Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 23-729

# Understanding High Schools' Effects on Longer-Term Outcomes

Preeya P. Mbekeani
American Institutes for Research

John P. Papay
Brown University

Ann Mantil
Brown University

Richard J. Murnane
Harvard University

Improving education and labor market outcomes for low-income students is critical for advancing socioeconomic mobility in the United States. We use longitudinal data on five cohorts of 9th grade students to explore how Massachusetts public high schools affect the longer-term outcomes of students, with a special focus on students from low-income families. Using detailed administrative and student survey data, we estimate school value-added impacts on college outcomes and earnings. Observationally similar students who attend a school at the 80th percentile of the value-added distribution instead of a school at the 20th percentile are 11% more likely to enroll in college, are 31% more likely to graduate from a four-year college, and earn 25% (or $10,500) more annually at age 30. On average, schools that improve students' longer-run outcomes the most are those that improve their 10th grade test scores and increase their college plans the most.

VERSION: February 2026

Suggested citation: Mbekeani, Preeya P., John Papay, Ann Mantil, and Richard J. Murnane. (2026). Understanding High Schools' Effects on Longer-Term Outcomes. (EdWorkingPaper: 23-729). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/qwp6-hk05

# Understanding High Schools' Effects on Longer-Term Outcomes

Preeya P. Mbekeani (corresponding author)
American Institutes for Research
pmbekeani@air.org

John P. Papay
Brown University
john_papay@brown.edu

Ann Mantil
Brown University
ann_mantil@brown.edu

Richard J. Murnane
Harvard University
Richard_Murnane@harvard.edu

February 19, 2026

## Abstract

Improving education and labor market outcomes for low-income students is critical for advancing socioeconomic mobility in the United States. We use longitudinal data on five cohorts of 9th grade students to explore how Massachusetts public high schools affect the longer-term outcomes of students, with a special focus on students from low-income families. Using detailed administrative and student survey data, we estimate school value-added impacts on college outcomes and earnings. Observationally similar students who attend a school at the 80th percentile of the value-added distribution instead of a school at the 20th percentile are 11% more likely to enroll in college, are 31% more likely to graduate from a four-year college, and earn 25% (or $10,500) more annually at age 30. On average, schools that improve students’ longer-run outcomes the most are those that improve their 10th grade test scores and increase their college plans the most.


**Keywords**: school effects, high schools, longer-run outcomes, value-added estimation, low-income students

## Introduction

Parents and policymakers have long cared about differences among schools in how they promote student learning, development, and well-being. State testing under No Child Left Behind (NCLB) served to publicize the widely understood fact that schools differ substantially in their students' average test scores, and such scores were treated as proxies for school quality. Test score levels, however, reflect students' family backgrounds and prior schooling as well as the effectiveness of their current schools. Moreover, test scores are only proxies for what many parents, educators, and policymakers really want to know – namely, whether schools promote the development of skills, knowledge, and attitudes that enable students to "flourish" as adults (Brighouse, et al., 2016). We know little about how high schools affect students' longer-run educational and labor-market outcomes and whether the schools in which students show more growth in measurable, short-term outcomes are those that also promote students' longer-run success.

These questions are of particular interest for students living in poverty, who tend not to enjoy the myriad advantages and opportunities for enrichment experienced by their peers from higher-income families (Kaushal, Magnuson, & Waldfogel, 2011). However, such students often attend under-resourced schools with less experienced teachers and limited access to high-quality instructional materials (Boyd et al. 2005; Clotfelter, Ladd, & Vigdor 2009; Kalogrides & Loeb, 2013; Lankford, Loeb, & Wyckoff, 2002). As a result, understanding the ways in which schools affect the longer-run outcomes and economic mobility of students living in poverty is central for efforts to promote equality of opportunity and social mobility (Duncan & Murnane, 2011; Chetty et al., 2020; Jackson et al., 2024).

These questions are critical for practitioners and policymakers. Understanding the important contributions that schools make to later-life outcomes and the substantial variation across schools, particularly those serving students from low-income families, can provide guidance for policymakers in assessing policies and practices aimed at improving outcomes. Furthermore, while standardized test scores have been the focus of state and federal accountability systems, there may be other pathways through which schools influence students' longer run trajectories. Information on whether and to what extent short-run measures of school effectiveness predict longer-run student success can be used to design more effective and meaningful systems of school improvement and accountability.

We make two key contributions to the existing literature on schools' impacts on student outcomes. First, we extend evidence on schools' impacts by examining two longer-run student outcomes of interest: four-year college completion and adult labor market earnings. Existing studies have primarily focused on college enrollment (e.g., Carrell et al., 2023; Jackson et al., 2020; Jackson et al., 2024; Jennings et al., 2015) or examined college completion or earnings among a narrow group of specialized schools (e.g., Deming et al., 2014; Kemple & Willmer, 2008). Second, we situate our study in a multi-dimensional conceptual framework of school effectiveness and examine the mechanisms for these longer-term impacts. Specifically, we explore whether schools that improve short-term outcomes more than expected also have larger-than-expected effects on longer-run educational attainment and earnings. We build on several studies (e.g., Carrell et al., 2023; Loeb et al., 2018; Jackson et al., 2020; Jackson et al., 2024) by examining schools' effects on multiple short-run measures and how these effects relate to college completion and earnings.

We use detailed longitudinal data on first-time 9$^{th}$ grade Massachusetts public-school students in the 2002-03 through 2006-07 school years, hereafter the 2003 to 2007 cohorts. We estimate school value-added models, conditioning on traditional measures found in state administrative datasets and a set of additional covariates from student surveys not typically found in administrative datasets. This permits us to better account for the sorting of students into high schools and to identify schools' causal impacts on these longer-run outcomes. We examine the impacts of schools on all students and also on students from low-income families. We find that the schools that are effective in serving low-income students are also effective in serving those from higher-income families.

We document substantial variation in high schools' effects students' longer-run outcomes. For the average student in our sample, the difference in the probability of enrolling in college associated with attending a school at the 20$^{th}$ percentile of the school effectiveness distribution as compared to one at the 80$^{th}$ percentile is 7.9 percentage points (or 11%), the difference in the probability of four-year college graduation is 13.6 percentage points (or 31%) and the difference in annual earnings is 25%, approximately $10,500 (in 2024 $) for students with earnings near the sample mean. Impacts are larger for low-income students, with earnings impacts of 35% (or $11,300) from attending a high school at the 20$^{th}$ percentile of the effectiveness distribution as compared to one at the 80$^{th}$ percentile. We show that our results are unlikely to be driven by differences in the unobserved characteristics of students enrolled in different high schools.

We find that schools that improve 10$^{th}$ grade test scores more have larger effects on both four-year college graduation and earnings. A one standard deviation increase in schools' value-added on 10th grade test-scores is associated with a 1.5% difference in adult earnings. We also find large and significant relationships between schools' impacts on students' college-going plans and their effects on students' longer-run outcomes. Relationships between other short-term outcomes (e.g., attendance; course-taking) and longer-run outcomes are also positive but smaller in magnitude.

Taken together, our results suggest that the high school a student attends has persistent impacts on their longer-term educational attainments and earnings. This is true for all students and particularly students from low-income families. Consequently, providing opportunities for students from low-income families to enroll in high value-added high schools has important implications for socioeconomic mobility and economic opportunity.

## Background

Efforts to assess school "quality" in the United States have largely focused on student standardized test scores. This proxy measure has been used extensively, both by researchers and in state accountability systems, particularly as a result of NCLB (Linn, 2008; Loeb et al., 2018). However, research has made clear that test score levels are highly influenced by students' family backgrounds, neighborhoods, and prior schooling experiences. In other words, student sorting, on both observed and unobserved characteristics, must be addressed to learn how schools differ in their effectiveness (Angrist, et al., 2021; Angrist, Hull, & Walters, 2022; Raudenbush & Willms, 1995).

These concerns have led to a second wave of analyses that use "value-added" or test score growth as measures of school quality. In the typical value-added model preferred by researchers, the outcome is a student test score and the covariates include a test score from a prior grade (Angrist, Hull, & Walters, 2022). Reardon (2018) and Reardon and colleagues (2019) have shown that, at the school level, the within-cohort growth in test scores is largely uncorrelated with the average test-score levels of entering students, and within-cohort growth is closely related to longitudinal growth measures. Policy efforts have also shifted to holding schools accountable for student academic growth occurring over the course of a year. For example, many states use student growth percentiles, a distinct but related approach, that measures student performance in a given year relative to the set of students with the same test scores in a prior year (Betebenner, 2009).

While these measures that aim to assess students' academic growth or learning over the course of a year are substantial improvements over systems that do not take into account students' prior achievement, relying exclusively on standardized test scores reflects an overly narrow view of what schools can do to improve students' long-run opportunities (Loeb et al., 2018; Jackson et al., 2020; Jennings et al., 2015). Some standardized tests measure only basic skills; most do not assess the full range of academic knowledge and skills in core subjects.

Furthermore, schools do many things beyond raising test scores that matter for students' longer-run success. Indeed, society's aims for schooling go beyond academic learning. Among other things, schools are expected to develop students' social and behavioral skills, including teamwork and personal responsibility; schools are also expected to prepare students to engage constructively in democratic citizenship (Brighouse et al., 2016; Gutmann, 1999). Research across a range of fields has demonstrated that a variety of skills, attitudes, and behaviors – such as self-discipline, motivation, and work habits – matter for educational attainment and labor market outcomes (Allensworth & Clark, 2020; Heckman & Rubenstein, 2001; Duckworth & Seligman, 2005; Deming, 2017). Several recent studies have used value-added approaches to explore how schools promote outcomes other than test scores, such as social-emotional learning and academic engagement (e.g., Jackson et al., 2020; Loeb et al., 2018). We contribute to this literature by examining high schools' impact on several shorter-run non-test score outcomes.

Of course, while short-term measures of development are important, they do not address the ultimate impacts of schooling on students as adults. In recent years, advances in state longitudinal data systems have made it possible to examine a greater variety of longer-run student outcomes. Some of the strongest evidence of schools' effects on college enrollment, college degree completion, and criminal justice involvement come from quasi-experimental studies exploiting school lotteries or other plausibly exogenous assignment mechanisms (e.g., Angrist et al. 2016; Angrist et al. 2017; Deming, 2011; Deming et al., 2014).

These studies show that school value-added models that incorporate rich sets of covariates, including prior test scores, generally recover the causal impacts of schools on student test scores. For example, using data from a school district in North Carolina, Deming (2014) found that nonexperimental estimates of school value-added nearly matched estimates from the randomized lottery-based study. The results indicate that the covariates typically available in state administrative datasets are largely sufficient to account for the sorting of students across schools.

One critical question is whether high schools' effects on longer-run outcomes, such as educational attainments and earnings, are driven by their impacts on student academic learning as measured on standardized tests. Few studies have examined this directly, and the existing evidence is mixed. Angrist and colleagues (2016) reported that charter school effects on four-year college enrollment were strongly correlated with those schools' effects on test scores. Jennings et al. (2015) examined a broader group of schools and found schools' effects on college attendance are larger than and imperfectly correlated with their effects on test scores. By contrast, Deming (2011) found no evidence of impacts on test scores for high-risk male lottery winners compared to their peers who lost out in the same lotteries, despite finding impacts on later criminal activity. Similarly, Kemple and Willner (2008) found that attending a Career Academy high school, which combined academic and technical curricula, had no measurable effect on student test scores but improved longer-run earnings by 11%. More recently, Jackson and colleagues (2020) found that high schools' impacts on students' social-emotional development were more predictive of postsecondary outcomes than schools' test-score value-added.

The validity of value-added models for examining longer-run outcomes, such as college enrollment, completion, and earnings, for which there are no lagged measures, is less established. However, several studies have used value-added models to examine college enrollment (Carrell et al., 2023; Jackson et al., 2020; Jackson et al., 2024; Jennings et al., 2015) and demonstrated little bias. We contribute to this literature by estimating models that include a richer set of covariates than past studies and find them very comparable to estimates from models using the more limited measures typically available in state administrative data. We also report the results of validity tests supporting our position that our results are not driven by selection bias.

**A Multi-Dimensional Conceptual Framework of School Effectiveness**

We ground our study in a multi-dimensional conceptual framework of school effectiveness related to students' longer-run outcomes, which we illustrate in Figure 1. We show in bold type the elements we examine in the present study. For simplicity, we name four overlapping goals of K-12 schooling – supporting students' educational attainments, developing students' ability to succeed in the labor market, promoting civic engagement and participation, and supporting students' long-term health and well-being.

<Insert Figure 1 about here>

Schools can promote student success by influencing a range of capacities and dispositions that contribute to these outcomes (Brighouse et al., 2016). The most obvious is academic knowledge and skills. Schools differ in the quantity and quality of learning opportunities available to students. They vary in the coherence, depth, and quality of instruction (Mehta & Fine, 2019; Rowan, 2011); access to advanced courses, such as Advanced Placement or dual enrollment in college courses; and how students are tracked into different course pathways, including the extent to which students receive advice and information about which courses to take.

Importantly, we can divide academic skills into two categories: those measured by state tests and those not. For example, state standardized tests wholly exclude some content areas (e.g., history/social studies in Massachusetts) and do not assess some critical skills in tested subjects (e.g., oral communication skills as part of English language arts). Unmeasured academic knowledge and skills are important for students' longer-run outcomes; in the present study, we use information on student course-taking as a proxy.

Beyond academics, schools influence a range of skills that pay off in the long run. For example, schools can develop students' employment skills; this may occur through hands-on learning experiences and opportunities that build capacities for specific careers. Classroom instruction, extracurricular activities, and school culture, including behavioral norms and disciplinary practices, shape students' social skill development. Schools may also have explicit curricula focused on building students' social-emotional skills. Finally, schools affect students' attitudes and dispositions about their futures. For example, schools influence students' knowledge of and ideas about postsecondary education and career options through counselors, teachers, peers, and other networks (Hoxby & Avery, 2013; Mantil, 2022; Mulhern, 2023). As such, they serve an important role in shaping students' aspirations and plans for future education and careers.

The effectiveness of schools in supporting the development of these skills, capacities, and dispositions depends on at least three factors: their priorities, their resources, and their students. In terms of priorities, schools may choose to emphasize some dimensions of learning or skill development more than others. Priorities may be influenced by a range of factors, including the goals of the local community, the priorities of the adults in the school, the school's position in a state's accountability system, and broader social forces. For example, a school facing sanctions may focus relatively more on raising high-stakes test scores than on developing students' socioemotional skills. By contrast, in the wake of the pandemic some schools have re-invested in efforts to support student well-being and social-emotional development.

School resources, broadly defined, are key determinants of how well schools can achieve their priorities. For example, between two schools equally interested in promoting academic skills development, one may do better because of greater financial resources, stronger leadership, more effective teachers, or a more robust curriculum. Most obviously, financial resources play a key role in shaping the quality of the educator workforce, physical facilities, and learning materials (Biasi, LaFortune, & Schönholzer, 2025; Jackson, Johnson & Persico, 2016; LaFortune, Rothstein, & Schanzenbach, 2018). Beyond the resources themselves, how efficiently they are deployed and the extent to which this deployment is aligned with school priorities shape how well schools can achieve their aims.

Finally, a key component of high schools is the student body. Peers matter, and student body composition has important implications not only for classroom learning but also for the social and relational capital students can access within and beyond school (Harris, 2010; Xu, Zhang, & Zhou, 2022). Possible mechanisms through which peers may positively impact a student's longer-run outcomes include promoting an engaged school culture, contributing to a college-going ethos, and providing connections to employment networks.

**Research Questions**

We focus on high schools' effects on three longer-term outcomes, college enrollment (2-year or 4-year), four-year college graduation, and adult labor market earnings. We then explore whether schools' impacts on these longer-term outcomes correlate with their impact on more proximal measures aligned to our framework. While we do not have direct measures of all the dimensions described above (e.g., employment skills), our data enable us to construct reasonable proxy measures for several shorter-term outcomes. As mentioned above, we examine the impacts of high schools on all students and note the cases in which the impacts of schools for students from low-income families differ from those for their peers from higher-income families. We address two questions: First, what is the variation in high schools' effects on college enrollment, four-year college graduation, and adult earnings? Second, are schools with larger than expected effects on these longer-run outcomes those that also improve test scores, improve students' attendance, influence their academic trajectories, and/or promote college going?

## Research Design

### Data & Sample

We use student-level data from the Massachusetts Department of Elementary and Secondary Education (DESE), including student demographic and enrollment information, standardized test scores, and rich data from questionnaires that students complete when they take state standardized tests in $8^{th}$ and $10^{th}$ grade. DESE collects information on college outcomes from the National Student Clearinghouse (NSC). NSC data include nearly all colleges and universities (public, private, 2-year, and 4-year) in the United States. The coverage rate in Massachusetts approached 95% in 2011 and has improved since then (Dynarski, Hemelt, & Hyman, 2015; National Student Clearinghouse, 2022). De-identified data on labor market earnings come from the Massachusetts Unemployment Insurance (UI) system, which matches records to directory information for former public-school students.

Our analytic sample consists of students who entered $9^{th}$ grade for the first time during the 2002-03 through 2006-07 school years. We focus on these years because they provide a sufficient time horizon to examine our longer-term outcomes of interest, particularly earnings. For our earnings analyses we focus on the two earliest cohorts, 2003 and 2004, due to limitations of the administrative data. Labor market earnings vary substantially throughout the decade after students graduate from high school, as youth are in and out of college and the workforce and may not have consistent employment. Individuals' placements in the earnings distribution tends to stabilize by age 30 (Chetty, Hendren, et al., 2014). Therefore, we examine earnings when the average student in our sample is 30 or 31.

We include students enrolled in all public high schools in Massachusetts, with the exception of those enrolled in career and technical education schools or in one of Boston's three exam schools. We also limit our sample of schools to those that are present in the data all five years, resulting in a total of 258 high schools including 21 charter schools and 54 schools located in urban areas. We use eligibility for free- or reduced-price lunch (FRPL) in 8$^{th}$ grade as our indicator of low family income and refer to students who qualify for FRPL as "low-income" and students who do not as "higher-income."[1]

We limit our analytic sample to students who were in a Massachusetts public school in 8$^{th}$ grade and have non-missing data on 8$^{th}$ grade test scores, FRPL eligibility, special education status, English learner status, and attendance as these are important covariates in our model. Our analytic sample excludes approximately 10% of first-time 9$^{th}$ graders in the state due to missing 8$^{th}$ grade test scores; nearly all of these are students who first entered Massachusetts public schools in 9$^{th}$ grade. In Table 1, we present descriptive characteristics for the students included in our study. Approximately one-quarter of these students are low-income. On average, students from low-income families in our sample have substantially lower 8$^{th}$ grade test scores (-0.57 standard deviations) than those from higher-income families.

[1] Ideally, we would use a finer-grained measure of family income, such as number of years of FRPL eligibility (Michelmore & Dynarski, 2017). However, because our dataset begins with the 2001-2002 school year, we are only able to observe one prior year of data for our first cohort. We opted to use a consistent approach across the five cohorts in our sample.

<Insert Table 1 about here>

**Measures**

We examine two longer-run educational outcomes. The first is whether a student enrolled in any college (a 2-year or 4-year institution) within six years of entering ninth grade, that is within two years of expected high school graduation. The second is whether a student graduated from a four-year college within ten years of entering ninth grade (by about age 24). More than 70% of students in our analytic sample enrolled in either a 2- or 4-year college and 43% of students had earned a four-year college degree. However, only 46% of students from low-income families enrolled in college and 15% earned a four-year degree. We focus on four-year college graduation rather than two-year degree attainment because of its stronger relationship with other longer-run outcomes, including increased civic participation, such as volunteering and voting, and positive health behaviors (Ma, Pender, & Welch, 2020). We find quite similar results if we examine whether a student earned any degree within the same time period.

Our third longer-run outcome is adult labor market earnings. We express in 2024 dollars individuals' log average quarterly earnings from the three years when students were approximately age 30 (14 to16 years after 9$^{th}$ grade).[2] We omit quarters with no reported earnings and multiply the average quarterly value by 4 to reflect implied annual earnings. For succinctness, we refer to this outcome interchangeably as adult earnings or earnings at age 30. The median earnings of the students in our sample was approximately $58,700 for students overall and $42,400 for low-income students.

We lack earnings data for approximately 30% of the analytic sample. The state UI data do

[2] This is our preferred earnings outcome because it avoids potentially underestimating earnings from students being in and out of the workforce over the course of a year. We create an alternative earnings outcome measure—the three-year average of students' total annual earnings, expressed in 2024 dollars and logged. We describe these supplementary results in Footnote 8.

not include earnings from self-employment, employment in the federal government or military, under-the-table pay, or earnings from employment in another state. As such, we cannot detect whether individuals with no reported earnings are unemployed, out of the labor force, have non-reported earnings, or were not matched. Rather than assuming no earnings among these individuals, we omit them from our analyses of earnings as recommended by Foote & Stange, 2022.[3] Unsurprisingly, students who leave Massachusetts for college and low-income students who do not earn a high school degree are more likely to be missing earnings data, but probably for different reasons. We are reassured that our estimates reflect impacts on true earnings for several reasons. First, Opportunity Insights' Opportunity Atlas reports earnings at approximately the same age for students who grew up in Massachusetts (Chetty et al., 2018). Nearly 70% of students remain in the same commuting zone, echoing the share of students in our sample for whom we have reported earnings. In their data, students who leave the commuting zone have reasonably similar earnings (about 5% higher) than those who stay in the zone. Furthermore, estimates from the American Community Survey (ACS) indicate that of 29–31-year-olds born in Massachusetts who still live in state and are employed, 94% are employed in state while 6% are employed out of state.[4] This suggests that relatively few students who remain in the state have out-of-state earnings.

We constructed four short-term high-school outcomes aligned to our framework of school effectiveness: (1) a test score index, (2) an index indicating whether the student is academically on-track, (3) an attendance index, and (4) an indicator of whether a student plans to attend a four-year college after high school. We standardize these four measures across the sample so that they

---

[3] We also fit models estimating the impact of high schools and short-run high school outcomes on the probability that a student was employed at age 30. The models had very little explanatory power and no striking results.

[4] Authors' calculations using ACS 2018 1-Year Public Use Microdata Sample (PUMS) available from the U.S. Census Bureau.

have a mean of zero and standard deviation of one. We do this for ease of interpretation and to facilitate comparisons across these outcomes. We construct these measures using the DESE data and 10th grade student questionnaire responses, as follows.[5]

- Test-score index: We form a standardized composite of students' 10th grade mathematics and ELA MCAS scores by averaging each student's standardized raw scores.
- Academic on-track index: We form the academic index from student reports of their 9th and 10th grade math courses and an indicator calculated from the administrative data of whether they took their 10th grade MCAS tests on time based on when they entered 9th grade. We create a composite of these correlated items by using weights derived from the first component of a principal components analysis (PCA).
- Attendance index: Following recent literature (Jackson, 2018; Kraft, 2019), we create an index that includes a student's attendance (logit function) and two indicators of whether they were suspended in-school and out-of-school in their test year. We again use weights derived from the first component of a PCA to form the index.
- College plans: We create a binary measure using information from the 10th-grade questionnaire on whether a student planned to attend a four-year college after high school. We refer to this measure as college plans.

Our central question predictor is the first public high school a student attended. Approximately 11% of students transfer between schools during their high school careers. Given that transfer is endogenous and may reflect the policies and practices of the first high school, we focus on first high school in an intent-to-treat-style analysis.

[5] The questions and response options are presented in the Online Appendix. All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's website and use the search engine to locate the article at http://onlinelibrary.wiley.com.

We measure all covariates in 8th grade, before students enroll in high school. As stated above, we use eligibility for FRPL in 8th grade as our indicator for family income. In 2004, the maximum annual income for reduced-price lunch eligibility for a family of four was $34,873 (equivalent to approximately $58,000 in 2024 dollars). We use information on gender, race/ethnicity, attendance, and whether a student received special education services or was classified as an English learner. As shown in Table 1, 4.5% of students identify as Asian, 7.9% as African-American or Black, 8.8% as Hispanic, less than 1% as Native American, and 79% as White. Fourteen percent of students received special education services (SPED), and about 3% were English learners (EL). We calculated each student's 8th-grade attendance rate by dividing days in attendance by days in membership. The average 8th-grade attendance rate among students in our sample was 95%.

A critical covariate is each student's score on the Massachusetts Comprehensive Assessment System (MCAS) grade-8 mathematics examination, the only MCAS test required for 8th graders in all of these cohorts. We standardize students' grade-8 raw scores across the population of test takers in the state for each cohort, such that the average score in each statewide cohort is zero, with a standard deviation of one. We refer to the previously described 8th grade measures (test scores, race/ethnicity, gender, FRPL, SPED, EL, and attendance rate) as "Standard" controls, as they are typical measures in statewide longitudinal data systems and common control variables in value-added models.

As part of the MCAS administration in these years, students responded to questionnaires that asked them about their course-taking, their plans for education after high school, and their parents' highest level of education (2003 and 2004 cohorts only). Given that the questionnaires changed somewhat each year, we identified a set of common survey items across the five cohorts, and we convert the categorical responses to indicator variables. We refer to these 8th grade survey data as "Expanded" controls, as they are not typically available in statewide data systems. We present the text of the common items we use in our analyses in Online Appendix. To create the lagged measure of college plans, we construct a binary measure that combines responses in which students stated they planned to continue their education beyond high school. Because we only have information on parent education for two of the cohorts, we use this important family characteristic in validity tests of our school value-added estimates. One limitation of these survey data is missingness due to both item and survey non-response. To retain the approximately 10% of students in our sample who are missing survey responses, we include an indicator for whether a student's response is missing. In the validity tests using parent education, we limit our sample to students who have non-missing information on parental education from the 8th grade survey.

Finally, to account for peers, we construct a composite measure of the student-body composition in grade 9, using peer achievement, peer income status, and peer college plans. Peer achievement is the leave-out average (i.e., omitting a student's own score from the average) of the 8th-grade mathematics test scores of students in the same cohort and high school, and peer income status is the leave-out proportion of students in the same cohort and high school who are not eligible for FRPL. To create the measures of peers' educational plans, we calculate the proportion of students in the same cohort and high school who indicated in 8th grade that they planned to attend a four-year college. We construct the composite of these three measures using PCA as described above.

**Analytic Strategy**

We estimate high-school effects on our longer-run outcomes using school value-added models (VAMs) that estimate a school's effect on a given student outcome conditional on a student's prior achievement and demographic characteristics (see e.g., Jackson et al., 2020; Jackson et al, 2024; Jennings et al., 2015; Loeb et al., 2018; Lloyd & Schachner, 2021). Conceptually, the value-added of a given school $j$ represents its contribution to student $i$'s outcome relative to the student's expected outcome in an average high school. These measures are relative – they do not assess the impact of a student attending a given high school compared to not attending school. Below, we provide evidence supporting the validity of these models for estimating the causal effects of high schools on students' short- and longer-term outcomes.

We first estimate school-average value-added, using models of the following the general form:

$$Y_{ijt} = \mathbf{B}\mathbf{X}_{ijt} + \mathbf{\Gamma}\mathbf{E}_{ijt} + \zeta_t + \alpha_j + \epsilon_{ijt} \quad \text{Eq. (1)}$$

where $Y_{ijt}$ represents one of our longer-term outcomes for student $(i)$ in school $(j)$ and cohort $(t)$. The standard lagged control variables, including a cubic function of test scores and attendance, are represented by the vector $\mathbf{X}_i$, the expanded controls from the 8th grade questionnaire are represented by the vector $\mathbf{E}_i$. We pool data across the five cohorts and include fixed effects for cohort, $\zeta_t$. To account for the correlation between schools and covariates, we include school-specific intercepts, $\alpha_j$, following Chetty, Friedman, and Rockoff (2014) and Jackson et al. (2024). The vector $\mathbf{B}$ is estimated across students within the same school, and $\epsilon_{ijt}$ is the within-school individual-level error term.

We extract the combined residual, which we denote as, $u_{ijt}$:

$$u_{ijt} = \alpha_j + \epsilon_{ijt} \quad \text{Eq. (2)}$$

We collapse the residuals to the school level to generate our estimate of the impact on a given outcome $Y$ of attending a given school, where $N_j$ is the total number of students in school $j$:

$$\hat{u}_j = \frac{1}{N_j}\sum_{i=1}^{N_j} u_{ijt} \qquad \text{Eq. (3)}$$

We first estimate school effects using all students in our analytic sample. However, since we are especially interested in schools' impacts on students from low-income families, we also fit separate models using the subsamples of low- and higher-income students. We limit these analyses to schools-income cells with at least 10 students.

Given the uncommonly rich data available to us, we are able to control for a fuller set of lagged covariates than is typical in the value-added literature. Ideally, we would also condition on two years of prior test scores; however, we do not have 7$^{th}$ grade test score data for the earliest cohort in our sample. Given that we have these data for some of our cohorts as well as information on parent education from the student questionnaire for two of our five cohorts, we use this additional information to conduct several validity checks of our value-added estimates.

Our second research question concerns whether a school's impacts on short-term measures explain the variation in a school's effects on longer-term outcomes. We explore this in two ways. First, we fit value-added models as described previously, but with each of the four near-term measures as the outcome. Here, we can control for lagged outcomes.[6] We estimate correlations between schools' effects on these measures and their effects on our longer-run outcome measures. These correlations provide descriptive evidence of whether schools with larger than expected effects on longer-run outcomes are also those that improve test scores, improve students' attendance, promote college plans, or support students to stay academically on-track. Given the imperfect reliability of these measures, our estimates of these correlations may be biased

[6] Each of these value-added models control for lagged outcomes; the one exception is that we lack information on prior suspensions because that information is not available in the state administrative data.

downwards. Following Jackson et al. (2024), we use a split-sample approach to disattenuate our correlations for measurement error. For each value-added measure, we estimate the reliability by correlating the value-added estimates in even years with those in odd years. Then, using these estimated reliabilities, we calculate adjusted correlations by multiplying the estimated raw correlation by the inverse of the square root of the product of the reliability of each of the outcomes.[7]

We also address our second question with a model-based approach, including the short-term school value-added estimates as predictors. Doing so allows us to examine whether, after conditioning on students' prior test scores and demographic and academic characteristics, school value-added on each of these measures predicts our longer-run outcomes of interest. When using value-added estimates of our shorter-run measures to predict longer-term outcomes for students in a particular cohort, we employ the value-added with the drift procedure developed by Chetty, Friedman, and Rockoff (2014). This approach uses school-by-year value-added estimates, which are calculated as in Equation 3 but averaged over the number of students in each year ($N_{jt}$). It then excludes data for that cohort when estimating value-added to avoid mechanical correlation and places more weight on value-added for other years that are more highly correlated with the prediction year to aid precision. We refer to these leave-year out predictions as drift value-added. Additional details on the approach can be found in Chetty, Friedman, and Rockoff (2014) and Jackson et al. (2024).

---

[7] The formula is as follows: $r_{xy}^{*} = \frac{\hat{r}_{xy}}{\sqrt{r_{xx}r_{yy}}}$. There are two caveats to note regarding this adjustment. First, the estimated reliability likely does not account for all the sources of error in these measures; consequently, even the disattenuated correlations may be biased downward (Boyd et al., 2013). Second, the adjustment assumes that the errors in each measure are uncorrelated, resulting in upwardly biased estimates of the disattenuated correlations (Kraft, 2019). The true correlations may be lower if the errors are correlated across outcomes in the same year.

## Results

### High Schools' Effects on Students' Longer-Run Outcomes

We begin by describing results for all students and then report results for students from low-income families where they differ from those for students from higher-income families. High schools differ substantially in their estimated effects on four-year college graduation and earnings. For a ninth-grade student with sample average characteristics, the probability of four-year college graduation ranges from 28% to 57% based solely on the high school they attended. This reflects more than a doubling in the probability of graduating from a four-year college. The difference in the probability of four-year college graduation associated with attending a school at the 20$^{th}$ percentile compared to a school at the 80$^{th}$ percentile of the distribution is 13.6 percentage points (pp), a difference of approximately 31%. In Table 2, we present the standard deviation of the school average value-added estimates for our longer-term outcomes. The standard deviation of school value-added on any college enrollment is 5.0 pp and on four-year college graduation is 6.7 pp. For low-income students, the standard deviation of the college enrollment value-added is larger, 8.9 percentage points, but the standard deviation of four-year college graduation value-added is quite similar. For comparison, we also present school-value added estimates from random effect models, which are generally smaller given that they are shrunken.

<Insert Table 2 about here>

We also find substantial differences among high schools in their estimated effects on age 30 earnings. The standard deviation of the school value-added on log earnings is 0.14, or approximately 14% difference in later earnings for the students in our sample who attend a high school one standard deviation higher in school quality. For students from low-income families, the comparable difference is 23%. The difference in annual earnings associated with attending a school at the 20th versus 80th percentile rank of the school quality distribution is $11,300 for the average low-income student in our sample.

To facilitate comparison across the longer-run outcome measures, we scale the estimated effects on four-year college graduation and earnings outcomes by their sample standard deviations. We present these estimates in column 2 in Table 2. A one standard deviation improvement in school value-added raises students' probability of four-year college graduation by 0.11 standard deviations and earnings by 0.15 standard deviations. Given all the factors that influence students' later-life outcomes, these estimated school effects are quite large.

We illustrate the magnitude of these school effects in Figure 2. The first bar in each series represents the standard deviation in each outcome from an unconditional model that includes all students with only a fixed effect for cohort. This suggests substantial variation across schools in students' ultimate outcomes, specifically 0.17 for college graduation and 0.23 for earnings. These differences represent both selection of students to schools and schools' effects on outcomes. The remaining bars represent estimates of the standard deviations from different models, sequentially adding (a) 8th grade mathematics test scores (cubic function), (b) standard controls, and (c) expanded controls including students' post-high school college plans. The standard deviation of our estimated school effects from the final model represent 40% (standard deviation of 0.067) and 60% (standard deviation of 0.14), respectively, of the difference in four-year college graduation and earnings across schools. Confirming the results discussed below and supporting the validity of our value-added models, estimates from models with standard controls are similar, albeit slightly larger, to those with expanded controls, which include lagged measures of course-taking, college plans, and other family and school resources from the 8th grade student questionnaire. We further address the validity of our school value-added estimates in a later section.

<Insert Figure 2 about here>

**Relationships Between Schools' Effects on Short-Run and Longer-Run Outcomes**

We next explore whether schools' impacts on short-term measures predict their impacts on longer-term outcomes. We look at schools' success in improving all students' 10th-grade test scores, promoting college plans, promoting academic progress, and improving student attendance. In Table 3, we present both raw (below diagonal) and disattenuated (above diagonal) correlations between schools' value-added on short-term and longer-term measures. The patterns are similar. In nearly all cases, correlations are positive and substantively meaningful. We report disattenuated correlations in the text.

<Insert Table 3 about here>

Looking first at the relationships among the short-term measures, we find positive correlations among schools' effects on all measures. However, many of the correlations are relatively modest, suggesting that the short-term measures are capturing somewhat different dimensions of schooling. In particular, schools that are most effective in raising 10th grade test scores are not necessarily the schools that are most effective in promoting other short-term outcomes. Test score value-added has a modest correlation ($r$=0.32) with school value-added on college plans and even more modest correlations with academic on-track value-added ($r$=0.23) and the attendance index value-added ($r$=0.30).

Looking at the longer-run outcomes, we find that schools that promote four-year college graduation rates more tend to be those that improve students' test scores and college plans; patterns are nearly identical for college enrollment. Schools that increase college plans more (net of students' 8th grade plans) appear to be setting students up to graduate from a four-year college ($r$=0.68), the strongest relationship that we observe. Schools' effects on 10th grade test scores are also positively correlated with four-year college completion (*r*=0.58). In Panel A of Figure 3, we show the relationships between schools' effects on four-year college graduation and these two short-run outcomes, with test scores on the left and college-going plans on the right. Estimated correlations between schools' effects on four-year college graduation and both the academic on-track index (*r*=0.08) and the attendance index (*r*=0.35) are positive, but smaller in magnitude.

<Insert Figure 3 about here>

We see broadly similar findings when we look at earnings. Schools that improve 10$^{th}$ grade test scores more than expected raise earnings more (*r*=0.29). Even stronger is the relationship between school value-added in promoting college plans and the school's impact on earnings (*r*=0.84). These relationships are illustrated in Panel B of Figure 3. The relationship between schools' estimated impacts on the attendance index and earnings (*r*=0.29) is similar to the relationship between the attendance index and college graduation (*r*=0.35). The correlations between schools' value-added on the academic index and the value-added on earnings (*r*=0.21) and the value-added on four-year college graduation (r=0.08) are smaller.

Correlations provide insights into the relationships between these measures but not into their relative magnitudes. In Table 4 we present estimates of how each of our short-term value-added measures individually predicts four-year college graduation conditional on our lagged controls (Columns 1-4). A one-standard-deviation difference in the test score value-added is associated with a 2.9 percentage point difference in the probability of graduating from a four-year college (Column 1). In Column 5, we show the parameter estimates when we include all our short-term measures in the model. Echoing the relatively modest correlations among the short-term measures, each of these measures, with the exception of the academic on-track value-added, predicts longer-run outcomes when they are entered simultaneously in the regression model. The related coefficients on the other near-term measures range from just over one percentage point for the attendance index value-added to three percentage points for the value-added of students' college plans, conditional on the other measures.

<Insert Table 4 about here>

We also find strong evidence that schools that enhance academic skills and college plans have bigger impacts on four-year college graduation than the sum of their individual effects would indicate. As seen in Column 6, the interaction of test score value-added and the value-added in college plans is positive and statistically significant, albeit small with a coefficient of 0.6 percentage points. The estimated impact on four-year college graduation of improving test scores by one standard deviation in schools that improve college-going plans by one standard deviation is a striking 5.4 percentage points. This suggests that improving test scores matters much more in schools that also promote college plans and vice versa. Interestingly, we find a small negative coefficient on the interaction of test-score value-added and the value-added for being on-track academically.

Finally, we explore the extent to which schools' effects on four-year college graduation operate through their value-added impacts on college enrollment. We add school value-added on college enrollment as a predictor of four-year college graduation in Column 7. A one standard deviation increase in college enrollment value-added increases the probability of four-year college graduation by a statistically significant 2.7 percentage points. The short-term value-added measures remain significant but smaller predictors, suggesting that some, but not all, of schools' impacts on four-year college graduation operates through impacts on college enrollment.

As shown in Table 5, the pattern of results is quite similar for earnings. The three short-term value-added measures each predict earnings alone, conditional on lagged covariates

(Columns 1-4). A one standard deviation difference in schools' effects on 10th grade test scores is associated with a 0.03 difference in log earnings (in other words, a 3.1% difference in adult earnings). A one standard deviation increase in college plans value-added increases earnings by 0.46, or 4.7%. We also see meaningful and statistically significant impacts on earnings of schools' effects on attendance value-added. Schools' short-term value-added estimates remain significant and are reasonably strong, independent predictors when used jointly to predict earnings at age 30. One difference from the results for college graduation is that there is no statistically significant interaction effects between schools' effects on test-score and on non-test-score value-added (Column 6).

<Insert Table 5 about here>

**Relationship Between Effects on Attainments and Earnings**

We know that a four-year college degree is a pathway to higher earnings (Autor, 2014). But do schools that improve college-going also boost students' earnings? We find that they do. In Table 3, the correlation between schools' estimated effects on students' earnings and college enrollment (*r*=0.50) and four-year college graduation (*r*=0.69) are positive and relatively strong. In Figure 4, we plot high schools' estimated effects on earnings (vertical axis) and their effects on four-year college graduation (horizontal axis).

<Insert Figure 4 about here>

We explore more directly the extent to which schools' effects on earnings operate through improving postsecondary attainment by adding the value-added of any (two- and four-year) degree completion to the regression model in Table 5, Column 7. Some of schools' impacts on earnings operate through their impacts on college graduation. We see that a one standard deviation increase in schools' value-added on college graduation increases earnings by 1.7% and is marginally significant. Test-score value-added is no longer significant, suggesting that schools test-score impacts on earnings operate through college attainment. However, college plans and attendance value-added remain significant predictors of earnings indicating that high schools influence earnings above and beyond impacts on postsecondary educational attainments.[8]

**The Role of Peers**

As suggested in our framework (Figure 1), one way that schools may affect student outcomes is through the effects of peers. As seen in Table 6, our peer composite measure predicts longer-run outcomes, conditional on our rich set of lagged covariates. A one-standard-deviation difference in the peer composite is associated with a 2.9 percentage point difference in the probability of enrolling in college (Column 1), a 5.0 percentage point difference in the probability of graduating from a four-year college (Column 3), and a 3.7% difference in later earnings (Column 5).

<Insert Table 6 about here>

Peer effects partly operate through our short-term value-added measures, particularly for earnings. In Columns 2, 4, and 6 of Table 6, we show results from models that include our short-run outcome measures. In the case of earnings, approximately half of the relationship between the peer composite and earnings is explained by the short-run measures. In the case of college enrollment and four-year college graduation, the coefficients on the peer composite are reduced by about one-quarter when the short-run measures are included. This suggests that for college outcomes, while some of the effect of peers is explained by the short-run measures, peers matter above and beyond these near-term outcomes.

[8] We also examined schools' effects on two alternative versions of our main outcomes: any college degree and average annual earnings. In each case, estimated effects are very similar to the results estimated with our preferred outcome measures.

**Effects By Family Income**

We have primarily focused so far on how schools serve all students. However, knowing how high schools affect low-income students' longer-run outcomes is important, as schools are critical vehicles for promoting socioeconomic mobility for these students. As we noted above, there are some differences in the magnitude of schools' value-added on longer-run outcomes for low- and higher-income students but in general, the patterns are similar (Table 2). In Table 7, we show the results of examining explicitly whether schools' effects are similar for the low- and higher-income students they serve. We do this by fitting our models separately for the two groups of students and estimating the correlations between the value-added estimates for the two groups. All correlations are positive, suggesting that schools that improve outcomes for higher-income students also improve outcomes for lower-income students. Disattenuated correlations are all between 0.85 and 1, providing support for the conclusion that high schools that are effective for higher-income students are also effective for low-income students.

<Insert Table 7 about here>

We also examine whether there are schools serving large percentages of low-income students that have high value-added for these students. Using estimates from the models fit on the subsample of lower-income students, in Figure 5 we plot our estimates of school value-added in promoting four-year college graduation (vertical axis, left panel) or earnings at age 30 (vertical axis, right panel) against the percentage of low-income students in the school (horizontal axis). Two patterns are evident. First, the schools with the highest value-added in promoting both long-run outcomes for low-income students are schools in which most students come from higher-income families. Second, there are schools serving large percentages of low-income students that are quite effective in promoting long-run outcomes for these students.

<Insert Figure 5 about here>

## Threats to Validity

### Student Sorting

The central threat to validity is that, even after controlling for observed student and family characteristics, we have not sufficiently accounted for the sorting of students to schools in ways that are related to outcomes. We leverage our rich data to conduct several tests providing evidence on this point. We present the results in Table 8.

<Insert Table 8 about here>

A long history of research documents that parents' education is a strong predictor of student outcomes, including educational attainments, and a key driver of inequality (Bradbury et al., 2015; Coleman et al., 1966; Reardon, 2011). This led us to examine whether our value-added estimates were sensitive to the inclusion of indicators of parents' education as control variables. We have this information for students in our 2003 and 2004 cohorts. We compare our estimated school effects as reported above to estimates from models that include as additional controls categorical indicators of parents' highest education attainments (we take the highest across the parents).[9] As shown in the first row of Table 8, we estimate correlations of at least 0.98 between models with and without controls for parent education. In other words, adding parental education does not affect estimates substantially.

We also conducted two tests following Chetty, Friedman, and Rockoff (2014) to assess the validity of our school value-added estimates. The specification test examines whether a change in estimated school value-added corresponds to a one-to-one change in student outcomes. We regress

[9] The categories are: did not finish high school; graduated from high school; graduated from a two-year college, business school, or technical school; graduated from a four-year college; has an advanced degree (such as a master's or doctorate); I don't know. We exclude responses of "I don't know".

student outcome residuals on school value-added. Coefficients of one suggest unbiasedness. As shown in the second row of Table 8, coefficients are quite close to one and we cannot reject that they are different from one. As a result, a change in school value-added appears to correspond to a one-to-one change in students' short-run and longer-run outcomes.

A second validity check, a forecast bias test, examines whether our value-added estimates are correlated with other pre-high school measures that are not included in our base value-added model. If our school value-added estimates are uncorrelated with these leave-out variables, this indicates that any bias from omitted variables would have to be from factors that are themselves uncorrelated with the leave-out variables. Here we rely on our rich administrative data and conduct two versions of this test using two "leave-out" variables: further lagged scores, specifically 7$^{th}$ grade English language arts scores, and parent education, as reported by students in 8$^{th}$ grade. For both tests, we limit our sample to students with the relevant measure. For lagged test scores, this is students in the 2004-2007 cohorts, and for parent education this is students in the 2003 and 2004 cohorts. Results from both the forecast bias tests show significant but very small correlations between our value-added estimates and the leave-out measures. While we cannot completely rule out potential bias from omitted variables, missing predictors would have to be much stronger than the important ones that we include to generate substantial bias (Altonji, Elder, & Taber, 2005; Frank et al., 2013; Oster, 2017).

A final piece of validity evidence comes from the reasonably high intertemporal stability of the school effects estimates. The estimated school effects are generally quite stable across cohorts, as shown in Table 9, with the one exception being effects on college plans. Given that schools may change their effectiveness at different rates, and even at different rates across each of these measures, we would not expect perfect correlations in school effects across years.

Nonetheless, we interpret the general stability of school effects across cohorts as evidence supporting their validity.[10]

<Insert Table 9 about here>

Finally, we note an important limitation of our interpretation. While we have described our estimates as "school value-added" or "school effects" they are, instead, more accurately effects of schools and the bundle of other educational activities experienced by students in the school. For example, if a school is located next to an excellent after-school tutoring program that many students attend, the "school effect" we observe may in fact be produced by this program. The same is true for athletics programs or other community organizations that serve students in particular schools. While these impacts are not, per se, the result of the schools' efforts, we can read them directly as average impacts of attending the school. We see this as a point of interpretation and not a threat to the validity of our analysis.

## Discussion

We advance the literature on high school effectiveness by examining schools' effects on students' longer-run educational attainments and earnings. Using longitudinal data on 9$^{th}$ grade students in public high schools in Massachusetts in 2003 through 2007, we show that high school quality matters for students' probability of four-year college graduation and labor market earnings at age 30. We also find that schools' impacts on three of our short-term measures predict longer-run outcomes. In particular, schools that have larger than expected effects on students' 10$^{th}$ grade test scores and college-going plans also have larger than expected effects on educational attainments and earnings. Our findings are consistent with other recent studies that have found that high schools with higher test-score value-added increase college enrollment by 1-2 percentage

[10] We also estimate Spearman rank correlations and find even greater intertemporal stability, with correlations ranging from 0.57 to 0.98.

points (Carrell et al. 2023; Jackson et al. 2024). Importantly, we find that schools that are especially effective in improving long-run outcomes for students from low-income families are the same schools that are especially effective for students from higher-income families. If anything, impacts on students from low-income families are greater.

Our research documents that high schools can make substantial differences in the life outcomes of students, including those from low-income families. Which high school a low-income student attends has real implications for socioeconomic mobility. Furthermore, the schools that improve outcomes for low-income students the most are not only schools that serve relatively few economically disadvantaged students. In other words, some schools that serve high proportions of economically disadvantaged students have substantial positive impacts on their later life outcomes. These estimates align with those from Reardon (2018) who found that test-score growth was not as correlated with school-level socio-economic status as test-score levels.

We also find strong evidence supporting our multi-dimensional model of school effectiveness. We highlight two key results. First, schools influence students' longer-run outcomes through pathways beyond tested academic knowledge and skills. We know that a wide range of skills are important for future educational attainments and pay off in the labor market. We show that schools that improve proxy measures of these skills more tend to have larger impacts on longer-run outcomes. In particular, schools that improve students' academic skills and knowledge in tested areas, improve their college-going plans (what some scholars have called "college press"), and improve their attendance in turn support students' educational attainments and labor market earnings. This has important implications for school improvement and accountability, which we discuss below.

Second, our results suggest that improvement efforts aimed at a particular outcome could in fact enhance multiple outcomes. In other words, schools that are especially effective in increasing students' interest in attending college also tend to be the schools that promote test-score improvements successfully (and vice versa). This suggests that schools should not treat skills as separate domains and that a holistic approach to students' academic and social-emotional development appears to be more effective. This is particularly relevant for school and district leaders focused on school improvement efforts.

For policymakers and practitioners, these findings suggest possible next steps in designing measures of school effectiveness for accountability and improvement. From a policy perspective, it is helpful to know that schools' impacts on short-run measures predict their effects on longer-run outcomes. Ultimately, education stakeholders care about students' longer-run success, but assessing schools on these longer-run outcomes is largely impractical. We provide some evidence of the validity of shorter-term outcomes as markers of school effectiveness.

Relatedly, while standardized test scores have formed the cornerstone of accountability policies for schools and teachers, our results suggest that examining school effects across a broader range of outcomes beyond test scores may better account for the multiple ways in which schools influence students' longer-run life outcomes. Our results are consistent with ongoing policy efforts at the federal level and in many states to move beyond accountability systems based only on test scores. Further work should be done to develop and refine measures of students' experiences in schools, including, for example, student behavior, engagement, and college plans. Identifying measures that are difficult to manipulate is an important avenue for future research.

Accountability is not the only way in which a broader understanding of schools' impacts can inform practice. Many of the measures we describe, or improved versions of them supported by state and district data efforts, can provide diagnostic information to school and district leaders to foster ongoing improvement. One clear example here is schools' impacts on students' college plans, which seems to be an important pathway through which schools improve longer-run outcomes. While survey measures of college plans may be too easily manipulable to be valid indicators in an accountability system, they can be valuable metrics to help educators gauge their success in supporting students.

In short, developing systems of school support and accountability that take into consideration the various ways schools may influence students' longer-run outcomes is important. The end goal of any such system should be to incentivize not only changes in the measures themselves but the development of students' underlying skills, abilities, and capacities. Supporting schools to improve the broader social and academic skills of low-income students has real potential to reduce persistent opportunity gaps and advance socioeconomic mobility.

**Data Availability Statement**

The data used in this study are not publicly available due to privacy restrictions. The authors obtained access to these data under a data use agreement with Massachusetts Department of Elementary and Secondary Education that prohibits sharing the data with third parties. Researchers interested in accessing the data would need to request access directly from Massachusetts Department of Elementary and Secondary Education.

**Acknowledgements**

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305H190035 to Brown University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. The authors thank the Massachusetts Department of Elementary and Secondary Education for their ongoing partnership and support. We also thank Emma Zorfass and Aubrey McDonough for excellent research support.

**References**

Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher*, *49*(3), 198–211. https://doi.org/10.3102/0013189x20902110

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, *113*(1), 151–184.

Angrist, J. D., Cohodes, S. R., Dynarski, S. M., Pathak, P. A., & Walters, C. R. (2016). Stand and deliver: Effects of Boston's charter high schools on college preparation, entry, and choice. *Journal of Labor Economics*, *34*(2), 275–318.

Angrist, J. D., Hull, P. D., Pathak, P. A., & Walters, C. R. (2017). Leveraging lotteries for school value-added: Testing and estimation. *The Quarterly Journal of Economics*, *132*(2), 871–919. https://doi.org/10.1093/qje/qjx001

Angrist, J. D., Hull, P. D., Pathak, P. A., & Walters, C. R. (2021). Race and the mismeasure of school quality (Working Paper 29608). National Bureau of Economic Research. http://www.nber.org/papers/w29608

Angrist, J. D., Hull, P. D., & Walters, C. R. (2022). Methods for measuring school effectiveness (Working Paper 30803). National Bureau of Economic Research. https://www.nber.org/papers/w30803

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent." *Science*, *344*(6186), 843–851.

Biasi, B., LaFortune, J., & Schönholzer, D. (2025). What works and for whom? Effectiveness and efficiency of school capital investments across the U.S. *Quarterly Journal of Economics*, *140*(3), 2329–2379.

Betebenner, D. (2009). Norm- and criterion-referenced student growth. *Educational*

*Measurement: Issues and Practice*, *28*(4), 42–51. https://doi.org/10.1111/j.1745-3992.2009.00161.x

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American Economic Review*, *95*(2), 166–171. https://doi.org/10.1257/000282805774669628

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. https://doi.org/10.3102/1076998613508584

Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2015). *Too many children left behind: The U.S. achievement gap in comparative perspective*. New York, New York: Russell Sage Foundation.

Brighouse, H., Ladd, H. F., Loeb, S., & Swift, A. (2016). Educational goods and values: A framework for decision makers. *Theory and Research in Education*, *14*(1), 3–25. https://doi.org/10.1177/1477878515620887

Carrell, S., Kurlaender, M., Martorell, P., Naven, M., & Sun, C. (2023). Do schools matter? Measuring the impact of California high schools on test scores and postsecondary enrollment (Working Paper). National Bureau of Economic Research. https://conference.nber.org/conf_papers/f191771.pdf.

Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R., & Porter, S. R. (2018). The opportunity atlas: Mapping the childhood roots of social mobility (Working Paper 25147). National Bureau of Economic Research. https://www.nber.org/papers/w25147

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review, 104*(9), 2593–2632. https://doi.org/10.1257/aer.104.9.2593

Chetty, R., Hendren, N., Jones, M. R., & Porter, S. R. (2020). Race and economic opportunity in the United States: An intergenerational perspective. *The Quarterly Journal of Economics*, *135*(2), 711–783. https://doi.org/10.1093/qje/qjz042

Chetty, R., Hendren, N., Kline, P., Saez, E., & Turner, N. (2014). Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *American Economic Review*, *104*(5), 141–147. https://doi.org/10.1257/aer.104.5.141

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

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfield F. D., & York, R. L. (1966). *Equality of Educational Opportunity*. Washington, DC: U.S. Department of Health, Education, and Welfare. Office of Education.

Deming, D. J. (2011). Better schools, less crime? *The Quarterly Journal of Economics*, *126*(4), 2063–2115. https://doi.org/10.1093/qje/qjr036

Deming, D. J. (2014). Using school choice lotteries to test measures of school effectiveness. *American Economic Review*, *104*(5), 406–411. https://doi.org/10.1257/aer.104.5.406

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics 132*(4), 1593–1640.

Deming, D. J., Hastings, J. S., Kane, T. J., & Staiger, D. O. (2014). School choice, school quality and postsecondary attainment. *American Economic Review*, *104*(3), 991–1013. https://doi.org/10.1257/aer.104.3.991

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*(12), 939–944.

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

Dynarski, S. M., Hemelt, S. W., & Hyman, J. M. (2015). The missing manual: Using National Student Clearinghouse data to track postsecondary outcomes. *Educational Evaluation and Policy Analysis*, *37*(1S), 53S–79S. https://doi.org/10.3102/0162373715576078

Foote, A. & Stange, K. M. (2022). Attrition from administrative data: Problems and solutions with an application to postsecondary data (Working Paper 30232). National Bureau of Economic Research. http://www.nber.org/papers/w30232

Frank, K. A., Maroulis, S. J., Duong, M. Q., & Kelcey, B. M. (2013). What would it take to change an inference? Using Rubin's causal model to interpret the robustness of causal inferences. *Educational Evaluation and Policy Analysis*, *35*(4), 437–460. https://doi.org/10.3102/0162373713493129

Gutmann, A. (1999). *Democratic education: Revised edition*. Princeton, NJ: Princeton University Press.

Harris, D. N. (2010). How do school peers influence student educational outcomes? Theory and evidence from economics and other social sciences. *Teachers College Record*, *112*(4), 1163–1197.

Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review*, *91*(2), 145–149.

Hoxby, C. M., & Avery, C. (2013). The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity*, *2013*(Spring), 1-65. https://doi.org/10.1353/eca.2013.0000

Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, *126*(5), 2072–2106.

Jackson, C. K., Johnson, R. C., Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics 131*(1), 157–218.

Jackson, C. K., Kiguel, S., Porter, S. C., & Easton, J. Q. (2024). Who benefits from attending effective high schools? *Journal of Labor Economics, 42*(3), 717–751.

Jackson, C. K., Porter, S. C., Easton, J. Q., Blanchard, A., & Kiguel, S. (2020). School effects on socioemotional development, school-based arrests, and educational attainment. *American*

*Economic Review: Insights*, *2*(4), 491–508. https://doi.org/10.1257/aeri.20200029

Jennings, J. L., Deming, D., Jencks, C., Lopuch, M., & Schueler, B. E. (2015). Do differences in school quality matter more than we thought? New evidence on educational opportunity in the twenty-first century. *Sociology of Education*, *88*(1), 56–82. https://doi.org/10.1177/0038040714562006

Kalogrides, D., & Loeb, S. (2013). Different teachers, different peers: The magnitude of student sorting within schools. *Educational Researcher*, *42*(6), 304–316. https://doi.org/10.3102/0013189x13495087

Kaushal, N. M., Magnuson, K. A & Waldfogel, J. (2011). How is family income related to investments in children's learning? In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 187–206). New York, NY: Russell Sage Foundation.

Kemple, J. J., & Willmer, C. (2008). *Career Academies: Long-term impacts on labor market outcomes, educational attainment, and transitions to adulthood.* MDRC. https://www.mdrc.org/publication/career-academies-long-term-impacts-work-education-and-transitions-adulthood

Kraft, M. A. (2019). Teacher effects on complex cognitive skills and social-emotional competencies. *Journal of Human Resources*, *54*(1), 1–36.

Lafortune, J., Rothstein, J., & Schanzenbach, D. W. (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics*, *10*(2), 1–26. https://doi.org/10.1257/app.20160567

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, *24*(1), 37–62. https://doi.org/10.3102/01623737024001037

Linn, R. L. (2008). Methodological issues in achieving school accountability. *Journal of Curriculum Studies*, *40*(6), 699–711. https://doi.org/10.1080/00220270802105729

Lloyd, T., & Schachner, J. N. (2020). School effects revisited: The size, stability, and persistence of middle schools' effects on academic outcomes. *American Educational Research Journal*, *58*(4), 748–784. https://doi.org/10.3102/0002831220948460

Loeb, S., Christian, M. S., Hough, H. J., Meyer, R. H., Rice, A. B., & West, M. R. (2018). School effects on social-emotional learning: Findings from the first large-scale panel survey of students (Working Paper). Policy Analysis for California Education. https://files.eric.ed.gov/fulltext/ED591089.pdf.

Ma, J., Pender, M., & Welch, M. (2020). *Education pays 2019: The benefits of higher education for individuals and society* (Trends in Higher Education Series). College Board. https://research.collegeboard.org/media/pdf/education-pays-2019-full-report.pdf

Mantil, A. (2022). Crossing district lines: The impact of urban–suburban desegregation programs on educational attainments. *Educational Evaluation and Policy Analysis*, *44*(1), 127–148. https://doi.org/10.3102/01623737211030504

Mehta, J., & Fine, S. M. (2019). *In search of deeper learning: The quest to remake the American high school*. Cambridge, MA: Harvard University Press.

Michelmore, K., & Dynarski, S. (2017). The gap within the gap: Using longitudinal data to understand income differences in educational outcomes. *AERA Open, 3*(1), 1–18. doi:10.1177/2332858417692958.

Mulhern, C. (2023). Beyond teachers: Estimating individual guidance counselors' effects on educational attainment. *American Economic Review*, *113*(11), 2846–2893.

National Student Clearinghouse (2022, March). "Enrollment Coverage Workbook." *Working With Our Data*. Retrieved December 12, 2022 from https://nscresearchcenter.org/workingwithourdata/

Oster, E. (2017). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204. https://doi.org/10.1080/07350015.2016.1227711

Raudenbush, S., & Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, *20*(4), 307–335.

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

Reardon, S. F. (2018). Educational opportunity in early and middle childhood: Variation by place and age. (Working Paper 17-12). Stanford Center for Education Policy Analysis. https://edopportunity.org/papers/wp17-12-v201803.pdf

Reardon, S. F., Papay, J. P., Kilbride, T., Strunk, K. O., Cowen, J., An, L., & Donohue, K. (2019). Can repeated aggregate cross-sectional data be used to measure average student learning rates? A validation study of learning rate measures in the Stanford Education Data Archive. (Working Paper 19-08). Stanford Center for Education Policy Analysis. https://edopportunity.org/papers/wp19-08-v201911.pdf

Rowan, B. (2011). Intervening to improve the educational outcomes of students in poverty: Lessons from recent work in high-poverty schools. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances.* (pp. 523–537). New York, NY: Russell Sage Foundation.

Xu, D., Zhang, Q., & Zhou, X. (2022). The impact of low-ability peers on cognitive and noncognitive outcomes. *Journal of Human Resources*, *57*(2), 555–596. https://doi.org/10.3368/jhr.57.2.0718-9637R2

## Tables

**Table 1**

Student characteristics first-time 9th grade cohorts, 2003-2007.

| | Analytic sample | | Low-income students | | Non-low-income students | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| Demographics and Prior Achievement (8th Grade) | | | | | | |
| Asian | 0.045 | 0.207 | 0.076 | 0.266 | 0.035 | 0.184 |
| Black | 0.079 | 0.269 | 0.210 | 0.407 | 0.037 | 0.190 |
| Hispanic | 0.088 | 0.283 | 0.273 | 0.446 | 0.030 | 0.170 |
| Native American/American Indian | 0.003 | 0.054 | 0.004 | 0.065 | 0.002 | 0.050 |
| White | 0.786 | 0.410 | 0.435 | 0.496 | 0.895 | 0.307 |
| Female | 0.504 | 0.500 | 0.495 | 0.500 | 0.507 | 0.500 |
| Free/Reduced-price lunch | 0.238 | 0.426 | 1.000 | 0.000 | 0.000 | 0.000 |
| Special education | 0.139 | 0.346 | 0.214 | 0.410 | 0.115 | 0.319 |
| English learner | 0.030 | 0.171 | 0.100 | 0.299 | 0.008 | 0.091 |
| Parent completed four-year college | 0.545 | 0.498 | 0.227 | 0.419 | 0.619 | 0.486 |
| Attendance rate | 0.947 | 0.060 | 0.925 | 0.077 | 0.954 | 0.051 |
| Math standardized MCAS score | 0.100 | 0.978 | -0.565 | 0.888 | 0.308 | 0.909 |
| Plan to attend college | 0.783 | 0.413 | 0.716 | 0.451 | 0.802 | 0.398 |
| Short-Term Outcomes | | | | | | |
| Math standardized MCAS score | 0.154 | 0.948 | -0.483 | 0.973 | 0.320 | 0.867 |
| ELA standardized MCAS score (10th grade) | 0.169 | 0.893 | -0.468 | 1.006 | 0.336 | 0.779 |
| Plan to attend four-year college (10th grade) | 0.758 | 0.429 | 0.626 | 0.484 | 0.789 | 0.408 |
| Algebra (9th grade) | 0.536 | 0.499 | 0.668 | 0.471 | 0.505 | 0.500 |
| Geometry (10th grade) | 0.513 | 0.500 | 0.603 | 0.489 | 0.492 | 0.500 |
| Took MCAS on-time | 0.970 | 0.169 | 0.917 | 0.276 | 0.985 | 0.122 |
| Attendance rate (10th grade) | 0.943 | 0.069 | 0.911 | 0.094 | 0.952 | 0.057 |
| Ever suspended (10th grade) | 0.144 | 0.351 | 0.286 | 0.452 | 0.101 | 0.301 |
| Longer-Run Outcomes | | | | | | |
| Any college enrollment | 0.705 | 0.456 | 0.462 | 0.499 | 0.781 | 0.413 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Four-year college enrollment | 0.536 | 0.499 | 0.237 | 0.425 | 0.629 | 0.483 |
| Any college degree | 0.477 | 0.499 | 0.203 | 0.402 | 0.562 | 0.496 |
| Four-year college degree | 0.432 | 0.495 | 0.154 | 0.361 | 0.518 | 0.500 |
| Probability of earnings | 0.668 | 0.471 | 0.675 | 0.468 | 0.666 | 0.472 |
| Median earnings at age 30 (2024$) | 58732.13 | 56848.08 | 42446.92 | 42012.87 | 65293.62 | 58658.92 |
| N (Students) | 284,686 | | 67,734 | | 216,952 | |

*Note.* The sample includes 258 high schools. Number of observations varies by variable due to missingness. The analytic sample excludes students attending exam schools and career and technical education schools and students missing 8th grade covariates. College enrollment is measured within two years after expected high school graduation; college graduation is measured ten years after high school entry. Parent college completion and earnings are only available for the 2003 and 2004 cohorts. The inter-quartile range (not standard deviation) is presented for median earnings. ELA = English language arts; MCAS = Massachusetts Comprehensive Assessment System; SD = standard deviation.

**Table 2**

Standard deviations of school value-added estimates on longer-run outcomes.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Longer-Run Outcome | Overall SD | Overall (Scaled) | Low-income students SD | Non-low-income students SD |
| A. School Fixed Effects Estimates | | | | |
| Any college enrollment | 0.050 | 0.106 | 0.089 | 0.051 |
| Four-year college degree | 0.067 | 0.138 | 0.067 | 0.072 |
| Earnings at age 30 | 0.140 | 0.149 | 0.214 | 0.146 |
| B. Random Effects Estimates | | | | |
| Any college enrollment | 0.045 | 0.137 | 0.066 | 0.041 |
| Four-year college degree | 0.064 | 0.092 | 0.045 | 0.067 |
| Earnings at age 30 | 0.111 | 0.117 | 0.124 | 0.107 |

*Note.* College outcomes are in percentage points and earnings are logged 2024 dollars. Fixed effects estimates are from a regression model with school fixed effects, cohort fixed effects, and standard and expanded controls. Random effects estimates are the shrunken model-based standard deviation of the school effects with our standard and expanded controls. In column 2, we scale the fixed effects results by the sample standard deviations of the outcomes for low- and higher-income students. Estimates for low- and higher-income students come from separate models and are restricted to cells with at least 10 students. SD = standard deviation.

**Table 3**

Disattenuated and raw correlations of schools' effects across value-added with estimated reliabilities on diagonal.

| | Earnings at 30 | Four-year college degree | Any college enrollment | Test-score index | College plans | Academic on-track index | Attendance index |
|---|---|---|---|---|---|---|---|
| Earnings at 30 | 0.523 | 0.693 | 0.499 | 0.288 | 0.840 | 0.208 | 0.286 |
| Four-year college degree | 0.454 | 0.822 | 0.875 | 0.583 | 0.693 | 0.084 | 0.345 |
| Any college enrollment | 0.315 | 0.693 | 0.763 | 0.527 | 0.675 | 0.149 | 0.289 |
| Test-score index | 0.192 | 0.487 | 0.424 | 0.848 | 0.322 | 0.147 | 0.225 |
| College plans | 0.504 | 0.522 | 0.489 | 0.246 | 0.688 | 0.234 | 0.299 |
| Academic on-track index | 0.141 | 0.071 | 0.122 | 0.127 | 0.182 | 0.883 | 0.013 |
| Attendance index | 0.187 | 0.283 | 0.229 | 0.187 | 0.224 | 0.011 | 0.818 |

*Note.* The sample includes 258 high schools. Value-added estimates represent school average estimates (estimated across all years). Disattenuated correlations are presented above the diagonal and raw correlations below. Estimated reliabilities are in gray cells on the diagonal and are calculated using a split-sample approach. Disattenuated correlations are adjusted for measurement error using the estimated reliabilities.

**Table 4**

Regression of four-year college graduation on short-term value-added.

| | (1) 4-Yr College Grad | (2) 4-Yr College Grad | (3) 4-Yr College Grad | (4) 4-Yr College Grad | (5) 4-Yr College Grad | (6) 4-Yr College Grad | (7) 4-Yr College Grad |
|---|---|---|---|---|---|---|---|
| Test-score VA | 0.029*** | | | | 0.018*** | 0.018*** | 0.010*** |
| | (0.003) | | | | (0.003) | (0.003) | (0.003) |
| College plans VA | | 0.036*** | | | 0.029*** | 0.030*** | 0.019*** |
| | | (0.003) | | | (0.003) | (0.003) | (0.003) |
| Academic on-track index VA | | | 0.004 | | 0.000 | 0.000 | -0.000 |
| | | | (0.004) | | (0.002) | (0.002) | (0.002) |
| Attendance index VA | | | | 0.016*** | 0.010*** | 0.011*** | 0.009*** |
| | | | | (0.003) | (0.002) | (0.002) | (0.002) |
| Test-score VA x College plans VA | | | | | | 0.006** | |
| | | | | | | (0.002) | |
| Test-score VA x Academic on-track VA | | | | | | -0.004** | |
| | | | | | | (0.002) | |
| Test-score VA x Attendance VA | | | | | | 0.002 | |
| | | | | | | (0.002) | |
| Any college enrollment VA | | | | | | | 0.027*** |
| | | | | | | | (0.003) |
| Four-Year College Graduation Mean | 0.477 | 0.477 | 0.477 | 0.477 | 0.477 | 0.477 | 0.477 |
| Observations | 284,686 | 284,686 | 284,686 | 284,686 | 284,686 | 284,686 | 284,686 |
| R-squared | 0.315 | 0.317 | 0.312 | 0.313 | 0.318 | 0.318 | 0.320 |

*Note*. Regression of probability of four-year college graduation on leave-year out standardized school value-added estimates Models include standard and expanded controls as described in the text and cohort fixed effects. Robust standard errors in parentheses. Grad = graduation; VA = value added; Yr = year.
*p<0.1; **p<0.05; ***p<0.01.

**Table 5**

Regression of earnings on short-term value-added and college graduation value-added, 2003-2004 cohorts.

| | (1) Earnings | (2) Earnings | (3) Earnings | (4) Earnings | (5) Earnings | (6) Earnings | (7) Earnings |
|---|---|---|---|---|---|---|---|
| Test-score VA | 0.031*** | | | | 0.015** | 0.014** | 0.011 |
| | (0.007) | | | | (0.007) | (0.007) | (0.008) |
| College plans VA | | 0.046*** | | | 0.039*** | 0.042*** | 0.033*** |
| | | (0.007) | | | (0.007) | (0.007) | (0.007) |
| Academic on-track index VA | | | 0.008 | | 0.002 | 0.001 | 0.003 |
| | | | (0.007) | | (0.006) | (0.006) | (0.006) |
| Attendance index VA | | | | 0.020*** | 0.014*** | 0.013** | 0.011* |
| | | | | (0.007) | (0.005) | (0.005) | (0.006) |
| Test-score VA x College plans VA | | | | | | 0.009 | |
| | | | | | | (0.006) | |
| Test-score VA x Academic on-track VA | | | | | | 0.010* | |
| | | | | | | (0.006) | |
| Test-score VA x Attendance VA | | | | | | -0.001 | |
| | | | | | | (0.004) | |
| Any college graduation VA | | | | | | | 0.017* |
| | | | | | | | (0.009) |
| Average earnings (logged) | 10.814 | 10.814 | 10.814 | 10.814 | 10.814 | 10.814 | 10.814 |
| Observations | 75,516 | 75,516 | 75,516 | 75,516 | 75,516 | 75,516 | 75,516 |
| R-squared | 0.139 | 0.140 | 0.138 | 0.138 | 0.140 | 0.140 | 0.140 |

*Note*. Regression of logged earnings on leave-year out standardized school value-added estimates. Models include standard and expanded controls as described in the text. Robust standard errors in parentheses. VA = value added.
*p<0.1; **p<0.05; ***p<0.01.

**Table 6**

Regression of longer-run outcomes on peer index and short-term value-added.

| | Any College Enrollment | | Four-Year College Degree | | Earnings | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Peer index | 0.029*** | 0.022*** | 0.050*** | 0.038*** | 0.037*** | 0.019** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.007) | (0.008) |
| Test score VA | | 0.007*** | | 0.007** | | 0.010 |
| | | (0.002) | | (0.003) | | (0.008) |
| College plans VA | | 0.015*** | | 0.027*** | | 0.039*** |
| | | (0.002) | | (0.002) | | (0.007) |
| Academic on-track index VA | | 0.001 | | 0.001 | | 0.003 |
| | | (0.002) | | (0.002) | | (0.006) |
| Attendance index VA | | -0.002 | | 0.005*** | | 0.012** |
| | | (0.002) | | (0.002) | | (0.006) |

*Note*. Models include standard and expanded controls as described in the text and cohort fixed effects. Standard errors in parentheses. VA = value added.

*p<0.1; **p<0.05; ***p<0.01.

**Table 7**

Correlations of school value-added across low- and not-low-income students in the same high school.

| | Raw correlation | Disattenuated correlation |
|---|---|---|
| Earnings | 0.385 | 0.993 |
| Four-year college degree | 0.493 | 0.868 |
| Any college degree | 0.472 | 0.897 |
| Four-year college enrollment | 0.634 | 0.889 |
| Any college enrollment | 0.592 | 0.942 |
| Test-score index | 0.834 | 1.000 |
| College plans | 0.477 | 1.000 |
| Academic on-track index | 0.741 | 0.913 |
| Attendance index | 0.763 | 1.000 |

*Note.* School effects are estimated separately for low- and non-low-income students. Earnings data are only available for the 2003 and 2004 cohorts. Reported results are restricted to school-income cells with at least 10 students. Models include standard and expanded controls and cohort fixed effect. Disattenuated correlations are disattenuated for measurement error using reliabilities calculated from a split-sample approach as described in the text and capped at 1.

**Table 8**
Validity tests of student sorting.

| | Earnings | Four-year college degree | Any college enrollment | Test-score index | College plans | Academic on-track index | Attendance index |
|---|---|---|---|---|---|---|---|
| Correlation of VA with and without parent education | 1.00 | 0.99 | 0.99 | 1.00 | 0.98 | 1.00 | 1.00 |
| Specification test | 0.986 | 0.992 | 1.013 | 1.021 | 0.979 | 1.003 | 0.990 |
| | (0.068) | (0.008) | (0.015) | (0.014) | (0.038) | (0.013) | (0.031) |
| FB: Leave out score | 0.000 | -0.001** | -0.003*** | -0.015*** | 0.000 | -0.000 | 0.000 |
| | (0.000) | (0.001) | (0.001) | (0.002) | (0.001) | (0.000) | (0.000) |
| FB: Parent education | 0.001 | -0.003*** | -0.002 | -0.001 | -0.008** | -0.000 | -0.000 |
| | (0.001) | (0.001) | (0.002) | (0.001) | (0.004) | (0.000) | (0.000) |

*Note*. The correlation test compares the correlation between value-added models that do and do not control for parent education. Parent education is only available for the 2003 and 2004 cohorts. Specification test coefficients are tested under the null hypothesis that the coefficient equals 1. Leave-out score is grade-7 English language arts and includes 2004-2007 cohorts. FB = forecast bias; VA = value-added.
*p<0.1; **p<0.05; ***p<0.01.

**Table 9**

Correlations of school value-added estimates within outcomes across cohorts.

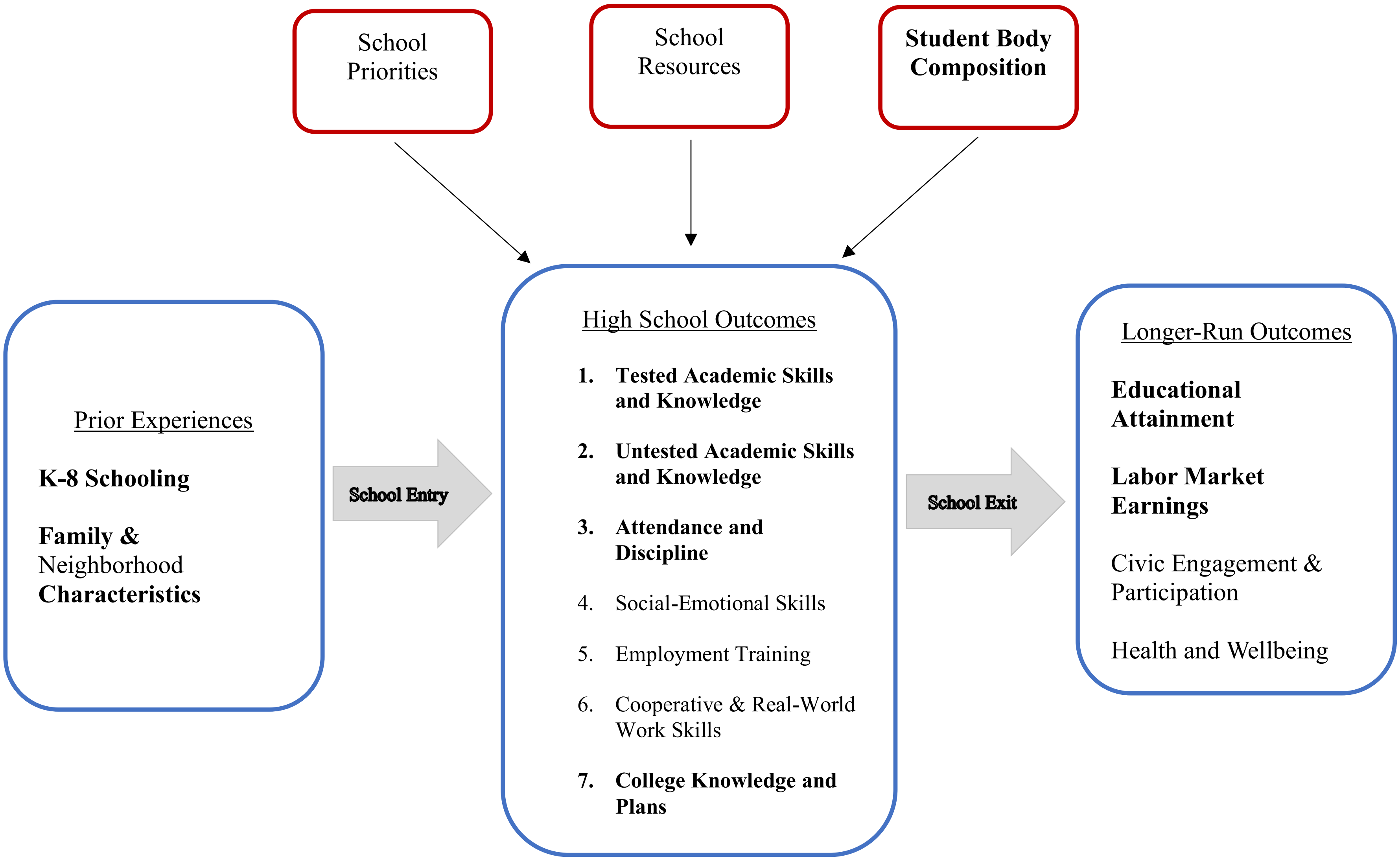
| | Annual earnings | Four-year college degree | Any college enroll-ment | Test-score index | College plans | Academic on-track index | Attendance index |
|---|---|---|---|---|---|---|---|
| t+1 | 0.543 | 0.760 | 0.633 | 0.732 | 0.368 | 0.763 | 0.688 |
| t+2 | | 0.743 | 0.611 | 0.661 | 0.254 | 0.651 | 0.617 |
| t+3 | | 0.744 | 0.551 | 0.596 | 0.357 | 0.577 | 0.599 |
| t+4 | | 0.718 | 0.515 | 0.535 | 0.179 | 0.571 | 0.599 |

*Note.* Correlation of school-by-year value-added models in year $t$ with years $t + i$. Models are estimated using standard and expanded controls.

**Figures**

**Figure 1**

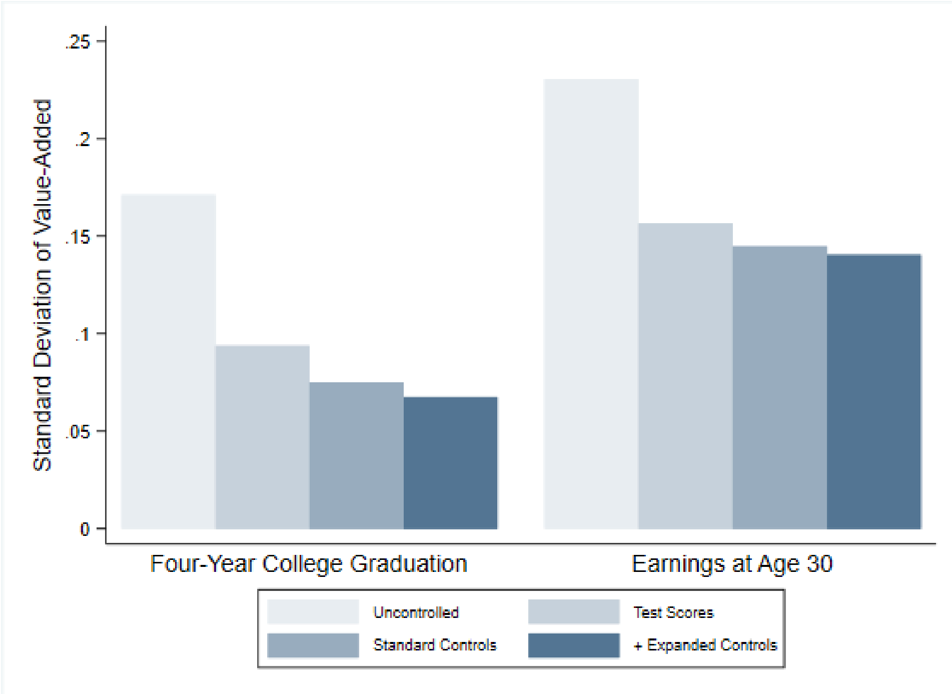
Multi-Dimensional Framework of School Effectiveness and Longer-Run Impacts.




*Note.* Bolded items are included in our analyses.

**Figure 2**

Estimated School Effects from Alternative Value-Added Models.

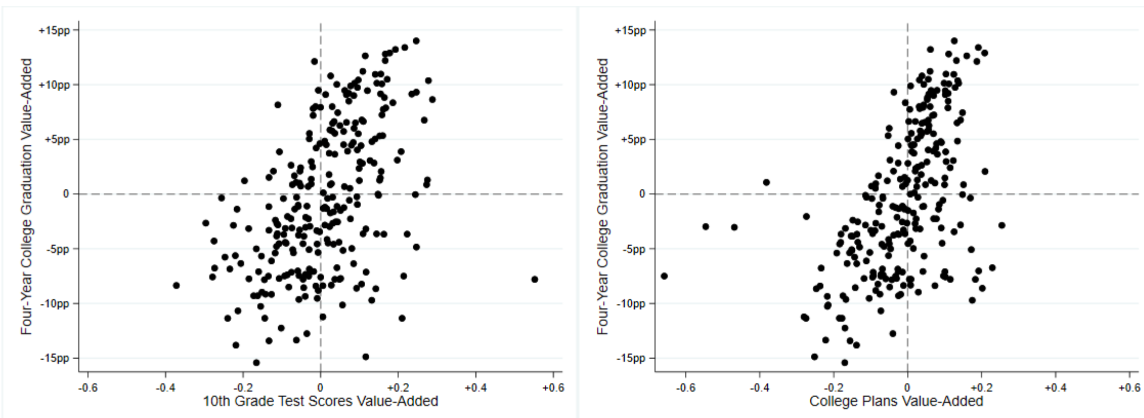



*Note.* The unconditional model includes a cohort fixed effect. Controls for other models are described in the text.

**Figure 3**

Relationship between Schools' Estimated Effects on Longer-Run Outcomes, Four-Year College Graduation (Panel A) and Earnings (Panel B), and Short-Run Outcomes, Test Score Value-Added (Left) and College Plans (Right).

Panel A: Four-Year College Graduation




Panel B: Earnings

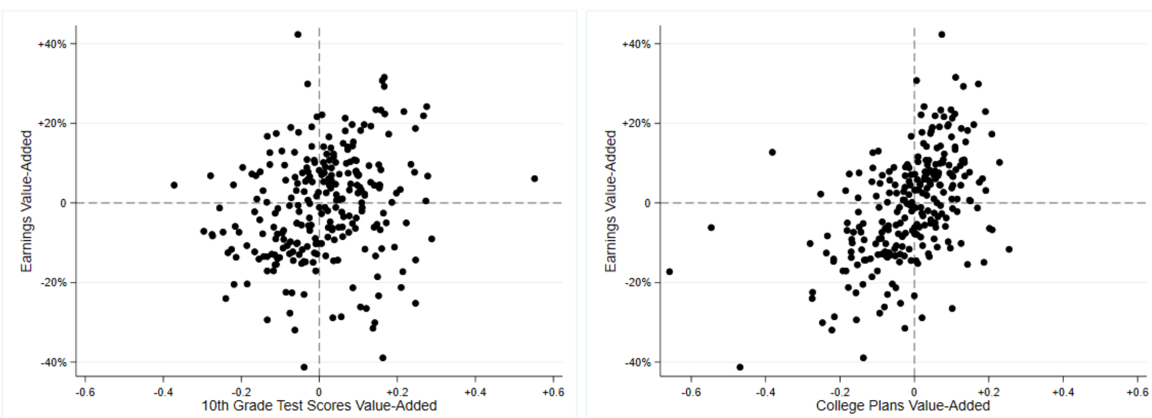


**Figure 4**

Relationship Between School Effects on Earnings and Four-Year College Graduation.

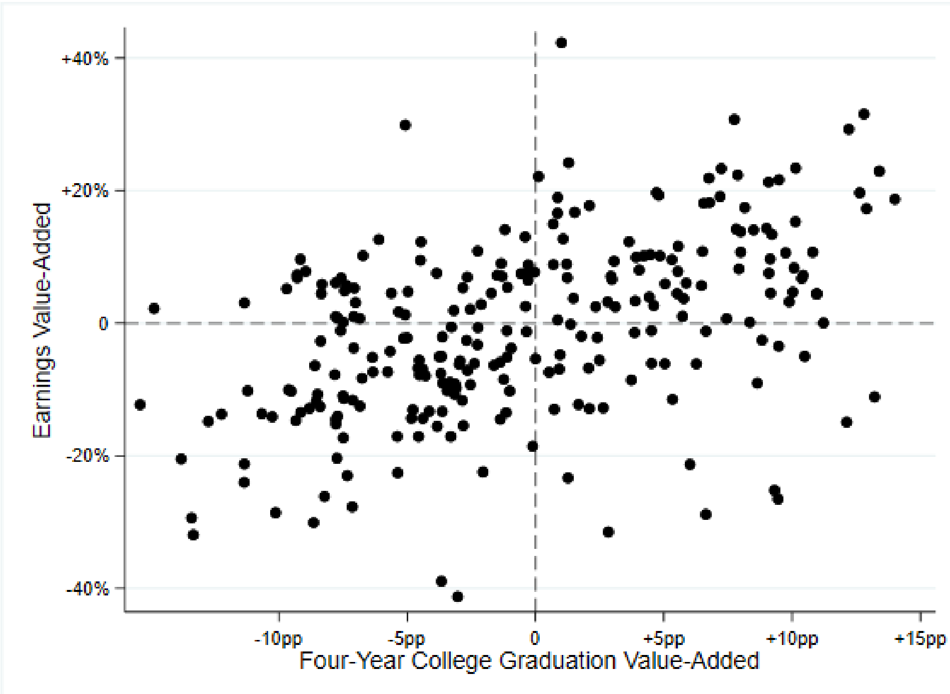


**Figure 5**

Relationship Between School Value-Added for Low-Income Students on Four-Year College Graduation (Left) and Earnings (Right) by School Percentage of Low-Income Students.