due to data limitations, it is not possible for us to go back further in time to better align our measures to the CHKS and CH measures). Second, CHKS (2014) estimate EM at the commuting zone (CZ) level, CH (2018) estimate EM more narrowly at the county level, and we estimate AM even more narrowly at the school-district level. We align our metrics at the same geographic levels of CHKS (2014) and

CH (2018) by re-estimating AM at the commuting-zone and county levels, respectively, in our sample states.[32]

An immediate question in the aggregation is how much variation in AM occurs within versus between levels of aggregation. To answer this question, we estimate district-level regressions where the dependent variable is $\bar{O}_{25d}$ and the independent variables consist of a vector of indicator variables for either commuting zones or counties. We interpret error-corrected R-squared estimates from these regressions as estimates of the cross-commuting-zone, or cross-county, variance in AM. This is the only variance in AM that we can feasibly connect to the estimates from CHKS (2014) and CH (2018); i.e., we must throw out all within-geography variance in AM in these comparisons.[33]

Appendix Table E1 shows results from the within-between variance decompositions of AM at the commuting zone and county levels. First, at the commuting zone level, it is clear that most of the variance in AM occurs within, not between, commuting zones. This is especially apparent when we measure AM in terms of test scores, where the average cross-CZ variance shares across states are just 0.13 and 0.16 for the 8th-grade and high-school tests, respectively. For graduation outcomes, the cross-CZ variance shares are larger but still small, at 0.24 and 0.29 for on-time and late graduation, on average across states. The implication is that most of the variance in AM across districts cannot be connected to variance in EM across commuting zones. This result is not entirely surprising. Commuting zones are large areas and education research

[32] Our seven sample states include at least some coverage of 188 CZs for which economic mobility metrics are available from CHKS. We focus on the 165 of these CZs for which at least 50 percent of the population resides in one of our sample states. We have full coverage over the county-level estimates from CH.

[33] For the error correction, we use output from the randomized inference procedure described in the main text to estimate the share of the variance in the district O25 estimates that reflects true variance (net of estimation error), then divide the raw R-squared values by this ratio, as in Aaronson, Barrow and Sander (2007). This adjustment re-scales the R-squared to be over the range of explainable variance in the dependent variable. Note that we also confirmed the results are similar if we use a simple "squared standard errors" approximation for the error variance in the dependent variable.

consistently shows the greatest impacts of interventions at narrower localities—e.g., individual differences between teachers are larger than differences between schools, which are larger than differences between districts, etc.[34]

Using the results from the within-between variance decompositions, we conduct a bounding exercise to show that cross-CZ variance in AM cannot meaningfully account for observed variation in EM, even assuming away other problems related to research design and causal inference. This is because the variance in AM across commuting zones, translated into plausible impacts on earnings using the extant literature, is too small to explain documented cross-CZ variance in EM from CHKS. Our findings in this regard are consistent with related evidence from Rothstein (2019), who uses different data and methods but reaches a similar conclusion. We provide details on the bounding exercise below in Section E.2.[35]

At the county level, Table E1 shows that there is more between variation in AM, which is consistent with our expectation because counties are smaller geographic units (typically much smaller). On average across states, the cross-county variance shares in district-level AM as estimated for 8th-grade test scores, high-school test scores, on-time graduation, and late graduation are 0.28, 0.31, 0.50, and 0.60, respectively. These larger variance shares at least permit the possibility of a substantive link between AM and EM in terms of the scope of the variance, and particularly for test-based AM.[36] Correspondingly, we report correlations between AM and EM at the county level in Appendix Table E2, which are positive and in the range of

[34] This result is also 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 that differences in schools are an important driver of place-based effects on students' educational attainment using narrowly-defined geographic areas.
[35] 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).
[36] For graduation-based AM, although the cross-county variance shares are larger, the link between high-school graduation and earnings is weaker, limiting the scope for a relationship with EM—see below for details.

0.42-0.53 on average across states. With the important caveat that these correlations are not causal, they at least leave open the possibility of a meaningful link between AM and EM at the county level.

E.2 Details for the bounding exercise

In this section we expand on the calculations we use to assess the prospects for connecting our estimates of AM to the EM estimates from CHKS and CH. We focus on whether there is a sufficient scope of variance to connect the metrics, noting that the other challenges referenced above remain (namely the time-inconsistency between the measures and the lack of a research design for use in establishing a causal connection).

Focusing first on the CZ-level analysis, it is straightforward to calculate that a one-standard-deviation move in the CZ-level distribution of AM for each outcome is equal to the district-level standard deviation (from Table 8 in the main text) multiplied by the square root of the cross-district variance share (from Table E1). For our four focal outcomes, this calculation indicates effect sizes corresponding to moves in the CZ-level distributions that are 36-54 percent as large as analogous moves in the district distributions. Thus, approximate "effect sizes" of a one-standard-deviation improvement in the CZ-level distributions of AM imply test percentiles that are 1.7 and 2.0 percentile points higher for the eighth grade and high school tests, respectively, and graduation rates that 2.7 and 2.6 percentage points higher, respectively, for on-time and late graduation.

In order to link variation in our AM metrics to variation in CHKS' EM metrics, we must convert changes in education outcomes to changes in earnings. Many studies map improvements in test scores to earnings later in life. These studies typically report values based on test standard deviations, not percentiles, so we must convert our percentile-based AM numbers to standard deviations. If we assume test scores are normally distributed, a 2-percentile point gain on the high school test—i.e., roughly one cross-CZ standard deviation of academic mobility—assessed at the 25th percentile of the distribution maps to a 0.06 standard deviation increase in test scores. The extant literature indicates that higher (late-grade) test scores of this magnitude correspond to higher earnings in adulthood on the order of about 0.7-0.8 percent.[37] Similarly, the literature suggests that graduation rate gains of 2.6-2.7 percentage points—again, about one cross-CZ standard deviation of academic mobility—would be expected to correspond to higher earnings of 0.0-0.2 percent.[38]

These earnings gains can be further converted into *mobility in the earnings distribution*, which is the metric used by CHKS. The standard deviation of CHKS's O25 values across the commuting zones in our sample states is just over 3 percentile points, which supplementary data files from CHKS show corresponds to a change in income of about 9.5 percent. This implies that for test scores, a one-standard-deviation move in the academic mobility distribution across CZs, converted to income gains, would map to a move in the economic mobility distribution estimated by CHKS of about 0.07-0.08 standard deviations, and 0.00-0.02 standard deviations for graduation outcomes. These are small numbers and simple *ad hoc* tests confirm they are well below the thresholds at which statistical relationships could be detected using CZ-level data.

These same calculations can be applied to assess the scope for detecting county-level relationships between AM and EM. Here the potential for detecting relationships is more promising because there is much more variation in AM across counties than across CZs. The variance shares in Table E1 imply effect sizes corresponding to moves in the county-level

[37] These back-of-the-envelope calculations are based on the correspondence between later-grade test scores and earnings reported in Lazear (2003), Mulligan (1999), and Murnane et al. (2000). These studies report that a one-standard-deviation increase in later-grade test scores corresponds to higher earnings on the order of 11-14 percent.
[38] These back-of-the-envelope calculations are based on estimates of the earnings returns to high-school graduation reported in Castex and Kogan Dechter (2014), Clark and Martorell (2014), and Ferrer and Riddell (2008). These studies estimate that the earnings-returns to obtaining a high school diploma ranges between 0 and 8 percent.

distributions of AM that are larger—53-77 percent as large as analogous moves in the district distributions. Effect sizes of this magnitude still suggest a modest translation to earnings (using the same calculations described above), but at least have the potential to be meaningfully linked to CH's estimates of EM. Correspondingly, we report correlations between AM and EM at the county level in in Table E2.

To summarize, the fact that there is very little variation across CZs in AM effectively rules out AM as a driver of the cross-CZ variance in EM documented by CHKS. This is instructive about the scope for AM to affect EM, as it establishes a level of geography across which there is clear variation in EM but no scope for impact of AM. Still, it does not directly follow that variation in AM is ignorable—indeed, most of the variation in AM is across districts within CZs. The variance of AM across counties is larger, and at least allows for the possibility of a non-spurious relationship with EM based on the scope of the variance.

.

Table E1. Cross-commuting-zone and cross-county variance shares of $\bar{O}_{25d}$.

	Grade-8 Test		HS Test		Grad		Grad +1	
Commuting Zones								
	Between CZ	Within CZ	Between CZ	Within CZ	Between CZ	Within CZ	Between CZ	Within CZ
All (Avg)	**0.13**	**0.87**	**0.16**	**0.84**	**0.24**	**0.76**	**0.29**	**0.71**
GA	0.24	0.76	0.38	0.62	0.25	0.75	0.27	0.73
MA	0.10	0.9	0.08	0.92	0.1	0.9	0.11	0.89
MI	0.07	0.93	0.04	0.96	0.13	0.87	0.15	0.85
MO	0.08	0.92	0.16	0.84	0.27	0.73	0.38	0.62
OR	0.13	0.87	Not Applicable		0.57	0.43	0.62	0.38
TX	0.14	0.86	0.13	0.87	0.23	0.77	0.30	0.70
WA	0.12	0.88	0.14	0.86	0.11	0.89	0.20	0.80

	Grade-8 Test		HS Test		Grad		Grad +1	
Counties								
	Between County	Within County	Between County	Within County	Between County	Within County	Between County	Within County
All (Avg)	**0.28**	**0.72**	**0.31**	**0.69**	**0.50**	**0.50**	**0.60**	**0.40**
GA	---	---	---	---	---	---	---	---
MA	0.17	0.83	0.16	0.84	0.31	0.69	0.30	0.70
MI	0.25	0.75	0.23	0.77	0.34	0.67	0.37	0.63
MO	0.28	0.72	0.40	0.60	0.77	0.23	1.00	0.00
OR	0.27	0.73	Not Applicable		0.71	0.29	0.76	0.24
TX	0.38	0.62	0.36	0.64	0.53	0.47	0.67	0.33
WA	0.34	0.66	0.39	0.61	0.35	0.65	0.47	0.53

Notes: The table reports average (error-corrected) R-squared values from a regression of $\bar{O}_{25d}$ on a vector of either CZ indicators or county indicators in each state. One minus these values gives the average within-CZ and within-county variance shares. In GA, districts are essentially synonymous with counties (with a handful of exceptions), rendering the county-level variance decomposition uninformative. Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results.

Table E2. Correlations of county-level estimates of $\bar{O}_{25}$ and county-level intergenerational economic mobility estimates from Chetty and Hendren (2018).

	Grade-8 Test	HS Test	Grad	Grad +1
All (Avg)	***0.42***	***0.49***	***0.54***	***0.53***
GA	0.43	0.53	0.38	0.29
MA	0.71	0.86	0.91	0.91
MI	0.51	0.35	0.60	0.60
MO	0.66	0.44	0.84	0.84
OR	0.07	Not Applicable	0.56	0.53
TX	0.06	0.07	0.14	0.07
WA	0.50	0.69	0.37	0.47

Notes: Oregon does not offer a high school test taken in a (near) universal grade, so Oregon is omitted from the HS test results.