Examining the Distributional Effects of School Quality on Student Outcomes

Colin Lilly & Laura Meinzen-Dick

Abstract

Value-added modeling is a popular approach for evaluating school and teacher effectiveness. Previous research focuses primarily on average treatment effects, finding a positive relationship between education quality and student outcomes. We estimate quantile treatment effects of school quality and find heterogeneity in the impact on long-run student outcomes, with worse off students disproportionately benefiting from gains to school quality. These effects differ by sex.

1 Introduction

Quality secondary education is critical to equip young people with the skills and knowledge they will need for a successful working life. Going beyond retaining facts, quality secondary education builds critical-thinking and communication skills that allow students to succeed in higher education, adapt to new circumstances, and advance in the workforce. There is a wealth of evidence that teacher, and more broadly, school quality matters for student achievement, both in educational settings and in long-term career trajectories (Hanushek 1997; Rivkin, Hanushek, and Kain 2005). Measuring school quality, however, is difficult, given the inherently multi-dimensional nature of students, teaching, and learning. Policymakers seek objective, quantifiable metrics of school effectiveness with which they can evaluate programs and identify tools to boost student success.

At the same time, there is an increasing recognition that higher-quality schools may not benefit all students equally. Disadvantaged and minority students are not only more likely to attend worse-quality schools, but in those schools they disproportionally bear the consequences of improper preparation for the workforce (Jimenez 2020). Scholars have explored heterogeneity in the returns to education by gender, race, parental income, and student ability (Barrow and Rouse 2005; Tobias and Koop 2004; Coady and Dizioli 2018; Henderson, Polachek, and Wang 2011; Harmon, Oosterbeek, and Walker 2003). This also matters for how we define quality education: creating superstar students may require different strategies than catering to their struggling peers.

One increasingly common measure of teacher quality is the value-added (VA) approach, which observes a teacher's impact through changes in students' test scores. Value-added measures (VAMs) have been touted due to their use of objective data, standardized test scores, and because they appear to capture variance in teacher quality unaccounted for previously (Hanushek and Rivkin 2010). As policymakers seek to narrow the student achievement gap, VAMs have been proposed to evaluate individual teachers' performance and better hold them accountable. VAMs are not just useful for calculating the effectiveness of individual teachers but can be aggregated across classrooms to the school-level, including for evaluation purposes (Fuhrman 1999). As Ladd and Walsh (2002) note, the proliferation of district and school-specific report cards to parents that include information on student outcomes and test performance has generated public pressure for school improvement, including financial incentives or sanctions in some districts (Ladd and Walsh 2002).

In this paper, we seek to explore the effects of quality education on students' success in the long term. We follow research from Chetty, Friedman, and Rockoff (2014b) that demonstrates that VA measures are not only unbiased (using a quasi-experimental design), but also have long-term impacts on college attendance, salaries, and teenage pregnancy. However, the effects of school quality may differ from teacher VA, compounding over multiple years as well as encompassing other resources, including mentoring by other adults such as coaches, available in the school. Furthermore, it is possible that teacher and school quality may benefit some students more than others, as opposed to an average effect spread evenly across the achievement distribution. It could be that while high-performers that engage fully with quality schools benefit in the long term from the VA, lower-performing students are left behind or enter occupations that do not reward quality education. Alternately, it could be that higher-achieving students perform well in the long term no matter the quality of their schooling, so VA disproportionately impacts the left of the distribution.

This paper will address the question of heterogeneity in the effects of school quality with a cross-sectional research design, using data published from the National Center for Educational Statistics' (NCES) Education Longitudinal Study of 2002 (ELS:2002). The ELS:2002 tracks the transition of a national sample of high school sophomores—first through their senior year, and eventually into adulthood. Using standardized test scores in mathematics and relevant covariates, we calculate a measure of school VA. Calculating a VAM requires estimating students' future test scores based on a baseline score, and then assessing the difference between students' actual and predicted scores, holding other relevant factors fixed which impact achievement by introducing demographic controls. We then explore the relationship between student outcomes and school quality by regressing long-term outcome data available in the ELS on school VA estimates. We also estimate quantile treatment effects (QTEs) to examine the distributional effects of VA across the long-term outcome distribution.

Using the ELS:2002 data, we find that school VA is significantly related to a suite of long-term outcomes for students. Students who attend better high schools are more likely to attend postsecondary school, attend more selective institutions, earn scholarships, and be employed as adults. In addition to these discrete outcomes, we also find that higher school VA is associated with substantially higher adult income and socioeconomic status (a measure which takes into account the prestige of an individual's occupation). These mean effects are broadly in line with those found elsewhere in the literature, such as those of Chetty, Friedman, and Rockoff (2014b).

When we turn to the distribution of these effects, we uncover important heterogeneity masked by standard OLS regression. In particular, the income gains from school VA are larger in the left-hand side of the distribution: for every one standard deviation improvement in school VA, individuals in the 20th percentile of adult income have approximately a 18.8 percent gain in earnings, compared with 2.5 percent at the median. This is an important result: school quality is more important in predicting outcomes for lower-income students, suggesting that effective schools can play an equity-enhancing role. This pattern is particularly marked for female students; for male students, patterns of heterogeneity in treatment effects are noisier.

We make two primary contributions to an active literature about the value of education, and particularly the quality of education. First, and most substantially, we use quantile regression to explore the role of VA across the distribution. To our knowledge, we are the first to identify these distributionally heterogeneous treatment effects of secondary school quality on long-term outcomes, rather than mean effects and coarse heterogeneity on observables. Additionally, we focus on school value-added rather than the more common teacher VA, which is more policy-relevant for resource allocation and choices made by parents. Together, our results suggest that higher-achieving students perform well regardless of the quality of their institution, so VA matters more for low-performing students. This implies research on school quality that focuses on mean impacts may be understating the benefit that effective schools have on 'at-risk' students that much educational reform policy seeks to target. Policymakers seeking to improve equity of outcomes may find improving school quality to be an effective lever.

2 Literature Review

This literature review is organized as follows. Section 2.1 outlines the economic significance of teacher and school quality. Section 2.2 provides an overview of one key measure of teacher and school quality, value-added, and discusses its reported strengths and shortcomings. Next, 2.3 details literature documenting the relationship between value-added and student outcomes. Finally 2.4 describes quantile treatment effects and its applications on education research.

2.1 Quality of Education

As an investment into a future worker's human capital, education is highly correlated to earnings. In one seminal paper, Angrist and Krueger (1991) analyzed the returns to education by exploiting experimental conditions related to compulsory education laws. Using U.S. Census data, they found that men born in the first quarter of the calendar year that can drop out in a lower grade tend to not only drop out sooner but have decreased average earnings compared to their peers born later.

Research suggests that the quality of education is just as critical to positive economic outcomes as the years of schooling (Hanushek and Kimko 2000; Jamison, Jamison, and Hanushek 2007). Teachers appear to be a particularly important input to education quality. To quote Hanushek (2011): "Literally hundreds of research studies have focused on the importance of teachers for student achievement." (467) The belief that the quality of instruction influences student achievement comes mainly from the fact that average learning gains across classrooms are starkly different; as demonstrated by Hanushek (1992) in his study of students in Gary, Indiana, a good teacher versus a bad teacher can potentially be the difference between one full year of standardized achievement. Two students starting the year at similar levels of prior achievement may end up in drastically different places based on the quality of their teacher.

2.2 Overview of Value-added

One proposed alternative to teacher evaluation already being used in some states and districts is the incorporation of value-added models (VAMs).VAMs are statistical models which input several educational inputs to measure the effect of individual teachers on student achievement. Using VAMs, teacher quality has generally been shown to vary dramatically in ways unaccounted for with prior evaluation techniques (Hanushek and Rivkin 2010). Evidence of the stability of VAMs over time and across classrooms has also been demonstrated in several sources (Aaronson, Barrow, and Sander 2007; Goldhaber and Hansen 2013; Jackson 2014; Chetty, Friedman, and Rockoff 2014a; Koedel, Mihaly, and Rockoff 2015)

Despite this fact, discussions on VA are fraught with controversy, with many critiquing the validity of VAMs as measures of teacher quality. As outlined by Baker et al. (2010), there are several issues that arise when assessing VAMs' validity due to the insufficiency of statistical controls. Although VAMs attempt to account for the backgrounds of different students and the context of classrooms, there are issues related to the nonrandom sorting of teachers to students across schools and students to teachers within schools. Affluent schools may appear to have more effective teachers just because their student body has more access to support resources outside of school; this would influence not just students' starting places, as VAMs aim to account for, but also the rate at which they progress.

Questioning VAMs' claim of causality, Rothstein (2010) worked with data on students in North Carolina public elementary schools to test for bias using falsification tests. Exploiting the fact that future teachers cannot influence students' past performance, his linear regression model found that fifth-grade teacher assignments appear to have large effects on students' third-grade gains, indicating bias from classroom assignments conditional on the typical controls. In a similarly designed study, Bitler et al. (2021) apply a VA model to estimate teachers' effect on an outcome they surely cannot control: students' height. After using VA models that fail to control for classroom-by-teacher level error, Bitler et al. (2021) find a 0.65 inch impact on height per standard deviation increase in teacher VA. These non-zero effects raise doubt about the precision and magnitude of VAM estimates.

Analyzing the placement of students into classrooms in Arizona public and charter elementary schools, Paufler and Amrein-Beardsley (2014) found that both parents and teachers play a prodigious role in student placement in most schools sampled. These findings appear to be consistent with previous scholarship, including Aaronson, Barrow, and Sander (2007), Clotfelter, Ladd, and Vigdor (2006), Jackson (2014), Koedel and Betts (2011), and Rothstein (2009, 2010).

Much of contemporary literature on VAMs has attempted to respond to criticism. Kane et al. (2013) addressed bias in VAMs with their Measures of Effective Teaching (MET) project, a study of teacher effectiveness from six districts around the country. Their analysis concluded that measures of effectiveness from the prior period correctly predicted student achievement growth based on teacher assignment, with the magnitude of achievement gains consistent with expectations. Despite these results, there are several caveats raised by researchers. For one, sample sizes and compliance constraints are such that findings are noisy–while the 2013 study did greatly expand the sampling size and thus increase statistical precision, bias at the 95 percent confidence interval cannot be ruled out (Rothstein and Mathis 2013). Additionally, these experiments were performed both inside schools that had consented to be part of the study and wherein teachers were amicable to random assignment, calling into question the potential to externalize these findings to a more general model of teacher evaluation (Koedel, Mihaly, and Rockoff 2015).

To estimate the degree of bias in the VAM, Chetty, Friedman, and Rockoff (2014a) generated predicted test scores based on parental characteristics obtained from the tax data and regressed these scores on teacher VA. Controlling for prior test scores, Chetty, Friedman, and Rockoff (2014a) found that "forecast bias from omitting parent characteristics is at most 0.3 percent at the top of the 95 percent confidence interval. Using a similar approach, we find that forecast bias from omitting twice-lagged scores from the VA model is at most 2.6 percent". (295) Due to this minimal bias, they concluded that the influence of household and parental characteristics on student/teacher sorting not controlled for in typical VAMs is limited. Given there may be unobservable variables still contributing to forecast bias, Chetty, Friedman, and Rockoff (2014a) also employed a quasi-experimental design based on teacher turnover in fourth-grade classrooms. They found that simply controlling for a student's lagged test scores produces a forecast bias of five percent not statistically different from zero. However, omitting these lagged scores exhibits a forecast bias of over 40 percent. Clearly, then, much of the sorting of students to teachers relevant for predicting future achievement is captured using prior test scores.

2.3 Value-added and Student Outcomes

In addition to establishing the validity and usefulness of VAMs, recent scholarship has attempted to prove the impact of VA on important student outcomes later in life. Numerous studies on the impact of test performance on earnings for young workers—including Lazear (2003), Mulligan (1999), and Murnane et al. (2000)—[1] have found that one-SD increase in mathematics performance in high school results in 10 to 15 percent higher annual earnings. Looking earlier into a student's life, Chetty et

1. All of which as cited by Hanushek (2011).

al. (2010) used data from the Tennessee STAR experiment to examine how kindergarten test scores affect adult outcomes. When controlling for a rich set of parental characteristics in their analysis, they discovered that a one-SD increase in the kindergarten end-of-year scores is correlated with an 18 percent increase in earnings at age 27.

A highly influential study on teacher quality and outcomes comes from Chetty, Friedman, and Rockoff (2014b). Employing both a cross-sectional comparison across classrooms and a quasi-experimental design based on teacher turnover, they found that a one SD improvement in teacher VA in a single grade raises the probability of college attendance at age 20 by 0.82 percentage points and increases the annual earnings by 1.3 percent at age 28 (the oldest age for which data was available at the time of publication); assuming this earnings effect remains constant, just one SD increase in teacher VA in a single grade would cause a lifetime increase of approximately $39,000.

In addition to average total effects, Chetty, Friedman, and Rockoff (2014b) analyzed heterogeneity in the impact of teacher VA across several subgroups, including demographics, test subjects, and grades, using gains to college quality at age 20. They found that the long-term impacts of VA are slightly greater for women than men, that improvements in teacher quality in English are more valuable than improvements in mathematics, that impacts of VA appear constant in percentage terms by parents' income, and that gains appear large and significant through grades four through eight.

Demonstrating the efficacy of VA estimates in settings outside of the U.S., Kirkebøen (2022) observes the impact of school quality in Norway, as it relates to student achievement and to outcomes in adulthood. Using population-level administrative data on student backgrounds to construct school VAMs, he observes student outcomes into their 30s, finding a persistent and strong relationship between school quality and long-term outcomes.

It must be acknowledged that peer quality may also play a role in influencing student achievement, and may be associated with school quality. Recent research has supported the idea that peer quality is an important determinant of student success, identifying positive relationships between exposure to high quality peers and labor market outcomes (Billings and Hoekstra 2023; Humlum and Thorsager 2021). Conflicting findings include those from (Belfield and Rasul 2020), who demonstrate that marginal students' achievement is adversely impacted by increased proximity to high quality peers. Finally, findings from (Feng and Li 2016) indicate that negative peer effects are ameliorated or reversed when in the presence of high quality teachers, suggesting that the two have a reciprocal relationship with one another on student achievement.

2.4 Distributional Literature

Most regression analysis is oriented around the mean; that is, the standard regression models only the average value of the response variable for a given value of the covariate. However, conditional-mean models are not infallible and come with some limitations. First, OLS regression assumes that the relationship between predictor and response variables is the same across all values, making it ill-equipped to observe effects at noncentral locations on the distribution (Cook and Manning 2013). Usually, it is acceptable to center estimates on the mean as it is assumed the mean treatment is the same treatment for everyone— but there may be times where this assumption does not hold up. Second, the models' attention solely to the mean ignores how variables change the total shape of a distribution. In the context of teacher quality, it is easy to imagine that the impact of teacher quality may be heterogeneous across the distribution of student achievement; in other words, perhaps good teachers make more of a difference for high-performing students than average or low-performers.

To better address the potential distributional effects of teacher quality as mentioned above, we employ a quantile regression approach, as first outlined by Koenker and Bassett (1978). Quantiles can be used to specify any position on a distribution, including those beyond the central location. Originally, this method sought to estimate conditional quantile functions, in which quantiles of the conditional distribution of the response variable are expressed as functions of the observed predictor variables (Koenker and Hallock 2001).

In recent years, there has been increasing attention to QTEs in the economics literature, specifically as it surrounds education policy. Eide and Showalter (1998) used a conditional QR to estimate the distributional effects of school quality on standardized test score gains. Their findings indicate that certain inputs to education may affect student achievement at points in the distribution of test score gains unrecognized by observing mean treatment effects; for example, increasing the marginal expenditure per pupil in a district seems to raise test score gains at the bottom of the impact distribution without changing the mean, while increasing the length of the school year improved performance for high-achieving students with a null average effect. Arias, Hallock, and Sosa-Escudero (2001) used data on identical twins to observe if people with differing levels of ability obtain different returns to education, to determine whether there is a causal effect of education on earnings exogenous from ability. Using twins allows the authors to control for ability bias caused by family effects. Estimating QTEs, they interpreted heterogeneity in the estimated returns to education, by which individuals of higher baseline ability become more educated, in turn seeing higher returns to schooling. However, the authors note that this may be due to an upward ability bias at high quantiles due to the endogeneity of schooling choices, an observation that is likely missed by observing only average treatment effects.

With data on U.K. students, Arulampalam, et al. (n.d.) estimated a QR of student performance to monitor the effects of class absences on performance across the conditional performance distribution. The results indicated that absence does have adverse effects on performance across the conditional distribution. However, when controlling for unobserved characteristics, the effects of absence are more varied across the distribution. When accounting for characteristics like innate ability and work ethic, top-performing students demonstrate a causal impact of absence on performance; for low-performing students, this is not the case. In their analysis of the famous Tennessee STAR experiment, Jackson and Page (2013) used QTEs to find heterogeneity in the effects of class-size reduction policies across the achievement distribution: test score gains for treatment recipients were the greatest at the top of the achievement distribution, and smallest at the bottom.

Andrews, Li, and Lovenheim (2016) use UQR to examine heterogeneity in the earnings returns to university quality. Using data on male students who graduated from Texas high schools, they find significant differences in earnings premiums across the flagship Texas universities: UT Austin has an earnings premium that increases at the top of the earnings distribution, while Texas A&M has decreasing earnings premium across the distribution, from a peak of 36.4 percent at the 1st percentile to 17.6 percent at the 84th percentile. Community colleges have an overall negative effect on earnings, but heterogeneity exists and there are notable returns at the very top of the earning distribution. Also using unconditional QTEs, Penner (2016) studies potential differential effects of Teach for America (TFA) teachers on classroom performance across the distribution of student achievement. She finds that while TFA teachers outperform their non-TFA counterparts in mathematics education across the student achievement distribution, results in English were heterogeneous; while students at the top of the reading achievement distribution made greater gains with TFA teachers, students at the bottom of the distribution scored worse than those with traditional, veteran teachers.

Most recently, Shea and Jenkins (2021) use QTEs to observe the impact of curricula interventions targeting socio-emotional (SE) skills on preschool students' development across the outcome distribution. They find that while there are positive impacts of curricula interventions on emotional intelligence and problem-solving ability on average, gains are unequal across the outcome distribution–children at the upper end of emotional knowledge and problem-solving skills gained more from the curricula than others, seemingly supporting a "skills beget skills" hypothesis of treatment effectiveness.

3 Data

In this paper, we use data from the Educational Longitudinal Study of 2002 (ELS:2002), a nationally representative study of 10th graders during the spring term of the 2001-2002 school year. The random sampling of students was first stratified at the school level, with 750 public, Catholic, and non-Catholic private schools included. The base year design included cognitive assessments, along with a survey of students, parents, teachers, and school administrators. In the base year, 15,400 sophomores answered the questionnaire, out of 17,590 eligible selected students; additionally, 13,500 of the students' parents and 7,100 of their teachers participated as well (Ingels et al. 2004). Student surveys collected information about the student's background, school and employment experiences, and language background, among other characteristics. Parental surveys solicited data on home backgrounds, family structure, and the degree of academic support students receive, while teacher questionnaires collected information on the teacher's professional background and credentials (Bozick, Lauff, and Wirt 2007). Of the group of students who completed the questionnaire, 14,540 of them also completed the cognitive assessment, including standardized tests in mathematics and reading (Ingels, Burns, et al. 2005). A nonresponse bias assessment was performed to limit identified biases, and key data that was missing in the survey and for test performance variables was imputed (Ingels, Burns, et al. 2005).

Measures of interest available from the base-year include the following: sex, race/ethnicity, English as a native language, having an individualized education plan (IEP), granted test accommodations for the cognitive assessment, number of siblings, family composition, the highest level of education between both parent's (if applicable), combined parental income, the type of school attended (Public, Catholic, or Other Private), total school enrollment, the school-wide rate of 10th grade students eligible for reduced price or free lunch, the highest educational degree earned by the math teacher, the school's number of full time teachers, and the standardized test scores in mathematics and reading, estimates of achievement that are relative to the total population of 2002 sophomores. One cause for concern in the survey design and available data is the lack of prior test scores. As described in the literature review above, most VAMs include a rich set of lagged student test scores prior to the base year. While the bias from omitting twice-lagged prior test scores in Chetty, Friedman, and Rockoff (2014a) cross-sectional study was minimal, they do assert that the most robust of VAMs control for students' achievement in the preceding years. However, work from Angrist et al. (2021) and Kirkebøen (2022) indicates that adjustment for family background provides VA estimates that are suitable even for schools with an insufficient record of standardized test performance.

After 2002, the first follow-up occurs in the spring of 2004: of the 16,500 students invited, 15,000 participated (a weighted response rate of 91 percent), a majority of whom were seniors (Bozick, Lauff, and Wirt 2007).[2] Similar to the initial survey design, students in the follow-up cohort were surveyed and administered a cognitive assessment, this time only in mathematics. Transfer students who did not remain in their base year schools did not participate in testing, but had scores imputed for them. Test scores are re-standardized to have a national mean of 50 and standard deviation of ten (Ingels, Planty, et al. 2005). The standardized test scores in mathematics, taken from both the base year and the first follow-up, provide the foundation for our calculated school VA, which is covered in-depth later in this paper. As stated by Bozick, Lauff, and Wirt (2007), this reassessment of mathematics two years later can provide an estimate of achievement gains as related to school processes.

Table 1 provides descriptive statistics for the sample, as well as a comparison at baseline between "low" and "high" VA schools. For simplicity's sake, this distinction is defined by having a VA score greater than one. These two groups appear remarkably similar, with highly analogous student characteristics, household incomes, parental educational attainment, and family compositions. Slightly less than half of the sample are men and more than half are white and non-Hispanic. Approximately 80 percent of students in both groups are native English speakers (80 and 79 percent in low and high VA schools, respectively). Finally, schools in both groups contained a similarly small minority of IEP students, at approximately six percent of students each. This seems to support existing literature that supports school VA as an adequate capture of school quality alone.

2. Technically, the 2004 survey of seniors can be considered a second cohort. Although heavily overlapping with the original 2002 sample, the 2004 cohort is a "freshened" sample that includes a small sample of students who were not high school sophomores, were not sophomores in the U.S., were recovering from serious illness, were institutionalized, or had temporarily dropped out during the time of the base year study.

Variable	(1) Low VA Mean/(SE)	(2) High VA Mean/(SE)	(1)-(2) Pairwise t-test Mean difference
Math Score	50.501	50.949	-0.448
	(0.252)	(0.297)	
Male (=1 if male)	0.475	0.479	-0.005
	(0.008)	(0.009)	
White non-Hispanic (= if white)	0.550	0.543	0.007
	(0.016)	(0.019)	
English as First Language	0.804	0.788	0.017
	(0.011)	(0.013)	
IEP	0.062	0.057	0.006
	(0.004)	(0.004)	
BY Test Accommodations	0.007	0.001	0.006***
	(0.002)	(0.000)	
Family Income	62821.230	63296.578	-475.348
	(1330.715)	(1507.585)	
Two parents/guardians	0.739	0.731	0.008
	(0.007)	(0.008)	
Parent with bachelor's degree or higher	0.395	0.396	-0.000
	(0.011)	(0.013)	
Number of observations	8469	7423	15892
Number of clusters	405	346	751

Table 1 Student Descriptive Statistics—Standard errors clustered by school. ***p<0.01,**p<0.05,*p<0.1

Table 2 displays summary statistics for the schools sampled, once again separated into low or high value-added. Schools in both groups are predominantly public, at approximately 80 percent. Low VA schools are slightly less likely to be considered urban at 33 percent, compared to high VA schools at nearly 35 percent. Both groups have a total enrollment of close to 1000 students, with an average sophomore class size of around 310. Schools across both groups have about 20 percent of sophomores on free or reduced school lunch. Finally, both groups have between 26 and 28 teachers on average, with slightly over 40 percent of them holding advanced degrees.

Later follow-ups first occurred approximately two years after most of the sample had graduated high school, in 2006. Of the 15,900 eligible sample members, 14,200 participated, for a weighted response rate of 88 percent (Bozick, Lauff, and Wirt 2007). Students were asked questions about their high school graduation status, post-secondary education, and labor force participation (Bozick, Lauff, and Wirt 2007). From this survey, we observe whether the student has attended college, if they have attended a highly selective four-year university, and if they have ever been employed since leaving high school.

The most recent update occurred in 2012, a decade after the original study and eight years removed from when most of the participants were in high school. A total of 13,250 individuals participated, out of the 15,724 eligible (Lauff and Ingels 2014).

Variable	(1) Low VA Mean/(SE)	(2) High VA Mean/(SE)	(1)-(2) Pairwise t-test Mean difference
Public (=1 if public)	0.784 (0.021)	0.788 (0.023)	-0.005
Urbanicity (=1 if urban)	0.329 (0.024)	0.349 (0.027)	-0.021
Total School Enrollment-2001	984.803 (39.445)	952.129 (43.579)	32.675
Grade 10 Enrollment 2001-2002	312.513 (10.429)	309.777 (11.867)	2.736
Grade 10 Percentage on Free Lunc	20.763 (1.207)	22.220 (1.376)	-1.457
Number of Teachers	27.982 (1.010)	26.369 (1.129)	1.614
Math Teachers with Advanced Degree	0.409 (0.015)	0.421 (0.017)	-0.012
Number of observations	8469	7423	15892
Number of clusters	405	346	751

Table 2 School Descriptive Statistics—Standard errors clustered by school. ***p<0.01,**p<0.05,*p<0.1

Respondents were questioned about their high school completion, postsecondary education, experience in the labor market, and current activities (Lauff and Ingels 2014). Specifically, we incorporate 2011 employment income, socioeconomic status, whether the respondent has ever been employed since leaving high school, whether they are a single parent, and residential mobility into our model.[3] Additional information from external sources was collected in the subsequent year, including financial aid data from the U.S. Department of Education, postsecondary transcripts (to complement high school transcripts made available), and SAT/ACT scores. While such measures would ideally be included in our analysis, much of this data is restricted and unavailable in the publicly accessible codebook.

Along with a lack of access to potential outcomes of interest due to data restrictions, another limitation to the dataset comes from its coding of variables. Most of the variables, including those that are numerical, are coded as nominal categories or grouped within intervals. This makes performing regressions, and obtaining precise estimates, more difficult. While perhaps two of the most important outcomes, income in adulthood and socioeconomic status, are continuous, there are several other variables worth observing that cannot be due to how the survey was conducted.

3. Socioeconomic status is a composite measure that reflects the average of 2011 earnings from employment, the prestige score associated with the respondent's current or most recent job, and educational attainment. This measure does not account for the income, occupation, or education of the respondent's partner, which may factor into overall household status (Lauff and Ingels 2014). Residential mobility variables compare the distance (in miles) between the ZIP code centroid associated with the respondent's residence during the base year and the ZIP code centroid associated with their follow-up residence (Lauff and Ingels 2014).

4 Empirical Strategy

To estimate school quality, we must first obtain each school's VAM. Below, we present a simple framework for calculating school VA. Much of this approach is borrowed from previous research, including Kane, Rockoff, and Staiger (2008), Chetty, Friedman, and Rockoff (2014a), and Deming (2014). In order to interpret our findings causally, we need to rely on the identification assumption, in that unobserved causes of student outcomes in adulthood must be unrelated to school VA conditional on our observed factors. While this assumption is strong, it is corroborated with findings from Chetty, Friedman, and Rockoff (2014a), who employed a quasi-experimental design to prove the validity of VA. They used a difference-in-differences (DiD) design by exploiting changes in teacher assignments for students who move from one grade to another or from one school to another, finding that mean test scores across cohorts change sharply when a low VA teacher is replaced by a high VA, and vice-versa. The estimates from this quasi-experimental design matched the estimates from a second design which mirrored our cross-sectional design, as described above. Additional evidence from Deming (2014), Kirkebøen (2022) on school VA and long-term outcomes supports this through a cross-sectional and quasi-experimental approach; therefore, we tentatively interpret our estimates as accurately representing the relationship between school effectiveness and student outcomes.[4]

4.1 Calculating School Value-added

To generate an estimate of school quality for gains in student achievement, we first use the following regression:

$$A_{is} = \beta_1 X_{is} + \beta_2 W_{is} + \Gamma_s + u_{is} \tag{1}$$

Here, the dependent variable A_{is} denotes the mathematics assessment score of student i in school s. X_{is} represents a vector of student characteristics including base year test scores, demographics, IEP status, testing accommodations, parental education, income, family composition, and number of siblings. W_{is} represents the characteristics of the student's school, like its size, urbanicity, type (public, Catholic, or other private), number of teachers, and percentage of students on free or reduced lunch. The residual is assumed to be comprised of Γ_s, the value-added effect from entering school s that is constant over time, and u_{is}, an idiosyncratic student effect varied across students and over time (Kane and Staiger 2008).

Using the estimated coefficients obtained from this regression, we calculate the predicted assessment score $\hat{A}_{ist}$, which we subtract from the student's actual assessment score A_{is}

$$\hat{\nu}_{is} = A_{is} - \hat{A}_{is} \tag{2}$$

Here, $\hat{\nu}_{is}$ is the difference in expected and actual standardized achievement. While $\hat{A}_{is}$ captures the expected achievement of student i in school s given their relevant

4. A conservative, correlational, interpretation of these results is also informative, however.

individual, parental, and school characteristics, A_{is} captures the student's actual performance. Therefore, $\hat{\nu}_{is}$ captures the value-added of school s for student i, reflecting unobserved contributions of teachers, peers, and student-teacher match, as well as idiosyncratic student-specific unobservables (Kirkebøen 2022). A positive value here indicates a positive estimated VA. From this estimate, we estimate average school VA, Γ_s, by taking the average of student residuals sorted in a school s.

4.2 Mean treatment effects of school quality on student outcomes

Next, we examine the relationships between estimated school VA, Γ_s, and long-term outcomes with the following general equation:

$$Y_{ist} = \beta_1 X_{is} + \beta_2 W_{is} + \beta_3 \Gamma_s + u_{ist} \tag{3}$$

Here, Y_{ist} reflects each long-term outcome of a student i from school s at time t. Coefficients β_1 and β_2 measure the effects from student, family, classroom, and school controls previously described while β_3 measures the estimated effect of school VA. Building from the groundwork of previous empirical work we assume that persistent differences between VA reflects school quality, and not any unobserved differences in student group characteristics.

4.3 Quantile Treatment Effects

To estimate the distributional effects of school VA on student outcomes, we apply an unconditional quantile approach, as first proposed by Firpo, Fortin, and Lemieux (2009). Their method utilizes the influence function (IF) tool commonly appropriated in econometric models: "The IF (Y;q_τ,F_γ) of a distributional statistic $\upsilon(F_\gamma)$ represents the influence of an individual observation on that distributional statistic. Adding back the statistic $\upsilon(F_\gamma)$ to the influence function yields what we call the recentered influence function (RIF)." ((Firpo, Fortin, and Lemieux 2009), 954)

At the qth quantile, the RIF is represented by the following equation:[5]

$$RIF(y; y_q) = y_q + \frac{q - 1(Y \leq y_q)}{f_\gamma(y_q)} \tag{4}$$

Here, $f_\gamma(y_q)$ is the density of at y_q. Assuming a linear relationship between RIF and the explanatory variables X, this approach can be implemented with an ordinary least squares (OLS) regression. To this end, we use the Stata command `rifhdreg`, dependent on the estimate of the density $F_{(y_q)}$.

We also apply a more traditional conditional approach, as first described by Koenker and Bassett (1978). The conditional quantile function at quantile τ given a vector of regressors, x_i is defined as:

$$Q_\tau(Y_i|X_i) = F_Y^{-1}(\tau|X_i) \tag{5}$$

5. Equation taken from Cameron and Trivedi (2022)

This approach estimates the values of the coefficients that minimize the sum of the absolute deviations between the actual values of the dependent variable and the predicted values from the quantile regression function (Angrist and Pischke 2009). Chernozhukov, Fernández-Val, and Melly (2020) extend this original concept while incorporating less computationally intensive processes in their algorithm, which we use to produce estimates of conditional QTEs using the Stata command `qrprocess`.

Although the results for either approach are likely to be similar when the estimated mean effects are small, Angrist and Pischke (2009) advise caution in interpreting conditional quantiles, particularly by distinguishing effects on the individual versus on the distribution. Studying conditional QTEs means interpreting coefficients as conditional on the entire outcome distribution. For instance, if we were to hypothetically observe that a high quality school raises the lower deciles of the wage distribution for our cohort, we could not suggest that someone who would be poor without attending a high VA school is now less poor; instead, it must be interpreted that those who are currently in the lower deciles of wage earnings and attended high quality schools are less poor than the poor would be who did not attend one (Angrist and Pischke 2009). Essentially, we are unable to know which individuals benefit from treatment. Unconditional quantile regression estimates, on the other hand, are interpreted like traditional OLS coefficients. Despite this, Angrist and Pischke (2009) acknowledge that unconditional quantile regression is rife with theoretical uncertainty and is thus an ongoing approach being developed through active research. Nevertheless, we prioritize the unconditional quantile results in our discussion for their interpretability.

For each regression, we obtain estimates by every decile, from the 10th to the 90th. We include the same vector of student and school covariates that were found in mean regression estimates.

5 Results

Turning to our analysis and findings, we first consider mean treatment effects on educational outcomes. Next, we assess school quality's impact on employment and earnings. Then, we briefly discuss estimated effects on miscellaneous outcomes. Finally, we observe any heterogeneity in treatment effects on income and socioeconomic status using the QR techniques described previously. Our treatment variable is school VA, normalized to zero mean and unit standard deviation. All regressions include base student and school covariates. We also run regressions with and without the inclusion of family covariates for several outcomes.

5.1 Education Outcomes

Table 3 displays the estimated mean effects of school quality regressed on several college-related outcomes. The dependent variable displayed in Columns (1) and (2) is a binary variable that measures whether the individual had ever attended college. One standard deviation increase in school VA is associated with a 4.49 percentage point increase in the likelihood of attending college, or a 4.32 percentage point increase after the inclusion of family background covariates. The second outcome tested in this table is whether the individual had attended a "highly-selective" four-year postsecondary

VARIABLES	(1) Attended a Postsecondary Institution	(2) Attended a Postsecondary Institution	(3) Attended a Highly Selective Institution	(4) Received Scholarship in First Year
School value-added	0.0449***	0.0432***	0.0190***	0.0234***
	(0.00524)	(0.00471)	(0.00371)	(0.00437)
Constant	0.573***	0.531***	0.241***	0.345***
	(0.0381)	(0.0625)	(0.0569)	(0.0647)
Observations	15,892	15,892	15,892	15,892
R-squared	0.118	0.170	0.188	0.061
Baseline controls	Yes	Yes	Yes	Yes
Family Controls	No	Yes	Yes	Yes

Table 3 Mean Effects on Education Outcomes - OLS Results, with standard errors clustered by school. The dependent variable in Columns (1) and (2) is an indicator for ever attending a post-secondary institution. Column (1) removes the controls on family characteristics that includes number of siblings, family composition, parental education, and household income. Column (3) indicates attendance of a highly selective institution, referred to those whose first-year students' test scores places them in roughly the top fifth of baccalaureate institutions. Column (4) indicates a grant or scholarship offer in the respondent's first year of study. ***p<0.01,**p<0.05,*p<0.1

institution. "Highly-selective" institutions refer to those whose student body has test scores that places the university in roughly the top fifth of baccalaureate institutions (Ingels, Burns, et al. 2005). As seen in Column (3), a one standard deviation increase in school VA is correlated with a 1.90 percentage point increase in the probability of attending a "highly-selective" university. Finally, Column (4) presents the estimated effects for an additional variable that measures whether students received a scholarship during their first year at their first postsecondary institution. It shows a 2.34 percentage point increase in the likelihood of receiving a scholarship for a one standard deviation increase in VA.

These results indicate statistically significant, yet relatively small effects. School quality appears to have the largest effect on the probability of attending college, with a mean college attendance rate of approximately 66 percent. The inclusion of additional covariates has only a modest effect on point estimates. As evident in Table A1, the differences in rates of college attendance and attendance of highly selective colleges between the high VA and low VA are rather limited to begin with, and controlling for the range of necessary covariates to isolate school VA's impact only shrinks the estimated effect further.

5.2 Employment and Earnings Outcomes

Table 4 shows school quality impacts outcomes in the labor market, along a couple dimensions. First, Columns (1-3) display the relationship between school VA and the dependent variable, income from employment transformed by an inverse hyperbolic sine. An inverse hyperbolic sine transformation is used to approximate the natural

VARIABLES	(1) Adult Income	(2) Adult Income	(3) Adult Income	(4) SE Status	(5) Job - 2006	(6) Job - 2011
School value-	0.182***	0.174***	0.166***	0.0897***	0.0150***	0.0125***
	(0.0548)	(0.0532)	(0.0517)	(0.0219)	(0.00427)	(0.00379)
SVA^2		-0.00864	-0.00836			
		(0.0289)	(0.0279)			
Constant	5.874***	5.896***	4.965***	-1.427***	0.632***	0.672***
	(0.434)	(0.441)	(0.788)	(0.319)	(0.0615)	(0.0596)
Observations	15,892	15,892	15,892	15,892	15,892	15,892
R-squared	0.035	0.035	0.047	0.080	0.047	0.042
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Family controls	No	No	Yes	Yes	Yes	Yes

Table 4 Mean Effects on Employment and Earnings - OLS results, with standard errors clustered by school. The dependent variable in Columns (1-3) is 2011 income from employment transformed by inverse hyperbolic sine. Column (1-2) omits the controls on family characteristics that includes number of siblings, family composition, parental education, and household income. Columns (2) and (3) include a quadratic term for school VA, SVA^2. The dependent variable in Column (4) is a socioeconomic composite measure. This measure is standardized at a 0 mean and 1 SD. The outcome variable in Columns (5) and (6) is an indicator for whether the respondent had held a job since leaving high school—first recorded in the 2006 questionnaire, and again in 2011. It is worth noting that the third follow-up had more missing respondents whose answers were coded as 0 than the second, likely skewing results slightly. ***p<0.01,**p<0.05,*p<0.1

logarithm of income while retaining zero-valued observations, for which there are many (Bellemare and Wichman 2020).[6]

Income information was collected during the third follow-up, when most of the participants were around 26 years old. The regression seen in Column (1) reports a 18.2 percentage point increase in wages for one standard deviation of VA. Given that the relationship may not be perfectly linear, we incorporate a quadratic term for school VA in Column (2); this produces a new coefficient of 17.4 percentage points, but the quadratic is insignificant. Additionally, Column (3) shows the inclusion of family controls. Now, a one standard deviation increase in school VA is associated with a 16.6 percentage point increase in wages, statistically significant at the 95 percent confident level. School quality appears to exhibit a significant impact on future earnings.

The dependent variable in Column (4) is socioeconomic status as of 2011, a composite measured created by ELS that accounts for income, the prestige score associated with the individual's latest job, and educational attainment. The measure is standardized to have a zero mean and standard deviation of one prior to averaging. A one standard deviation increase in VA is correlated with a 0.0897 standard deviation increase in socioeconomic composite scores, including parental controls.[7]

Finally, Columns (5) and (6) display the impact of school quality on whether the respondent had held a job since leaving high school, as of the second follow-up for (5)

6. For a given variable x, the inverse hyperbolic sign can be calculated with the following equation: $arcsinh(x) = ln(x + \sqrt{x^2 + 1})$

7. It is worth noting that the composite measure does not account for the background of respondent's spouses or partner. Due to positive assortative matching, it is likely that joint SES is higher.

VARIABLES	(1) Single Parent	(2) Res. Mobility F3 vs. BY	(3) Res. Mobility F3 vs. F2	(4) Res. Mobility F2 vs. F1
School value-added	-0.00255*	-4.953	0.885	1.528
	(0.00145)	(3.962)	(3.940)	(1.818)
Constant	0.0170	-28.34	45.71	0.232
	(0.0221)	(63.98)	(57.77)	(20.94)
Observations	15,892	15,892	15,892	15,892
R-squared	0.030	0.051	0.047	0.012
Baseline controls	Yes	Yes	Yes	Yes
Family controls	Yes	Yes	Yes	Yes

Table 5 Mean Effects on Other Outcomes - OLS results, with standard erros clustered by school. Column (1) indicates respondents who are currently unmarried or not in a marriage-like relationship, with a biological or adopted child that lives with them in the household. Columns (2-4) indicate the distance in miles between the respondent's base year and F3 residence (2), F3 and F2 residences (3), and F2 and F1 residences (4), respectively. All regressions include controls on baseline and parental characteristics. ***p<0.01,**p<0.05,*p<0.1

and third for (6). Both have a coefficient of less than two percentage points, indicating a minimal effect of school quality on job attainment generally. Intuitively, this makes sense, as job attainment generally seems more contingent on education quantity than quality; that is, graduating high school (as almost everyone in this cohort did) is likely a better predictor of an individual's ability to get a job than the quality of the high school attended.

5.3 Other Outcomes

Table 5 displays the effects of school quality on other outcomes of interest. In Column (1), the dependent variable is whether the individual is a single parent, with a coefficient close to zero. Columns (2-4) report the effect on residential mobility, a variable which produces the distance in miles between the respondent's residences during two different time periods. Column (2) reports the difference between residences from the base year to third follow-up, (3) from the second follow-up to the third, and (4) from the first follow-up (the cohort's senior year) to the second.

Generally, these results have a statistically null estimated effect, indicating that school quality does not appear to be particularly predictive for either outcome. Given that only a very small portion of the sample are single parents (under three percent), it is difficult to obtain an unbiased estimator. For the residential mobility measures, it cannot be ignored that data for one-fourth of the sample is missing, with about another quarter having a zero-mile difference between their two residences. Because of this, it is equally hard to get unbiased estimations on the impact of school quality on individual's freedom to move.

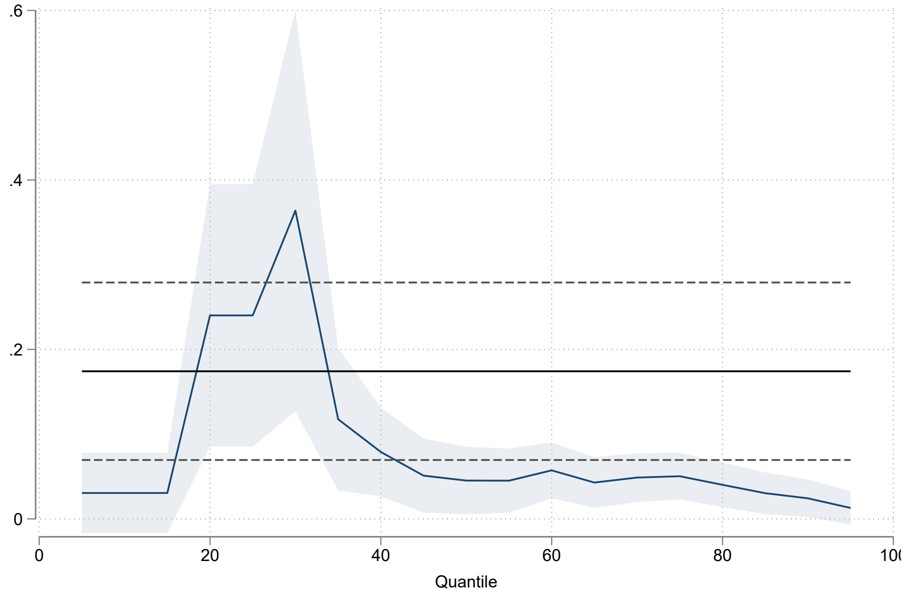

Fig. 1 Unconditional Quantile Treatment Effects on Adult Income

	(log)Adult Income								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
School value-added	0.0306 (0.0241)	0.240** (0.0789)	0.364** (0.121)	0.0790** (0.0267)	0.0453* (0.0203)	0.0573*** (0.0167)	0.0488*** (0.0147)	0.0403** (0.0134)	0.0243* (0.0112)
Observations	15892	15892	15892	15892	15892	15892	15892	15892	15892

Table 6 Unconditional Quantile Effects on Adult Income –Columns (1-9) represent each decile, 10th-90th. Standard errors clustered by school. The dependent variable is adult income, transformed by an inverse hyperbolic sine.***p<0.001,**p<0.01,*p<0.05

5.4 Quantile Effects

Now, we move on to a discussion of quantile treatment effects. Fig. 1 introduces the unconditional quantile treatment effect, as estimated using `rifhdreg`.[8] The dependent variable here is again the inverse hyperbolic sine of income. The dark blue line graphed shows the estimated effect of school VA for an individual at that quantile in the earnings distribution. The light blue area is the confidence interval at the 95 percent level. For comparison, the mean treatment effect is displayed as the solid horizontal line, with dashed lines as the 95 percent confidence intervals. From looking at this figure, it is evident that the effects of school quality are positive, or perhaps non-negative, throughout the outcome distribution, but are also heterogeneous.

While the confidence intervals are generous, there is a clear pattern: gains to school VA are substantially larger on the left of the distribution. Interestingly, the treatment

8. Current quantile regression methods are not capable of estimating treatment effects on a binary outcome, limiting the scope of our analysis to continuous variables.

effect of school quality on income seems most pronounced at the 2nd through 3rd decile, as displayed in Table 6. This may be the case because at the smallest quantiles, individuals have absolute incomes of zero and experience no gain from school VA to income as a result. Regardless, these results suggest that for those at the bottom of the income distribution, gains from school VA are slightly greater than for those at the top, who school quality may not benefit much. High-achieving people may thrive no matter what environment they are put in—whether due to innate ability or some other unobserved characteristic—and as such are highly likely to be successful regardless of school quality.

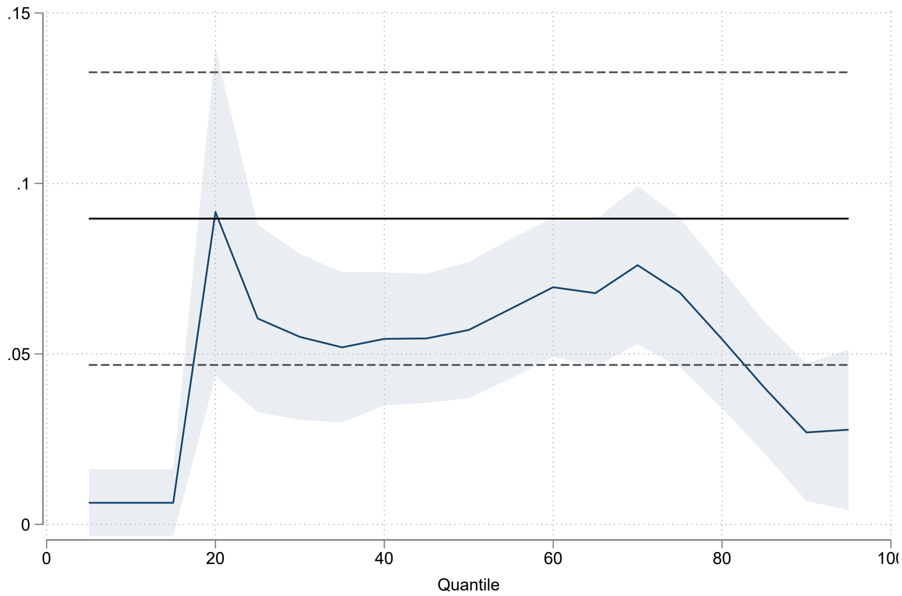

Fig. 2 Unconditional Quantile Treatment Effects on Socioeconomic Status

Fig. 2, which illustrates the effects on socioeconomic status as captured using `rifhdreg`, support the findings from Figure 1. This time, however, gains from school quality tapers off once near the middle of the distribution, rises between the 60th and 80th percentile, then declines again at high quantiles. The estimated effect is greatest at the 20th percentile. Those at the very bottom of the socioeconomic status distribution are impacted minimally by school VA gains; however like Figure 1, this is likely because the first decile of the distribution has an income of zero, a measure which factors into the socioeconomic composite score.

School quality may be especially ineffective for improving socioeconomic status at the top quintiles due to their non-school advantages. Individuals in the upper echelons of society were likely to be in a high social class regardless of the school they attended, unlike those near the other tail of the distribution who are more dependent on the resources and opportunities provided by a quality education. In fact, it appears that

	Socioeconomic Status								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
School value-added	0.00635 (0.00501)	0.0916*** (0.0245)	0.0550*** (0.0124)	0.0544*** (0.00994)	0.0570*** (0.0102)	0.0696*** (0.0104)	0.0760*** (0.0118)	0.0543*** (0.0103)	0.0270** (0.0103)
Observations	15892	15892	15892	15892	15892	15892	15892	15892	15892

Table 7 Unconditional Quantile Effects on Socioeconomic Status — Standard errors clustered by school. The dependent variable is a socioeconomic composite score. This measure is standardized at a 0 mean and a unit standard deviation.***p<0.001,**p<0.01,*p<0.05

across much of the distribution, the mean estimated effect may be overstating the actual impact of school quality on gains to socioeconomic status.

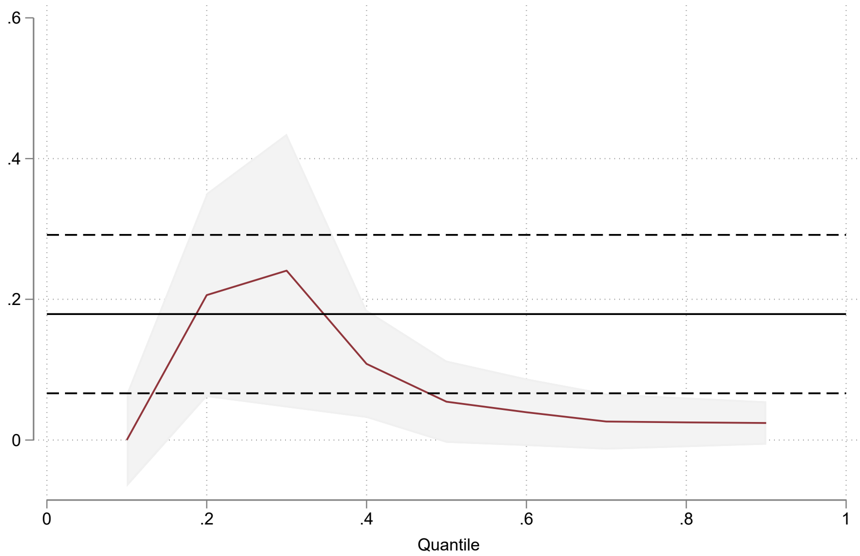

Fig. 3 Conditional Quantile Treatment Effects on Adult Income

Next, we estimate conditional QTEs using the `qrprocess` command. Fig. 3 displays an estimate for school quality's impact on the conditional income distribution. Here, the maroon line indicates estimated QTEs at each quantile with a 95 percent confidence interval (light grey area), while the solid and dashed black lines represent the mean treatment effect and its 95 percent confidence interval. The pattern shown here is generally similar to the unconditional distribution shown in Figure 1, but with a more gradual curve and less pronounced peak. Once again, the difference in effects between the low and high quantiles of the distribution appear to be different and statistically significant. The largest estimated effect from increased school VA appears at or around the 30th percentile on the conditional earnings distribution—therefore, among those who received an additional unit of school quality, the lower range of earners disproportionately benefits.

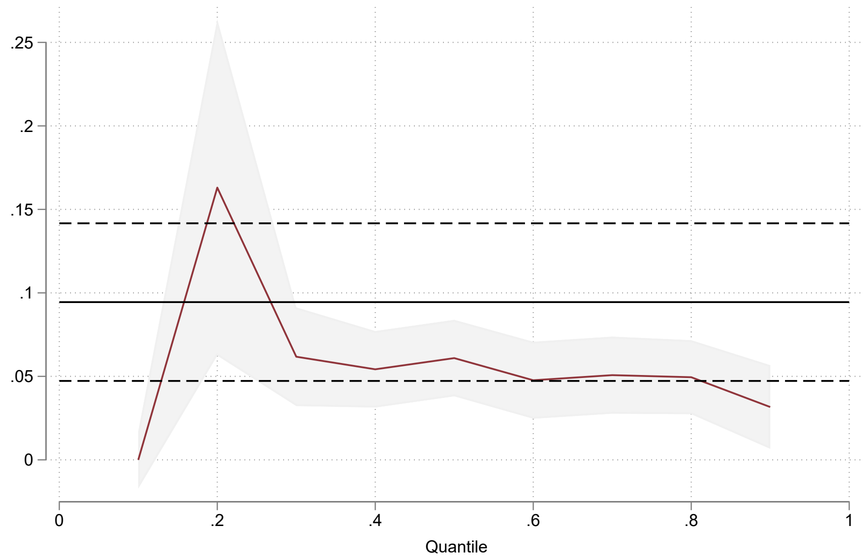

Fig. 4 Conditional Quantile Treatment Effects on Socioeconomic Status

Fig. 4 maps school VA's impact on the conditional socioeconomic composite distribution. Here, the shape is broadly similar to the one presented in Figure 3, but with a noticeably larger estimated effect at the 20th conditional percentile and a flatter slope at increasing quantiles. Those at around the 20th percentile, conditional on having an increased unit of school VA, demonstrate a gain to their socioeconomic composite score that is higher than the mean treatment effect's upper bound at the 95 percent confidence interval. Those beyond the 80th percentile, meanwhile, seem to receive a reduced (albeit still positive) impact of increased school quality on their socioeconomic composite score, corroborating the results of the unconditional QR. While overall it appears that the results from conditional quantile regression provide more conclusive evidence of heterogeneity in treatment effects, due to the difficulty in interpretation the unconditional quantile estimation is better suited to provide estimates of distributional gains from school VA not just conditional on attending a school with VA, but on the overall distribution.

Finally, we compare differences in quantile treatment effects between subgroups. Figures 5 and 6 plot estimated effects of VA on income between males and females. While the pattern between the two is once again similar, there is a drastic difference in the magnitude of the estimated quantile treatment effect between the sexes. Despite having a relatively similar, if not slightly lesser, mean treatment effect compared to women, men at the 20th through 30th percentiles seem to experience sizeable gains to income. Even with generous confidence intervals—dropping approximately half of the

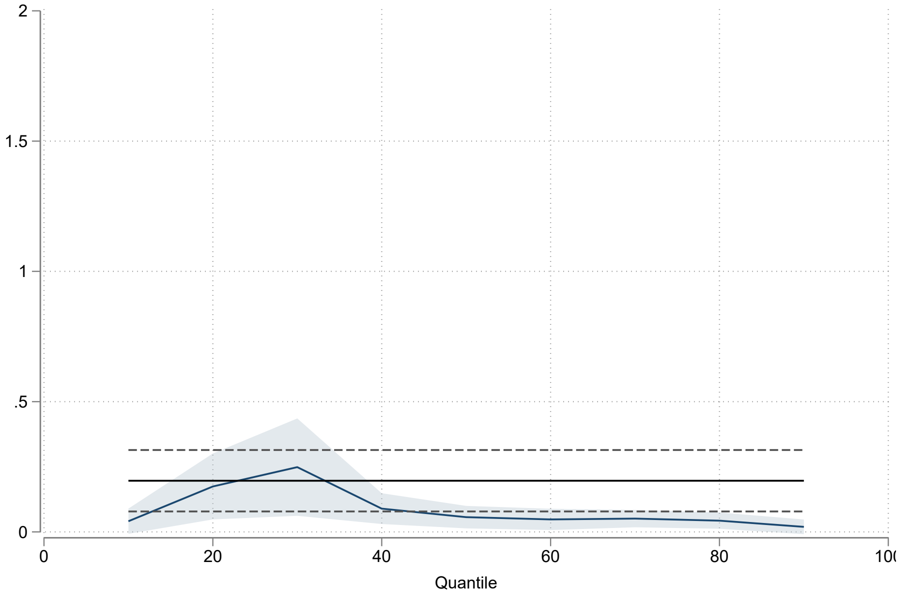

Fig. 5 Unconditional Quantile Treatment Effects on Adult Income for Females

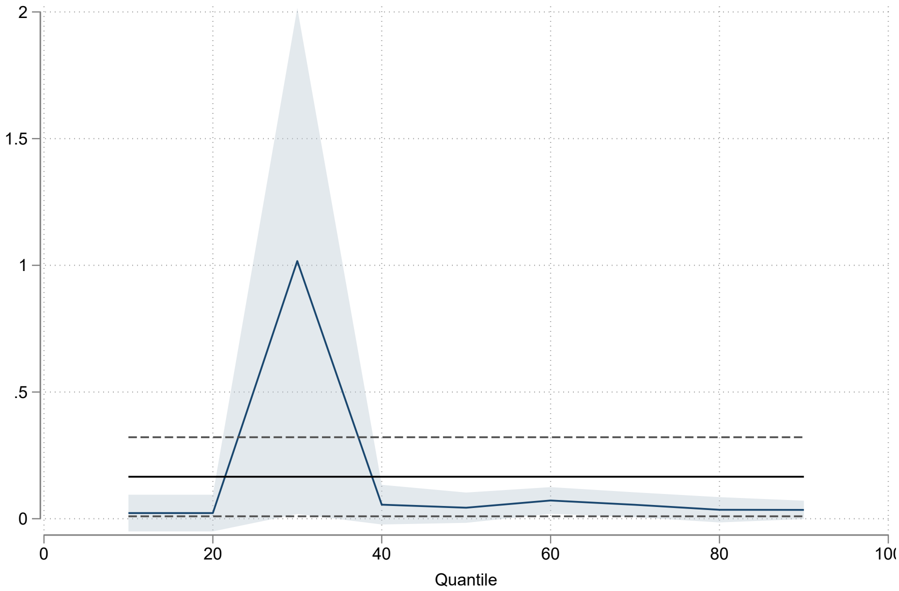

Fig. 6 Unconditional Quantile Treatment Effects on Adult Income for Males

sample size has increased variance— there appears to be an estimated gain to school quality for male earners around the 25th percentile that is more than double the effect to female earners at the same point in the distribution. While both groups experience the greatest estimated effect at around the 30th quantile, a one SD increase in school VA is associated with around a 100 percent increase in earnings for men, compared to around a 25 percent increase for women.

6 Conclusion

Most existing research on school effectiveness suggests that, on average, high VA schools positively impact student achievement and influence long-term success. This study moves beyond mean treatment effects and observes heterogeneity in the impact of school VA on key outcomes in adulthood, including earnings from employment and socioeconomic status. While individuals of all levels of success appear to benefit from attending high quality schools, school VA appears to be most beneficial for those close to the bottom of the outcome distribution. Furthermore, this heterogeneity differs along gendered lines: men see significantly higher spikes in estimated effects on income at around the 25th percentile. Our analysis indicates that subgroup analyses for mean effects may inadequately capture the relevant distributional effects of school quality, and that quantile regression approaches have their role in VA research.

There may be opportunity for further research. As mentioned by Chetty, Friedman, and Rockoff (2014b), adopting a similar quantile regression methodology to teacher VA could help reveal which teachers are effective at working with high or low achieving students. Given that some research suggests that teacher quality may be the most important individual input to achievement (Hanushek 2011), knowing teachers' distributional impacts on student outcomes may improve classroom assignments and bring up overall VA. It is also worth noting that the surveyed student cohort attended college and entered the workforce during a particularly turbulent time, the 2008 Great Recession. It is possible that during this time, broader macroeconomic conditions played a greater role in labor market sorting, dampening gains to school quality for the upper tail of earners. Some evidence suggests that this is not the case, with wage growth appearing most stagnant for the middle quantile of earners (Smeeding 2012). Still, our findings should be interpreted in the context of the recession, with additional research needed to determine the long-term impacts on outcomes beyond this period.

While these findings are encouraging, they present problematic implications for policy. To best serve disadvantaged students, one proposal is the increased promotion and distribution of school voucher programs, which offer families the opportunity to send their children to private or charter schools in the area. On one hand, this has the potential to match disadvantaged students to high quality teachers and grant them access to high-aptitude peers. These students stand to disproportionately benefit from attending better schools. On the other hand, voucher programs fundamentally benefit a select few, redistributing tax dollars away from already-struggling schools to help a limited number of voucher recipients. Therefore, underlying the debates on helping needy students is a deep tension between individual and collective well-being. Although this paper does not claim to solve the contradictions at the heart of education reform, it is ultimately encouraging that quality schools can make a positive impact in the lives of their worse-off students.

References

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." Publisher: The University of Chicago Press, *Journal of Labor Economics* 25, no. 1 (January 1, 2007): 95–135. ISSN: 0734-306X, accessed April 3, 2023. https://doi.org/10.1086/508733. https://doi.org/10.1086/508733.

Andrews, Rodney J., Jing Li, and Michael F. Lovenheim. 2016. "Quantile treatment effects of college quality on earnings." Publisher: University of Wisconsin Press, *Journal of Human Resources* 51 (1): 200–238.

Angrist, Joshua, Peter Hull, Parag A. Pathak, and Christopher Walters. 2021. "Credible School Value-Added with Undersubscribed School Lotteries." *The Review of Economics and Statistics* (December 20, 2021): 1–46. ISSN: 0034-6535, accessed March 2, 2023. https://doi.org/10.1162/rest_a_01149. https://doi.org/10.1162/rest_a_01149.

Angrist, Joshua D., and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" Publisher: Oxford University Press, *The Quarterly Journal of Economics* 106 (4): 979–1014. ISSN: 00335533, 15314650, accessed April 3, 2023. https://doi.org/10.2307/2937954. http://www.jstor.org/stable/2937954.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press. ISBN: 978-0-691-12034-8, accessed March 31, 2023. https://doi.org/10.2307/j.ctvcm4j72. https://www.jstor.org/stable/j.ctvcm4j72.

Arias, Omar, Kevin F. Hallock, and Walter Sosa-Escudero. 2001. "Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins data." *Empirical Economics* 26, no. 1 (March 1, 2001): 7–40. ISSN: 1435-8921, accessed March 30, 2023. https://doi.org/10.1007/s001810000053. https://doi.org/10.1007/s001810000053.

Arulampalam, et al. n.d. "Am I missing something? The effects of absence from class on student performance — Elsevier Enhanced Reader." Accessed November 3, 2022. https://doi.org/10.1016/j.econedurev.2011.12.002. https://reader.elsevier.com/reader/sd/pii/S0272775711001786?token=47DA4CB72133F11B6A2FA7675525C196313B8394D918A7942591C852D4400D7F23238B5D05E368D5AFBE29B34A1BF028&originRegion=us-east-1&originCreation=20221103170506.

Baker, Eva L., Paul E. Barton, Linda Darling-Hammond, Edward Haertel, Helen F. Ladd, Robert L. Linn, Diane Ravitch, Richard Rothstein, Richard J. Shavelson, and Lorrie A. Shepard. 2010. *Problems with the Use of Student Test Scores to Evaluate Teachers. EPI Briefing Paper #278.* Publication Title: Economic Policy Institute ERIC Number: ED516803. Economic Policy Institute, August 29, 2010. Accessed October 30, 2022. https://eric.ed.gov/?id=ED516803.

Barrow, Lisa, and Cecilia Elena Rouse. 2005. "Do Returns to Schooling Differ by Race and Ethnicity?" *American Economic Review* 95 (2): 83–87. https://doi.org/10.1257/000282805774670130. https://www.aeaweb.org/articles?id=10.1257/000282805774670130.

Belfield, Christopher, and Imran Rasul. 2020. "Cognitive and Non-Cognitive Impacts of High-Ability Peers in Early Years*." _Eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-5890.12216, *Fiscal Studies* 41 (1): 65–100. https://doi.org/https://doi.org/10.1111/1475-5890.12216. https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-5890.12216.

Bellemare, Marc F., and Casey J. Wichman. 2020. "Elasticities and the Inverse Hyperbolic Sine Transformation." _Eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/obes.12325, *Oxford Bulletin of Economics and Statistics* 82 (1): 50–61. ISSN: 1468-0084, accessed April 1, 2023. https://doi.org/10.1111/obes.12325. https://onlinelibrary.wiley.com/doi/abs/10.1111/obes.12325.

Billings, Stephen B., and Mark Hoekstra. 2023. "The Effect of School and Neighborhood Peers on Achievement, Misbehavior, and Adult Crime." _Eprint: https://doi.org/10.1086/720323, *Journal of Labor Economics* 41 (3): 643–685. https://doi.org/10.1086/720323. https://doi.org/10.1086/720323.

Bitler, Marianne, Sean P. Corcoran, Thurston Domina, and Emily K. Penner. 2021. "Teacher Effects on Student Achievement and Height: A Cautionary Tale." Publisher: Routledge, *Journal of Research on Educational Effectiveness* 14, no. 4 (October 2, 2021): 900–924. ISSN: 1934-5747, accessed April 4, 2023. https://doi.org/10.1080/19345747.2021.1917025. https://www.tandfonline.com/doi/citedby/10.1080/19345747.2021.1917025.

Bozick, Robert, Eric Lauff, and John Wirt. 2007. *Education Longitudinal Study of 2002 (ELS:2002): A First Look at the Initial Postsecondary Experiences of the High School Sophomore Class of 2002.* Publication Title: National Center for Education Statistics ERIC Number: ED498655. National Center for Education Statistics, October. Accessed March 2, 2023. https://eric.ed.gov/?id=ED498655.

Cameron, Adrian Colin, and Pravin K. Trivedi. 2022. *Microeconometrics using stata.* Second edition. Vol. 2. College Station, Texas: Stata Press. ISBN: 978-1-59718-359-8.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2020. *Fast Algorithms for the Quantile Regression Process,* arXiv:1909.05782, April 6, 2020. Accessed March 31, 2023. https://doi.org/10.48550/arXiv.1909.05782. arXiv: 1909.05782[econ,stat]. http://arxiv.org/abs/1909.05782.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2010. "$320,000 Kindergarten Teachers." Type: 10.1177/003172171009200306, *Phi Delta Kappan* 92 (3): 22–25. https://journals.sagepub.com/doi/abs/10.1177/003172171009200306.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104, no. 9 (September): 2593–2632. ISSN: 0002-8282, accessed October 31, 2022. https://doi.org/10.1257/aer.104.9.2593. https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2593.

———. 2014b. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104, no. 9 (September): 2633–2679. ISSN: 0002-8282, accessed October 28, 2022. https://doi.org/10.1257/aer.104.9.2633. https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2633.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System], *The Journal of Human Resources* 41 (4): 778–820. ISSN: 0022166X, accessed April 3, 2023. http://www.jstor.org/stable/40057291.

Coady, David, and Allan Dizioli. 2018. "Income inequality and education revisited: persistence, endogeneity and heterogeneity." ISBN: 0003-6846 Publisher: Taylor & Francis, *Applied Economics* 50 (25): 2747–2761.

Cook, Benjamin Lê, and Willard G. Manning. 2013. "Thinking beyond the mean: a practical guide for using quantile regression methods for health services research." Publisher: Shanghai Mental Health Center, *Shanghai Archives of Psychiatry* 25, no. 1 (February): 55. ISSN: 1002-0829, accessed November 20, 2022. https://doi.org/10.3969/j.issn.1002-0829.2013.01.011. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4054530/.

Deming, David J. 2014. "Using School Choice Lotteries to Test Measures of School Effectiveness." *American Economic Review* 104, no. 5 (May): 406–11. https://doi.org/10.1257/aer.104.5.406. https://www.aeaweb.org/articles?id=10.1257/aer.104.5.406.

Eide, Eric, and Mark H. Showalter. 1998. "The effect of school quality on student performance: A quantile regression approach." *Economics Letters* 58, no. 3 (March 1, 1998): 345–350. ISSN: 0165-1765, accessed November 20, 2022. https://doi.org/10.1016/S0165-1765(97)00286-3. https://www.sciencedirect.com/science/article/pii/S0165176597002863.

Feng, Han, and Jiayao Li. 2016. "Head teachers, peer effects, and student achievement." *China Economic Review* 41:268–283. ISSN: 1043-951X. https://doi.org/https://doi.org/10.1016/j.chieco.2016.10.009. https://www.sciencedirect.com/science/article/pii/S1043951X16301377.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. “Unconditional Quantile Regressions.” _Eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA6822, *Econometrica* 77 (3): 953–973. ISSN: 1468-0262, accessed November 20, 2022. https://doi.org/10.3982/ECTA6822. https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA6822.

Fuhrman, Susan. 1999. “The New Accountability.” *CPRE Policy Briefs* (January 1, 1999). https://repository.upenn.edu/cpre_policybriefs/14.

Goldhaber, Dan, and Michael Hansen. 2013. “Is it Just a Bad Class? Assessing the Long-term Stability of Estimated Teacher Performance.” _Eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecca.12002, *Economica* 80 (319): 589–612. ISSN: 1468-0335, accessed November 3, 2022. https://doi.org/10.1111/ecca.12002. https://onlinelibrary.wiley.com/doi/abs/10.1111/ecca.12002.

Hanushek, Eric A. 1992. “The Trade-off between Child Quantity and Quality.” *Journal of Political Economy* 100 (1): 84–117. ISSN: 00223808, 1537534X, accessed April 5, 2023. http://www.jstor.org/stable/2138807.

———. 1997. “Assessing the Effects of School Resources on Student Performance: An Update.” *Educational Evaluation and Policy Analysis* 19 (2): 141–164. ISSN: 01623737, 19351062, accessed April 5, 2023. http://www.jstor.org/stable/1164207.

———. 2011. “The economic value of higher teacher quality.” *Economics of Education Review* 30, no. 3 (June 1, 2011): 466–479. ISSN: 0272-7757, accessed October 30, 2022. https://doi.org/10.1016/j.econedurev.2010.12.006. https://www.sciencedirect.com/science/article/pii/S0272775710001718.

Hanushek, Eric A., and Dennis D. Kimko. 2000. “Schooling, Labor-Force Quality, and the Growth of Nations.” *American Economic Review* 90, no. 5 (December): 1184–1208. ISSN: 0002-8282, accessed November 6, 2022. https://doi.org/10.1257/aer.90.5.1184. https://www.aeaweb.org/articles?id=10.1257/aer.90.5.1184.

Hanushek, Eric A., and Steven G. Rivkin. 2010. “Generalizations about Using Value-Added Measures of Teacher Quality.” *American Economic Review* 100 (2): 267–71. https://doi.org/10.1257/aer.100.2.267. https://www.aeaweb.org/articles?id=10.1257/aer.100.2.267.

Harmon, Colm, Hessel Oosterbeek, and Ian Walker. 2003. “The Returns to Education: Microeconomics.” Publisher: John Wiley & Sons, Ltd, *Journal of Economic Surveys* 17, no. 2 (April 1, 2003): 115–156. ISSN: 0950-0804, accessed July 21, 2023. https://doi.org/10.1111/1467-6419.00191. https://doi.org/10.1111/1467-6419.00191.

Henderson, Daniel J., Solomon W. Polachek, and Le Wang. 2011. "Heterogeneity in schooling rates of return." *Economics of Education Review,* Special Issue: Economic Returns to Education, 30, no. 6 (December 1, 2011): 1202–1214. ISSN: 0272-7757, accessed July 21, 2023. https://doi.org/10.1016/j.econedurev.2011.05.002. https://www.sciencedirect.com/science/article/pii/S0272775711000720.

Humlum, Maria Knoth, and Mette Thorsager. 2021. "The Importance of Peer Quality for Completion of Higher Education." *Economics of Education Review* 83:102120. ISSN: 0272-7757. https://doi.org/https://doi.org/10.1016/j.econedurev.2021.102120. https://www.sciencedirect.com/science/article/pii/S027277572100039X.

Ingels, Steven J., Laura J. Burns, Stephanie Charleston, Chen Xianaglei, and Emily Forrest Cataldi. 2005. "A Profile of the American High School Sophomore in 2002: Initial Results From the Base Year of the Longitudinal Study of 2002." Publisher: National Center for Education Statistics, March 10, 2005. Accessed April 3, 2023. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005338.

Ingels, Steven J., Michael Planty, Robert Bozick, and Jeffrey A. Owings. 2005. *A Profile of the American High School Senior in 2004: A First Look. Initial Results from the First Follow-Up of the Education Longitudinal Study of 2002 (ELS:2002). NCES 2006-348.* Publication Title: National Center for Education Statistics ERIC Number: ED486298. ED Pubs, P, October. Accessed March 30, 2023. https://eric.ed.gov/?id=ED486298.

Ingels, Steven J., Daniel J. Pratt, James E. Rogers, Peter H. Siegel, and Ellen S. Stutts. 2004. *Education Longitudinal Study of 2002: Base Year Data File User's Manual. NCES 2004-405.* Publication Title: National Center for Education Statistics ERIC Number: ED484410. ED Pubs, P, February. Accessed April 3, 2023. https://eric.ed.gov/?id=ED484410.

Jackson, C. Kirabo. 2014. "Teacher Quality at the High School Level: The Importance of Accounting for Tracks." Publisher: The University of Chicago Press, *Journal of Labor Economics* 32, no. 4 (October): 645–684. ISSN: 0734-306X, accessed October 31, 2022. https://doi.org/10.1086/676017. https://www.journals.uchicago.edu/doi/full/10.1086/676017.

Jackson, Erika, and Marianne E. Page. 2013. "Estimating the distributional effects of education reforms: A look at Project STAR." *Economics of Education Review* 32 (February 1, 2013): 92–103. ISSN: 0272-7757, accessed November 3, 2022. https://doi.org/10.1016/j.econedurev.2012.07.017. https://www.sciencedirect.com/science/article/pii/S0272775712001033.

Jamison, Eliot A., Dean T. Jamison, and Eric A. Hanushek. 2007. "The effects of education quality on income growth and mortality decline." *Economics of Education Review,* Economics of Education: Major Contributions and Future Directions - The Dijon Papers, 26, no. 6 (December 1, 2007): 771–788. ISSN: 0272-7757, accessed November 21, 2022. https://doi.org/10.1016/j.econedurev.2007.07.001. https://www.sciencedirect.com/science/article/pii/S0272775707000611.

Jimenez, Laura. 2020. *Preparing American Students for the Workforce of the Future.* Center for American Progress, September 14, 2020. Accessed March 2, 2023. https://www.americanprogress.org/article/preparing-american-students-workforce-future/.

Kane, Thomas J., Daniel F. McCaffrey, Trey Miller, and Douglas O. Staiger. 2013. *Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment. Research Paper. MET Project.* Publication Title: Bill & Melinda Gates Foundation ERIC Number: ED540959. Bill & Melinda Gates Foundation, January. Accessed October 31, 2022. https://eric.ed.gov/?id=ED540959.

Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2008. "What does certification tell us about teacher effectiveness? Evidence from New York City." *Economics of Education Review* 27, no. 6 (December 1, 2008): 615–631. ISSN: 0272-7757, accessed March 31, 2023. https://doi.org/10.1016/j.econedurev.2007.05.005. https://www.sciencedirect.com/science/article/pii/S0272775707000775.

Kane, Thomas J., and Douglas O. Staiger. 2008. *Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation,* 14607, December. Accessed November 16, 2022. https://doi.org/10.3386/w14607. https://www.nber.org/papers/w14607.

Kirkebøen, Lars J. 2022. *School Value-Added and Long-Term Student Outcomes,* 4123829, Rochester, NY. Accessed March 23, 2023. https://doi.org/10.2139/ssrn.4123829. https://papers.ssrn.com/abstract=4123829.

Koedel, Cory, and Julian R. Betts. 2011. "Does Student Sorting Invalidate Value-Added Models of Teacher Effectiveness? An Extended Analysis of the Rothstein Critique." *Education Finance and Policy* 6, no. 1 (January 1, 2011): 18–42. ISSN: 1557-3060, accessed April 3, 2023. https://doi.org/10.1162/EDFP_a_00027. https://doi.org/10.1162/EDFP_a_00027.

Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff. 2015. "Value-added modeling: A review." *Economics of Education Review* 47 (August 1, 2015): 180–195. ISSN: 0272-7757, accessed October 30, 2022. https://doi.org/10.1016/j.econedurev.2015.01.006. https://www.sciencedirect.com/science/article/pii/S0272775715000072.

Koenker, Roger, and Gilbert Bassett. 1978. "Regression Quantiles." *Econometrica* 46 (1): 33–50. ISSN: 00129682, 14680262, accessed April 5, 2023. http://www.jstor.org/stable/1913643.

Koenker, Roger, and Kevin F. Hallock. 2001. "Quantile Regression." *Journal of Economic Perspectives* 15, no. 4 (December): 143–156. https://doi.org/10.1257/jep.15.4.143. https://www.aeaweb.org/articles?id=10.1257/jep.15.4.143.

Ladd, Helen F, and Randall P Walsh. 2002. "Implementing value-added measures of school effectiveness: getting the incentives right." *Economics of Education Review* 21, no. 1 (February 1, 2002): 1–17. ISSN: 0272-7757, accessed March 2, 2023. https://doi.org/10.1016/S0272-7757(00)00039-X. https://www.sciencedirect.com/science/article/pii/S027277570000039X.

Lauff, Erich, and Steven J. Ingels. 2014. *Education Longitudinal Study of 2002 (ELS:2002): A First Look at 2002 High School Sophomores 10 Years Later. First Look. NCES 2014-363.* Publication Title: National Center for Education Statistics ERIC Number: ED544672. National Center for Education Statistics, January. Accessed March 30, 2023. https://eric.ed.gov/?id=ED544672.

Paufler, Noelle A., and Audrey Amrein-Beardsley. 2014. "The Random Assignment of Students into Elementary Classrooms: Implications for Value-Added Analyses and Interpretations." Publisher: SAGE Publications ERIC Number: EJ1021926, *American Educational Research Journal* 51, no. 2 (April): 328–362. ISSN: 0002-8312, accessed April 6, 2023. https://doi.org/10.3102/0002831213508299.

Penner, Emily K. 2016. *Teaching for All? Teach for America's Effects across the Distribution of Student Achievement.* Publication Title: Grantee Submission ERIC Number: ED601018. Accessed December 2, 2022. https://eric.ed.gov/?id=ED601018.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." _Eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0262.2005.00584.x, *Econometrica* 73 (2): 417–458. ISSN: 1468-0262, accessed November 27, 2022. https://doi.org/10.1111/j.1468-0262.2005.00584.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0262.2005.00584.x.

Rothstein, Jesse. 2009. "Student Sorting and Bias in Value-Added Estimation: Selection on Observables and Unobservables." *Education Finance and Policy* 4, no. 4 (October 1, 2009): 537–571. ISSN: 1557-3060, accessed October 31, 2022. https://doi.org/10.1162/edfp.2009.4.4.537. https://doi.org/10.1162/edfp.2009.4.4.537.

———. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement*." *The Quarterly Journal of Economics* 125, no. 1 (February 1, 2010): 175–214. ISSN: 0033-5533, accessed October 30, 2022. https://doi.org/10.1162/qjec.2010.125.1.175. https://doi.org/10.1162/qjec.2010.125.1.175.

Rothstein, Jesse, and William J. Mathis. 2013. *Review of "Have We Identified Effective Teachers?" and "A Composite Estimator of Effective Teaching: Culminating Findings from the Measures of Effective Teaching Project".* Publication Title: National Education Policy Center ERIC Number: ED539299. National Education Policy Center, January. Accessed November 17, 2022. https://eric.ed.gov/?id=ED539299.

Shea, Zhiling Meng, and Jade Marcus Jenkins. 2021. "Examining Heterogeneity in the Impacts of Socio-Emotional Curricula in Preschool: A Quantile Treatment Effect Approach." *Frontiers in Psychology* 12. ISSN: 1664-1078, accessed November 4, 2022. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.624320.

Smeeding, Timothy. 2012. "Income, wealth, and debt and the Great Recession." *Report. New York, NY, The Russell Sage Foundation and the Stanford Center on Poverty and Inequality.*

Tobias, Justin, and Gary Koop. 2004. "Learning about heterogeneity in returns to schooling." *Journal of Applied Econometrics* 19 (November 1, 2004). https://doi.org/10.1002/jae.744.

A Appendix

A.1 Descriptive Statistics

Variable	(1) Low VA Mean/(SE)	(2) High VA Mean/(SE)	(1)-(2) Pairwise t-test Mean difference
Adult Income	20997.573 (343.164)	21974.961 (379.198)	-977.389*
Socioeconomic status	-0.832 (0.032)	-0.689 (0.032)	-0.142***
Attended PS Institution	0.627 (0.010)	0.689 (0.010)	-0.062***
Attended highly selective PS	0.169 (0.009)	0.209 (0.011)	-0.040***
Ever held a job-2006	0.794 (0.006)	0.804 (0.006)	-0.010
Ever held a job-2011	0.796 (0.005)	0.812 (0.006)	-0.016**
Single parent	0.030 (0.002)	0.026 (0.002)	0.004
Residential Mobility - F3 vs BY	156.127 (7.540)	145.532 (7.360)	10.595
Number of observations	8469	7423	15892
Number of clusters	405	346	751

Table A1 Descriptive Statistics for Outcomes in Adulthood—-Descriptive statistics, with standard errors clustered by school.

A.2 Probit Regressions

This section presents probit regression results, as a robustness check for our mean regression findings which use a linear probability model. Complementing findings in the main paper, probit regressions were run for all relevant binary outcome variables.

VARIABLES	(1) Attended a Postsecondary Institution	(2) Attended a Postsecondary Institution	(3) Attended a Highly selective Institution	(4) Received Scholarship in First Year
School value-added	0.0449***	0.131***	0.0926***	0.0727***
	(0.00524)	(0.0141)	(0.0188)	(0.0142)
Constant	0.573***	0.0769	-0.612**	-0.405**
	(0.0381)	(0.180)	(0.262)	(0.196)
Observations	15,892	15,892	15,885	15,892
R-squared	0.118			
Baseline controls	Yes	Yes	Yes	Yes
Family controls	No	Yes	Yes	Yes

Table A2 Probit Regression analogue to Table 1's OLS results, with standard errors clustered by school. The dependent variable in Columns (1) and (2) is an indicator for ever attending a post-secondary institution. Column (1) omits the controls on family characteristics that includes number of siblings, family composition, parental education, and household income. Column (3) indicates attendance of a highly selective institution, referred to those whose first-year students' test scores places them in roughly the top fifth of baccalaureate institutions. Column (4) indicates a grant or scholarship offer in the respondent's first year of study.

VARIABLES	(1) Job-2006	(2) Job-2011	(3) Single Parent
School value-added	0.0541***	0.0446***	-0.0347*
	(0.0141)	(0.0130)	(0.0189)
Constant	0.286	0.382**	-2.046***
	(0.193)	(0.193)	(0.377)
Observations	15,885	15,892	15,781
Baseline controls	Yes	Yes	Yes
Family controls	Yes	Yes	Yes

Table A3 Probit Regression analogue to Table 2's OLS results, with standard errors clustered by school. Column (1) and (2) indicate whether the respondent had held a job since leaving high school—first recorded in the 2006 questionnaire, and again in 2011. Column (3) indicates respondents who are currently unmarried or not in a marriage-like relationship, with a biological or adopted child that lives with them in the household. All regressions include controls on baseline and parental characteristics. ***p<0.01,**p<0.05,*p<0.1

A.3 Mean Outcomes by Subgroup

This section of the appendix expands on our analysis by calculating mean impact by subgroup. To ensure that quantile treatment effects specifically capture heterogeneity across demographic groups, we run standard OLS regressions and restrict the sample to various subgroups discussed in the paper: men, women, black students, and white students.

Interestingly, women appear to have a slightly greater estimated average gain to earnings from school quality compared with men. Black students on average continue to see significantly fewer gains to earnings from school quality compared to their white peers, although the sample of Black students in the ELS survey is relatively small.

VARIABLES	(1) Female	(2) Male	(3) Black	(4) White
School value-added	0.196***	0.166**	0.00285	0.177***
	(0.0601)	(0.0795)	(0.113)	(0.0607)
Constant	4.751***	4.317***	2.918	1.757
	(1.273)	(1.376)	(2.057)	(1.914)
Observations	7,666	7,578	2,020	8,682
R-squared	0.057	0.050	0.070	0.052
Baseline controls	Yes	Yes	Yes	Yes
Family controls	Yes	Yes	Yes	Yes

Table A4 Mean Impacts by Demographic Group—OLS Results, with standard errors clustered by school. Dependent variable is adult income, transformed by an inverse hyperbolic sign. *** p<0.01, ** p<0.05, * p<0.1

A.4 Quantile Regression Results

In this section, we display the unconditional quantile regression results for the demographic groups discussed in the main paper. Tables A6 and A5 show the output for unconditonal quantile regression results on adult income for women and men. For the graphs illustrating these findings, refer to Figures 5 and 6 in the main paper. Tables A7 and A8 do a similar exercise for Black and White students, respectively, with the graphs shown in figures A1 and A2, although the sample of nonwhite students in the ELS:2002 survey is too small for credible quantile analysis.

	(log)Adult Income								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
School value-added	0.0410	0.174**	0.249**	0.0894**	0.0567*	0.0480*	0.0511**	0.0432**	0.0195
	(0.0247)	(0.0644)	(0.0952)	(0.0301)	(0.0220)	(0.0208)	(0.0170)	(0.0159)	(0.0143)
Observations	7666	7666	7666	7666	7666	7666	7666	7666	7666

Table A5 Unconditional Quantile Effects on Adult Income for Women — Standard errors clustered by school. The dependent variable is (log)adult income. ***p<0.001,**p<0.01,*p<0.05

	(log)Adult Income								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
School value-added	0.0223	0.0223	1.017*	0.0554	0.0435	0.0720**	0.0550*	0.0353	0.0348
	(0.0369)	(0.0369)	(0.508)	(0.0399)	(0.0305)	(0.0268)	(0.0250)	(0.0254)	(0.0185)
Observations	7578	7578	7578	7578	7578	7578	7578	7578	7578

Table A6 Unconditional Quantile Effects on Adult Income for Men — Standard errors clustered by school. The dependent variable is (log)adult income. ***p<0.001,**p<0.01,*p<0.05

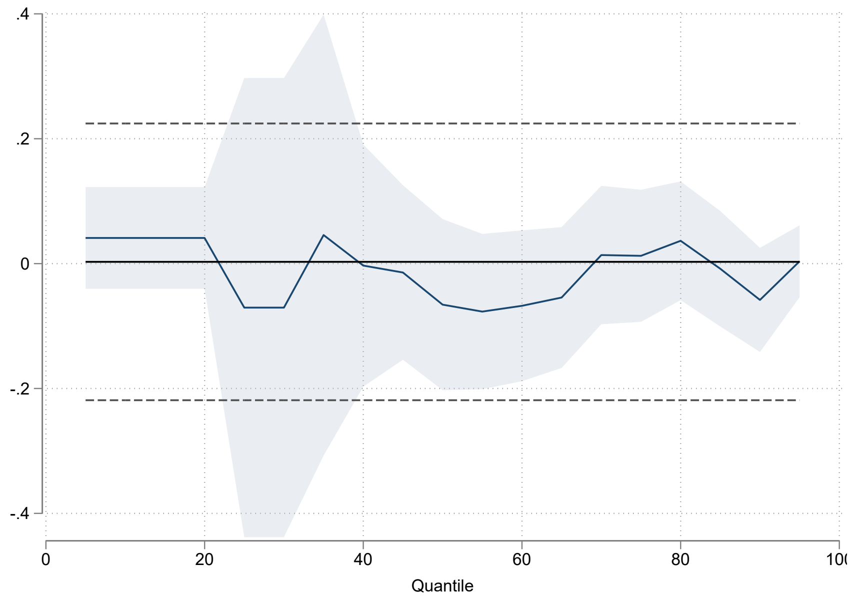

Fig. A1 Unconditional Quantile Treatment Effects on Adult Income for Blacks

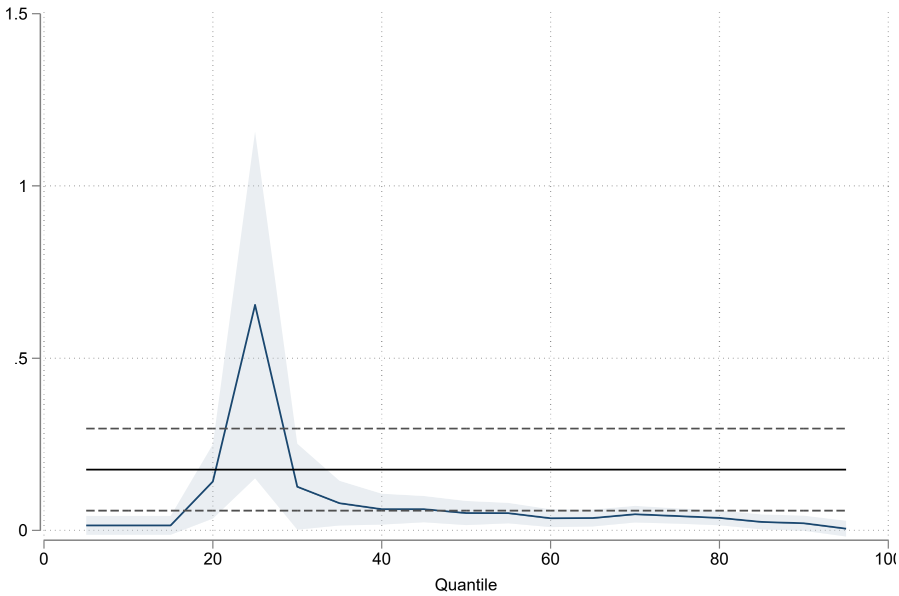

Fig. A2 Unconditional Quantile Treatment Effects on Adult Income for Whites

	(log)Adult Income								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
School value-added	0.0411 (0.0415)	0.0411 (0.0415)	-0.0705 (0.187)	-0.00312 (0.0987)	-0.0657 (0.0697)	-0.0675 (0.0615)	0.0137 (0.0564)	0.0366 (0.0484)	-0.0581 (0.0425)
Observations	2020	2020	2020	2020	2020	2020	2020	2020	2020

Table A7 Unconditional Quantile Effects on Adult Income for Blacks—Standard errors clustered by school. The dependent variable is (log)adult income. ***p<0.001,**p<0.01,*p<0.05

	(log)Adult Income								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
School value-added	0.0145 (0.0140)	0.142* (0.0550)	0.127* (0.0636)	0.0615** (0.0229)	0.0503** (0.0179)	0.0352** (0.0132)	0.0469*** (0.0123)	0.0363** (0.0110)	0.0206 (0.0113)
Observations	8682	8682	8682	8682	8682	8682	8682	8682	8682

Table A8 Unconditional Quantile Effects on Adult Income for Whites—Standard errors clustered by school. The dependent variable is (log)adult income. ***p<0.001,**p<0.01,*p<0.05