Human-Capital Formation During Childhood and Adolescence: Evidence from School Quality and Postsecondary Success in California

Naven, Matthew

Washington and Lee University

19 December 2019

Human-Capital Formation During Childhood and Adolescence: Evidence from School Quality and Postsecondary Success in California

Matthew Naven*

Washington and Lee University

Draft Date: December 19, 2019

Download the most recent version at

http://matthewnaven.com/research/school_va_timing.pdf

Abstract

This paper investigates the role of school quality in human-capital formation. Specifically, I investigate how the timing of school quality differentially affects long-run outcomes. Using individual-level data on the universe of public-school students in California, I estimate elementary-, middle-, and high-school quality using a value-added methodology that accounts for the fact that students sort to schools on observable characteristics. I then determine the impact of school quality on future K–12 and postsecondary outcomes. I find that high-school quality has the largest impact on postsecondary enrollment, while elementary- and middle-school quality play a larger role in college readiness. In other words, early human-capital investments are important for future postsecondary success, but the unique timing of the college decision process allows for later human-capital investments to also play a significant role. [JEL Codes: **H75, I21, I23, J24**]

*Contact the author at mnaven@wlu.edu or visit www.matthewnaven.com for more information. I am grateful to Scott Carrell, Marianne Bitler, Shu Shen, Michal Kurlaender, Paco Martorell, Sherrie Reed, and other members of the California Education Lab for helpful comments and suggestions. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305E150006 to the Regents of the University of California. The opinions expressed are those of the author and do not necessarily represent views of the Institute, the U.S. Department of Education, or of the agencies providing data. I gratefully acknowledge support from the Russell J. and Dorothy S. Bilinski Educational Foundation Dissertation Writing Fellowship. I would like to thank the participants at the University of California, Davis, the University of Nevada, Reno, the All-California Labor Economics Conference, Sonoma State University, APPAM, Analysis Group, Cornerstone Research, Kennesaw State University, IMPAQ International, the Federal Communications Commission, Washington and Lee University, AEFP, and the APPAM California Regional Student Conference seminars for their insights.

1 Introduction

The timing of human-capital investments is an important determinant of their efficacy. Because early human-capital investments both augment and improve the productivity of later human-capital investments (Cunha et al., 2006), early interventions have been shown to be the most effective and efficient (Heckman, Krueger and Friedman, 2003; Heckman, 2006; Doyle et al., 2009). However, the effects of early human-capital interventions tend to fade out rather quickly with regards to cognitive achievement (Currie and Thomas, 1995; Bitler, Hoynes and Domina, 2014), although they often reappear when examining long-run outcomes (Garces, Thomas and Currie, 2002; Ludwig and Miller, 2007; Deming, 2009; Duncan and Magnuson, 2013). Dynamic complementarities may play a large role in the persistence of early interventions, as evidence suggests that the benefits of an early educational intervention are larger when followed by access to better-funded public K–12 schools (Johnson and Jackson, 2017).

Nevertheless, human-capital formation and interventions that occur after pre-school have also been shown to have lasting effects on long-run outcomes. Elementary- and middle-school students assigned a high-quality teacher have a lower likelihood of having children as teenagers, a higher likelihood of attending college, and earn higher salaries. The effects are substantial, as replacing a teacher in the bottom five percent of the distribution with an average teacher would increase the present value of students' lifetime income by about $250,000 (Chetty, Friedman and Rockoff, 2014*b*).[1] Later interventions can matter too. Students given college advising and mentoring as late as their senior year of high school have been shown to be more likely to enroll and persist in college (Carrell and Sacerdote, 2013; Castleman and Goodman, 2018; Barr and Castleman, 2019). Effective interventions can also occur in college (Carrell, Page and West, 2010; Barr and Castleman, 2019) or the labor market (Schochet, Burghardt and McConnell, 2008).

This paper investigates how the timing of human-capital investments affects their efficacy within the context of school quality. Schools are a large contributor to human-capital formation. While teachers can have important impacts on the long-run outcomes of their students[2], there are other factors within a school that may determine student outcomes as well, such as principals (Clark, Martorell and Rockoff, 2009; Horng, Klasik and Loeb, 2010; Loeb, Kalogrides and Horng, 2010; Grissom and Loeb, 2011; Ladd, 2011; Béteille, Kalogrides and Loeb, 2012; Branch, Hanushek and Rivkin, 2012; Loeb, Kalogrides and Béteille, 2012;

[1]Despite these lasting effects, the persistence of teacher-induced learning is low from grade to grade (Jacob, Lefgren and Sims, 2010). This paradox may be explained by the fact that teachers may have long-term effects that are not initially apparent on contemporaneous test scores but manifest themselves in the future (Carrell and West, 2010).

[2]There is a large literature on teacher quality. See, for example, Rockoff (2004); Hanushek et al. (2005); Jacob and Lefgren (2005); Rivkin, Hanushek and Kain (2005); Hanushek and Rivkin (2006); Kane, Rockoff and Staiger (2008); Kane and Staiger (2008); Ishii and Rivkin (2009); Rothstein (2009); Carrell and West (2010); Corcoran (2010); Hanushek and Rivkin (2010); Jacob, Lefgren and Sims (2010); Rothstein (2010); Hanushek (2011); Kinsler (2012); Bacher-Hicks, Kane and Staiger (2014); Bitler et al. (2014); Chetty, Friedman and Rockoff (2014*a*,*b*); Staiger and Kane (2014); Guarino et al. (2015); De Vlieger, Jacob and Stange (2017); Chetty, Friedman and Rockoff (2017); Rothstein (2017).

Gates et al., 2014; Grissom, Blissett and Mitani, 2018), counselors (Carrell and Carrell, 2006; Reback, 2010; Carrell and Hoekstra, 2014), curricula (Altonji, 1995; Yu and Mocan, 2018), and expenditures (Hanushek, 1989; Hanushek, Rivkin and Taylor, 1996; Hanushek, 1997, 2003; Martorell, Stange and McFarlin Jr, 2016; Lafortune, Rothstein and Schanzenbach, 2018). Additionally, high-quality teachers may sort to schools based on school or location characteristics (Lankford, Loeb and Wyckoff, 2002; Ladd, 2011), further reinforcing the role that schools play in providing high-quality instruction to their students. Moreover, while parents have some limited influence on teacher assignments within a school, they have a much larger influence on which school their children attend. Thus, it is important to understand how broad measures of school quality affect longer-run outcomes.

Despite a growing literature on estimating causal school quality (Abdulkadiroğlu et al., 2011; Dobbie and Fryer Jr, 2011; Pop-Eleches and Urquiola, 2013; Deming et al., 2014; Deming, 2014; Dobbie and Fryer Jr, 2015; Angrist et al., 2016; Dobbie and Fryer Jr, 2016; Abdulkadiroğlu et al., 2017; Angrist et al., 2017; Hubbard, 2017), it is not clear *when* attending a high-quality school matters most. Although there are studies on school quality at the elementary-, middle-, and high-school level, to my knowledge this is the first paper that compares the long-run effects of school quality across school levels. Using individual-level data on the universe of California public-school students linked to postsecondary records, I explore how elementary-, middle-, and high-school quality affect both the extensive and intensive margins of postsecondary outcomes. The extensive margin, postsecondary enrollment, may be affected by aspects of a school's quality beyond the cognitive skills the school teaches (such as non-cognitive skills, information on the college application process, or a culture of college attendance). The intensive margin, measured by a student's college readiness and persistence, is much more likely to be affected by cognitive skills alone. Schools that play a large role in one margin may not necessarily impact the other, as each school level may impart different skills throughout a student's education.

I calculate school quality by extending the value added with drift methodology, as in Chetty, Friedman and Rockoff (2014*a*), to schools. The drift methodology, which allows value added to change from year to year, is particularly suited to the school quality setting, as schools experience faculty and staff turnover that could lead to changes in quality from year to year. I estimate how school value added on standardized test scores translates to postsecondary success as well as estimate a school's total value added on postsecondary enrollment, which includes both test score and non-test score factors.

Results show that high-school quality has the largest impact on the extensive margin of postsecondary enrollment. A one standard deviation increase in high-school value added increases postsecondary enrollment by 2.2 percentage points (3.4%) and 4-year enrollment by 2.8 percentage points (10.3%). However, elementary- and middle- school quality have the largest effects on the intensive margin, such as persistence and the need for remedial classes upon enrollment. A one standard deviation increase in elementary-school value added increases persistence to year two at four-year colleges by 1.2 percentage points (1.4%). A one standard deviation increase in middle-school value added reduces the need for English and math remediation by 2.2 percentage points (9.5%) and 3.2 percentage points (14%) respectively. Thus, results indicate that earlier grades give students the tools to succeed in college while high schools play the largest role in the postsecondary education decision process.

I then correlate the value-added estimates with observable school characteristics in order to determine which school inputs are correlated with school quality. Surprisingly, there appears to be little to no pattern to these inputs and my value-added estimates. One exception is that funding for after-school programs is correlated with higher value added on postsecondary enrollment, thus, after-school supervision may have important long-term effects. While I find few patterns in the characteristics of schools and school value added, I find that high value added schools tend to cluster in populous areas surrounding California's major metropolitan areas such as Los Angeles and the Bay Area.

This paper adds numerous important contributions to the literature on human capital broadly and education quality specifically. First, this is the first paper to study how school quality differentially contributes to human-capital formation at various points during a student's educational career. Second, this paper is unique in that it links the universe of public-school students in California, which has the largest public-school population in the United States, to their postsecondary outcomes. California is a particularly relevant state in which to study postsecondary outcomes because California has a robust postsecondary infrastructure that includes two-year community colleges, teaching universities, and globally-ranked research universities. Finally, this paper provides new insights on the relationship between K–12 school quality and measures of the intensive margin of postsecondary enrollment, which informs us about how schools contribute to college readiness.

2 Data

My study uses individual-level data on the universe of public-school students in the state of California. Standardized test score information comes from the California Standards Test (CST). Data from the CST spans the 2002–2003 to 2012–2013 school years[3] and tests students in English language arts (ELA) and math during grades 2–11. The data also include demographic information on each student, such as sex, race, economic disadvantage status, limited English proficiency status, and whether or not the student has

[3]Due to the fact that I use test scores from two grades prior as a control variable, I only calculate value-added estimates for the years 2004–2005 to 2012–2013.

a disability. State student IDs can be used to link students to prior test scores across time. Each cohort consists of about 475,000 students, which makes this the largest ever study on school quality.

Starting in the 7th grade, students have the option of taking different math assessments based on the math subject in which they are enrolled. This makes calculating school value added in math difficult, because scores are not directly comparable between the various math subjects within grades. Because all students take the same ELA exam in each grade, my primary analyses will investigate school quality on the ELA exam. In appendix section D I present results for math value added in elementary school, where there is a common test.[4] Although studying differences between subjects for all levels of schooling would be ideal, Master, Loeb and Wyckoff (2017) show that ELA value added persists into future test scores on both ELA and math exams while math value added only persists to future math scores. Thus, it is likely that school-induced learning on ELA subject matter imparts long-term skills that are broadly applicable, which may be important for postsecondary success.

Table D.1 gives summary statistics for the CST data by school level for the test-score value-added sample and includes all the dependent and independent variables used in the value added analyses. Appendix Table D.12 shows the limitations that are imposed in order to form the value-added sample, which are similar to those made in the teacher value added literature. The vast majority of students in the CST data that cannot be included in the value added estimation are excluded because they lack prior test scores, although in high school an almost equal number of students are excluded because they attend alternative[5] high schools. For my analyses elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11.[6] I exclude grades 2–3 because they lack sufficient prior test scores in order to estimate value added.[7]

[4]The standard deviation of elementary-school math value added is about twice the size of that for ELA, which is consistent with prior studies of school and teacher value added. Unsurprisingly, elementary-school math value added has a larger positive impact on math scores in the next grade than ELA scores (while elementary-school ELA value added impacts future scores in both subjects similarly). Elementary-school math value added also has a positive impact on the math subject that students take, as students are more likely to take the hardest math subject when they first track to different math subjects in grade 7 as well as for their final math exam in grade 11. Interestingly, elementary-school math value added has a smaller impact on the 11th grade math subject than elementary-school ELA value added, although elementary-school math value added explains a larger proportion of the variation in math subject. With regards to the impact of school value added on postsecondary outcomes, elementary-school ELA value added has a larger impact on overall postsecondary enrollment, four-year university enrollment, CSU English remediation, CSU math remediation (surprisingly), CSU persistence, and transfer from a CCC to a four-year university than math.

[5]This includes schools in the following categories: Special Education Schools (Public), County Community, Youth Authority Facilities (CEA), Opportunity Schools, Juvenile Court Schools, Other County or District Programs, State Special Schools, Alternative Schools of Choice, Continuation High Schools, District Community Day Schools, Adult Education Centers, and Regional Occupational Center/Program (ROC/P).

[6]California's elementary-school grade spans are somewhat equally split between K-5 and K-6 schools (in 2018–2019 2,545 schools taught grades K-5 and 1,951 schools taught grades K-6). This leads to a fair amount of grades 7–8 middle schools (in 2018–2019 879 middle schools taught grades 6–8 and 332 taught grades 7–8). As the K-5, 6–8, and 9–12 model is the most common, I elected to use these grade splits. However, this does lead to some "middle schools" that are simply the 6th grade cohort of an elementary school. A similar problem would have arisen if I elected to use a K-6, 7–8, and 9–12 grade split, as some "elementary schools" would simply be the 6th grade cohort of a middle school.

[7]Prior test scores are necessary in order to obtain unbiased estimates when using value added methodologies (Kane and Staiger, 2008; Deming, 2014).

Hispanics are the largest racial group in California, followed by whites, Asians, blacks, and other-race[8] students. Almost 60% of students in elementary school are socioeconomically disadvantaged[9], although this percentage declines slightly in more advanced school levels. Around a quarter of students are limited English proficient in elementary school, although this also declines as students age, likely due to the fact that students are reclassified as English proficient or higher dropout rates for limited English proficient students. About 4% of the sample has some type of disability. As is the case in other value added studies, the value-added sample is positively selected on prior test scores, as they score anywhere from 0.06 to 0.14 standard deviations above average on their current test scores.[10] Appendix section A gives more information on the data.[11]

Postsecondary data comes from the National Student Clearinghouse (NSC), the California State University (CSU) system, and the California Community College (CCC) system. The NSC data includes enrollment and degree receipt data for the cohorts of students that graduated high school between the spring of 2010 and 2017, inclusive.[12] The NSC data includes all types of universities in the United States and, in particular, accounts for the lack of data from the University of California (UC), private California universities, and out of state universities that the CSU and CCC data do not account for. The CSU files include application and enrollment files from fall 2001 to spring 2017 and degree receipt files from fall 2001 to spring 2016. The CCC files include enrollment files from fall 1992 to spring 2017 and degree receipt files from fall 1992 to spring 2016. Appendix section B explains the details of the match between the K–12 and postsecondary data. Table D.2 gives an overview of all of the datasets used in this paper.

3 School Value Added

3.1 Model

In this section I describe a model of student learning in order to better describe which factors contribute to a school's value added measure. Suppose that the outcome of a student i in grade g of school s in year t is determined according to equation (1), such that a student's endowment ι_i, contemporaneous learning ℓ_{ig}, prior learning ℓ_{ik} (depreciated by a factor δ_k), and idiosyncratic school-level shocks θ_{st} all contribute.

[8] The other category includes Native Americans and two or more races.

[9] Defined by the California Department of Education (CDE) as "a student neither of whose parents have received a high-school diploma or a student who is eligible for the free or reduced-price lunch program, also known as the National School Lunch Program (NSLP)."

[10] Test scores are standardized to have mean zero and standard deviation one at the grade-by-year level on the entire population of students taking the CST.

[11] The value-added sample differs from the overall population of students on a few demographic characteristics due to sample restrictions. The high-school value-added sample is 8% less likely to be male, 37% less likely to be black, 43% less likely to be limited English proficient, and 57% less likely to have a disability than the students who are excluded from the value-added sample. Appendix Table D.13 gives a comparison between the included and excluded students.

[12] The cohorts matched were actually spring 2009 to spring 2016 11th grade students, because we do not observe high-school graduation data nor the students in 12th grade.

Assume that students take each grade only once, so that g and t are interchangeable within student.

$$z_{isgt} = \underbrace{\iota_i}_{\text{Endowment}} + \sum_{k=0}^{g-1} \underbrace{\delta_k \cdot \ell_{ik}}_{\text{Prior Learning}} + \underbrace{\underbrace{\ell_{ig}}_{\text{Learning}} + \underbrace{\theta_{st}}_{\text{Shocks}} + \underbrace{\varepsilon_{isgt}}_{\text{Noise}}}_{r_{isgt}} \tag{1}$$

Assume that the portion of outcome z_{isgt} that is due to learning is modeled by equation (2) such that teachers τ_{sgt} and other school factors ψ_{st} (such as principals, counselors, curricula, extracurricular activities, and peers) contribute to student learning.

$$\ell_{ig} = \underbrace{\tau_{sgt}}_{\text{Teachers}} + \underbrace{\psi_{st}}_{\text{School Factors}} \tag{2}$$

While other studies have investigated the impact of τ_{sgt} on long-run outcomes, studying school quality allows ψ_{st} to also have an impact. This may be particularly important when studying the effects of education on postsecondary enrollment, as high schools are much more likely to have counselors dedicated to the postsecondary decision process and some schools may have better resources on the application process, such as college fairs or mandatory SAT/ACT testing, than others.

Note that by regressing the test score in grade g on the test score in grade $g-1$ it is possible to control for ι_i and $\sum_{k=0}^{g-1} \delta_k \cdot \ell_{ik}$, the performance a student would achieve even in the absence of school input. This leaves us with the residual term r_{isgt}, which captures the portion of student performance that is not related to the student's prior achievement.

3.2 Methodology

To estimate ℓ_{ig}, I extend the value-added methodology that allows for drift over time described in Chetty, Friedman and Rockoff (2014*a*) to the school level. The value-added methodology accounts for the fact that schools receive students of varying backgrounds.[13] Hence, schools that receive only the lowest-performing students should not be penalized for the fact that the students they receive will likely have lower outcomes on average. Instead, they should be evaluated on how much they improve the outcomes of those students, regardless of the students' prior achievement. Thus, a school that improves the test scores of the lowest-performing students would be determined to have a higher value added than a school that made no change to the test scores of the highest-performing students, even though the latter school's students may perform better on average.

A school's value added is calculated by first removing the portion of each student's test score that is due

[13] Value-added methodologies were first pioneered in estimating school and hospital quality (Willms and Raudenbush, 1989; McClellan and Staiger, 1999, 2000). Meyer (1997) and Everson (2017) provide some background on the methodology.

to non-school factors. To do so, I regress student test scores z_{isgt} on cubic polynomials in prior test scores $\boldsymbol{z}_{ig}$, demographic characteristics $\boldsymbol{X}_{it}$, and the number of students in a student's cohort[14] $\boldsymbol{W}_{sgt}$ (defined as school-by-grade-by-year) as in equation (3). The cubic polynomials in prior test scores account for mean reversion and the fact that students with low test scores have more room to improve than students with high test scores. I also include grade fixed effects γ_g and year fixed effects ψ_t. The demographic characteristics $\boldsymbol{X}_{it}$ contain a linear term for age and fixed effects for sex, ethnicity[15], socioeconomic disadvantage, limited English proficiency, and disability status.

$$z_{isgt} = \boldsymbol{z}_{ig}\boldsymbol{\delta}_g + \boldsymbol{X}_{it}\boldsymbol{\beta}_X + \boldsymbol{W}_{sgt}\boldsymbol{\beta}_W + \gamma_g + \psi_t + r_{isgt} \tag{3}$$

Because there could be idiosyncratic shocks that are uncorrelated with school quality but influence the performance of all students within a school in each year, such as the proverbial dog barking outside of the school on the day of the test, the residual term r_{isgt} will contain school value added μ_{st}, idiosyncratic shocks θ_{st}, and a student-level error term ε_{isgt} as in equation (4).

$$r_{isgt} = \mu_{st} + \theta_{st} + \varepsilon_{isgt} \tag{4}$$

Under the assumptions that ε_{isgt} is a mean zero random error term and students do not sort to schools in each year on *unobservable* characteristics, the student-level error terms have expected value zero conditional on school and year, which gives us equation 5.

$$\mathbf{E}[r_{isgt}|s,t] = \mu_{st} + \theta_{st} \tag{5}$$

I therefore average the residual r_{isgt} to the school-by-year level in order to eliminate the student-level error term. However, because value added and idiosyncratic shocks are the same for all students at this level, the average residual will contain both school value added and the school-level idiosyncratic shock as in equation (6).

$$\bar{r}_{st} = \mu_{st} + \theta_{st} \tag{6}$$

In order to reduce the variation from the idiosyncratic shocks while retaining the variation in school value added, I project the average residual in year t onto the residuals in all other years t' (jackknife projection)

[14]Due to the inclusion of a school fixed effect (as described in equation (10) of appendix section C) this controls for year-to-year changes in cohort size *within* schools.

[15]Asian, Hispanic, black, and other; white is omitted.

as in equation (7).

$$\bar{r}_{st} = \bar{\boldsymbol{r}}_{st'}\boldsymbol{\beta}_{\bar{r}t'} + \epsilon_{st} \tag{7}$$

The value-added estimates that I use in this paper are the predicted values from equation (7), $\hat{\mu}_{st} = \bar{\boldsymbol{r}}_{st'}\hat{\boldsymbol{\beta}}_{\bar{r}t'}$. However, I rescale the estimates so that they have mean zero for each school level-by-subject combination, thus schools with positive value added are above average and vice versa.[16] I outline additional methodological details in appendix section C.

This projection strategy has several advantages. Under the assumptions that school value added is correlated across years ($\text{cov}(\mu_{st}\mu_{st'}) \neq 0$), the school-level common shocks are uncorrelated across years ($\text{cov}(\theta_{st}\theta_{st'}) = 0$), and the school-level common shocks are not correlated with school value added across years ($\text{cov}(\mu_{st}\theta_{st'}) = 0$), the projection will utilize variation from school value added and remove variation from the common shocks when using school value-added estimates to predict long-run outcomes. In practice, the finite sample size in the number of years may lead to violations of the last two assumptions regarding θ_{st}, which is why the projection will reduce the variation from the idiosyncratic shocks instead of completely eliminating it.[17] If the common shocks are uncorrelated with long-run outcomes, then this strategy reduces attenuation bias due to measurement error. If the common shocks are correlated with long-run outcomes, then this strategy reduces bias that results from the coefficient on estimated school value added measuring the combined effect of school value added and the common shock. This strategy is also useful because it prevents the same estimation errors on both the left- and right- hand side of the regression when examining long-run outcomes, which would be the case if we used the average residual for a set of students to predict future outcomes for those same students.

3.3 Results

Figure D.1 shows the distributions of school value added. The standard deviation of school value added ranges from 0.066 for high school to 0.087 for middle school. This tells us, for example, that a one standard deviation increase in high-school value added increases the average test score of its students by 6.6% of a student-level standard deviation. The magnitudes are similar in size to those found for the distribution of school value added using charter school lotteries in Deming (2014) and for the distribution of teacher value

[16] This rescaling has no impact on the results to follow.

[17] If the common shocks are truly idiosyncratic, then the last two assumptions regarding θ_{st} are likely to hold as the number of years goes to infinity. Furthermore, to the extent that good or bad events happen continuously at the same schools, these should be considered part of a school's value added, which further reinforces that the common shocks are idiosyncratic. As for the first assumption, schools will experience some faculty and staff turnover, but school value added is likely to be correlated from year to year as the majority of the personnel will remain in the same school from one year to the next. Empirical evidence that this is true and that the correlation in school value added decreases as the gap in years increases is presented in Figures D.2 and D.10.

added in Chetty, Friedman and Rockoff (2014*a*).[18]

The drift methodology, which allows a school's value added to change from year to year, is only an improvement over prior value added methodologies if a school's value added actually varies across time. To illustrate that this is true in practice, figure D.2 shows the correlation between a school's value added estimate in year t and year t', where the horizontal axis gives the number of years between years t and t' and the vertical axis gives the correlation. Here we can see the importance of using the drift methodology. While a school's value added is highly correlated within a two-year window, the correlation begins to drop off as the number of years between estimates grows.

3.3.1 Validity Tests

There are three potential concerns regarding the validity of the value-added estimates. The first is that the estimates may be picking up noise due to sampling error and small sample variability. This would be the case if test scores are sufficiently noisy that student-level residual test scores, ε_{isgt}, do not average out to zero at each school even when schools have no effect on student performance (Bitler et al., 2019). If this were the case, we would attribute value added to schools when we were in fact just observing sampling error.

In order to measure how much of the estimated variation in school value added is due to noise, I calculate school value-added estimates after randomly assigning students to schools. I call these value-added estimates permuted value added, as I permute the school assignment vector within a grade by year cell. Figure D.3 shows the distributions of permuted value added, and I plot the distributions on the same axes as figure D.1 so that their variability can be directly compared. As can be seen, there is essentially no variation in school quality when students are randomly assigned to schools in this way. The largest permuted value added standard deviation relative to the actual value added standard deviation is 0.001 for high school, which is only 1.5% of the size of the actual value added standard deviation. These results alleviate concerns that the value-added estimates are merely an artifact of noisy test score measures or small sample variability.

Another concern is that the value-added estimates are the incorrect magnitude. Specifically, the issue is whether a one unit increase in school value added actually is associated with a one standard deviation increase in student test scores. In order to test this, I run a bivariate regression of residualized test scores r_{isgt} on the school value-added estimates μ_{st}, where the residualized test scores are calculated using equation (3). This follows the procedure used in Chetty, Friedman and Rockoff (2014*a*) and Rothstein (2017) and calculates by how much a school's estimated value added actually increases the test scores of its students. We expect the coefficient to equal one, which would indicate that a one unit increase in school value added

[18] The standard deviations are about a quarter of the size of those found for school value added in Angrist et al. (2017) and about half the size of those for teacher value added in Kane and Staiger (2008).

increases student test scores by one standard deviation on average.

The first row of Table D.3 provides this estimate along with its 95% confidence interval. The coefficient estimates range from 1.010 to 1.019, which are economically indistinguishable from one. Chetty, Friedman and Rockoff (2014*a*) obtain a coefficient estimate of 0.998. This gives evidence that the school value-added estimates have the correctly-sized effect on student test scores. Furthermore, figure D.4 graphs the relationship between μ_{sdt} and r_{isgt} in 20 equally sized bins. Results show that the value-added estimates and test score residuals have an almost perfectly linear relationship throughout the value added distribution.

The final concern, and potentially most problematic, involves the potential sorting of students to schools based on unobserved ability. If students with high unobserved ability sort to specific schools, such that $cov(\varepsilon_{isgt}, \mu_{st}) \neq 0$, then these schools' estimated value added will be higher than their true value added. However, this is only an issue if the sorting occurs on *unobserved* ability. Hence, there is no issue if students sort to schools on observed ability, because this will be controlled for with the inclusion of prior test scores and demographic controls. For example, if students with high test scores tend to attend the same schools, as occurs in practice, then we can still obtain unbiased estimates of school value added as long as prior test scores are included in the control vector so that $\mathbf{E}[\varepsilon_{isgt}|s,t] = 0$. In fact, research comparing value-added estimates to estimates obtained using random assignment to schools (Deming, 2014; Angrist et al., 2017) or teachers (Kane and Staiger, 2008) shows that once you control for prior test scores, even the inclusion of demographic characteristics in the control vector is essentially irrelevant because prior test scores are a sufficient statistic for student ability.

The primary threat to this assumption would be if students or parents *changed* their level of input into academic preparation between the student's prior grade and current grade and students sorted to schools based on this change in behavior. For example, if all students of parents who received an *increase* in income between grades, where the *extra* income was used to purchase academic assistance, attended the same school, then the estimated value added of this school would be positively biased. This is due to the fact that the prior test scores and demographic controls of those students would not control for this change in academic assistance, so $\mathbf{E}[\varepsilon_{isgt}|s,t] > 0$. If students whose parents *consistently* have high income sort to the same schools there would not be the same issue, because the students' prior test scores would also reflect their high socioeconomic status.

The issue in determining to what degree students sort to schools on unobserved ability is that, by definition, we have no measures of unobserved ability. However we can approximate unobserved ability using variables in our data that likely would be correlated with ability but that were not included as a control variable in equation (3). Given the available data, the best possible measure of unobserved student ability is an additional prior test score. Under the assumption that this omitted variable is the only component of

ε_{isgt} that is correlated with contemporaneous student test scores, we can obtain an estimate of $\frac{cov(\varepsilon_{isgt},\hat{\mu}_{st})}{var(\hat{\mu}_{st})}$. Chetty, Friedman and Rockoff (2014*a*) call this value forecast bias, which gives an estimate of what proportion of the variation in school value added is due to sorting on unobserved ability.[19]

The second row of Table D.3 provides the estimate of forecast bias along with its 95% confidence interval. Here we expect an estimate of zero, which would give evidence that there is no sorting of students to schools on unobservable characteristics. The estimates suggest that between 0.9% (middle school) and 3.9% (high school) of the variance in school value added is due to sorting on unobserved ability, thus selection on unobservables does not appear to be a large issue.[20] Chetty, Friedman and Rockoff (2014*a*) estimate forecast bias of 2.2%. Given that the forecast bias estimates are all negative, this would suggest that students who are unobservably *worse* tend to attend schools with higher value added. This would result in value-added estimates that are biased towards zero, thus the value-added estimates are slightly conservative if anything. Figure D.5 shows that this relationship holds throughout the distribution of school value added.

3.4 Value Added Versus Average Test Scores

Given the evidence shows my value-added estimates likely provide an unbiased measure of school quality, one might wonder whether the average test scores at a school could provide the same information. After all, parents interested in the academic performance of a school will most likely look at the average level of test scores at the school. Figure D.6 plots a school's value added against average test scores for those students. This figure shows that average test scores are not sufficient to predict value added. While average test scores and value added are positively related, as would be expected if value added causally impacted student test scores, average test scores do not account for the majority of the variation in school value added. In fact, the slope on the bivariate regression ranges from 0.059 to 0.084, depending on the school level, which would imply that only up to 8.4% of a school's increase in average test scores is due to the value added that that school provides. Furthermore, average test scores explain at most 24% of the variation in school value added. This indicates that a large proportion of the average test scores at a school is simply due to the type of students that enroll as opposed to any benefits the school provides.

[19]Similar to Chetty, Friedman and Rockoff (2014*a*), I estimate forecast bias using the following steps. First, I obtain the portion of contemporaneous test scores that projects onto three-grade prior test scores by adding three-grade prior test scores to equation (3). The projection is equal to the predicted value using only the test score from three grades prior. I then regress this projection on school value added.

[20]As with Chetty, Friedman and Rockoff (2014*a*), the coefficients from this test are statistically different than zero even though they are not economically different than zero, likely due to the large sample size.

4 Long-Run Outcomes

While I've established the variability and validity of school value added on test scores, test scores have no inherent meaning unless they have lasting effects that eventually translate to labor-market outcomes. I now examine whether school value added on test scores affects future K–12 outcomes, as well as the extensive and intensive margins of postsecondary enrollment. To do so, I run a regression of a student's outcome y_i on the student's school's value added as in equation (8). I run these regressions for each school level separately.

$$y_i = \hat{\mu}_{st}\beta_\mu + \boldsymbol{z}_{ig}\boldsymbol{\delta}_g + \boldsymbol{X}_{it}\boldsymbol{\beta}_X + \boldsymbol{W}_{sgt}\boldsymbol{\beta}_W + \gamma_g + \psi_t + \nu_{isgt} \tag{8}$$

In all regressions I also include all of the control variables from equation (3) used in the estimation of school value added, as they will likely also contribute to postsecondary outcomes. I scale the value-added estimates by the standard deviation of the estimated value added distribution, $\sigma_{\hat{\mu}_{st}}$, so that the coefficient β_μ can be interpreted as the effect of a one standard deviation increase in school value added. I cluster bootstrap the standard errors at the school level to account for the fact that $\hat{\mu}_{st}$ is a generated regressor.

Because each student's postsecondary outcomes do not vary over time but their school's value added is allowed to drift over time, the regressions may contain multiple observations for a student with identical outcome values but differing school value added. For example, a student observed in 6th, 7th, and 8th grade who enrolls in college will have three distinct middle-school value-added estimates but will have a value of 1 for enrolling in college for all of those observations. In order to assure that all students contribute equally, I weight each observation by the inverse of the number of observations per student. Thus, a student observed in 6th, 7th, and 8th grade would have a weight of $\frac{1}{3}$ for each observation while a student observed only in 7th and 8th grade would have a weight of $\frac{1}{2}$.

4.1 K–12 Outcomes

First, I explore whether school value added impacts future K–12 performance. The outcomes I examine are ELA and math test scores one grade later, whether a student enrolled in a public school one grade later, and whether a student took the most advanced math subject in future grades. Table D.4 shows that school value added persists to future test scores. In elementary school a one standard deviation increase in school value added increases ELA test scores in the next grade by 8.8% of a standard deviation. The effects for middle and high school are also similarly large. The effect sizes at all school levels are close to the effect sizes on contemporaneous scores, which contrasts with evidence of fade out in other environments (Currie and Thomas, 1995, 1999; Bitler, Hoynes and Domina, 2014), although part of this may be due to students remaining in the same school, because school value added is highly correlated one year apart (as seen in figure D.2). Interestingly, ELA value added has an even larger effect on future math scores than on future ELA scores. School value added has an economically insignificant effect on remaining in the public-school system, which combines the effect of transferring to a private school, dropout, and moving to another state. Finally, school value added has a positive impact on the math subject that students take, as students are more likely to take the hardest math subject for their final math exam in grade 11. Appendix table D.14 shows that while elementary-school math value added is more likely to track students to the most difficult math track in 7th grade, elementary-school ELA value added actually has a larger effect on whether students eventually take the most difficult math exam by the time they graduate.

4.2 Extensive Margin of College

Much more important than K–12 outcomes, however, is the effect of school value added on postsecondary enrollment, because attending college has proven to be a worthwhile investment for both the average and marginal student (Oreopoulos and Petronijevic, 2013). Hoekstra (2009) finds that attending a flagship university increases the earnings of white men by 20%, while Zimmerman (2014) shows that admission to a 4-year university for the marginal student gives a wage premium of 22% and bachelor's degree receipt for the marginal admission increases wages by 90%.

I define postsecondary enrollment as enrolling in any institution in the NSC data within one year of high-school graduation.[21] I code two-year and four-year enrollment as mutually exclusive, so if students enroll in both a two-year and four-year institution within a year of graduating high school (such as if they take a summer course at a community college) then I code them as only enrolling in a four-year institution. Table D.5 gives summary statistics of postsecondary enrollment by school level for the sample of students for whom I later estimate value added on postsecondary enrollment (in section 4.4). About two-thirds of students enroll in any postsecondary institution, and about 40% of college attendees enroll in a four-year university. The vast majority of college enrollees attend a public institution and an in-state California institution, which is not surprising given the quality and cost of the California public university systems.

Table D.6 shows the results from the regressions of postsecondary enrollment on school value added. The results show that high-school value added has the largest impact on postsecondary enrollment, as a one standard deviation increase in value added increases overall enrollment by 2.2 percentage points (3.4%). This is about 2 percentage points smaller than the effect of 11th grade value added on postsecondary enrollment

[21] I also include any student who enrolls in a CSU or a CCC within one year of high-school graduation in order to account for any missing data in the NSC data. The sample consists solely of students who could potentially be matched to the NSC data, as students who did not enroll in a CSU or CCC and could not be potentially matched to the NSC data may still have enrolled in a postsecondary institution, such as a UC, but I would not observe this.

found in Hubbard (2017). High school also has the largest impact on 4-year enrollment, with an effect size of 2.8 percentage points (10.3%). Elementary- and middle-school value added have smaller, but still positive, effects on overall and 4-year enrollment, although elementary school has a somewhat larger effect on 4-year enrollment than middle school. High value added elementary and high schools appear to induce students to enroll in a 4-year university instead of a 2-year community college, which should provide a higher wage premium (Kane and Rouse, 1995).

As a robustness check, I run horse-race regressions that include school value added from all levels of schooling for the subset of students that I observe in elementary, middle, and high school. These regressions take the form of equation 8, but instead of including the value added for a student's specific school level in the different years for which the student was enrolled in that level of school, these regressions include the student's average value added estimate for elementary, middle, and high school. I also use each student's average value of the other control variables to account for the fact that these values may change from grade to grade. I include interaction terms between the school levels in order to test whether there are benefits to attending multiple high value added schools in succession. It should be noted that this is a unique sample, because these are the students that I observe for at least five consecutive grades. For this reason, the sample size is much smaller than that from the regressions in Table D.6.

Table D.7 confirms that high-school value added consistently has the largest positive effect on post-secondary enrollment. This is likely due to the fact that high-school enrollment is so close to the college decision process, which requires a concentrated effort at a very specific point in time. I find very little evidence that there are benefits to attending multiple high value added schools in succession, as the interaction terms between school levels are an order of magnitude smaller than the direct effect of high-school value added. Furthermore, many of the interaction terms are negative, which would indicate that there are actually decreasing returns to attending multiple high value added schools.

4.3 Intensive Margin of College

Next, I explore how school value added on test scores affects CSU and CCC outcomes that are conditional upon enrollment at one of those institutions. For CSU these outcomes include acceptance (conditional on application), remediation, STEM major, undecided major, persistence, degree receipt, and STEM degree receipt. For CCC these outcomes include remediation, persistence, transfer to a four-year university, degree receipt, and associate's degree receipt. I measure degree receipt within 6 years for 4-year degrees and within 3 years for 2-year degrees. The need for remedial classes is a negative outcome, because students are paying college tuition for courses that they had the opportunity to take for free while enrolled in high school. STEM major is a positive outcome, because STEM majors earn more than any other major with the exception of business (Arcidiacono, 2004; Melguizo and Wolniak, 2012; Kinsler and Pavan, 2015) and the premium has increased over time (Gemici and Wiswall, 2014). Both 2-year and 4-year degrees are positive outcomes because they provide a wage premium for workers (Kane and Rouse, 1995).

The regressions for CSU outcomes are shown in Table D.8. Interestingly, high value added schools decrease a student's likelihood of being accepted conditional on application, although the effect is extremely small. This is likely due to increases in CSU application on the extensive margin, where students have a low likelihood of acceptance, that dominate any increases in the probability of acceptance on the intensive margin. Encouragingly, high value added elementary and middle schools also reduce a student's need for remedial classes upon enrolling at a CSU. A middle school with value added one standard deviation above average decreases the need for remedial ELA and math classes by 2.2 percentage points (9.7%) and 3.2 percentage points (13.9%) respectively. Interestingly, as seen in appendix table D.16, elementary-school ELA value added has a much larger effect on the need to take remedial math classes than elementary-school math value added, which suggests that school-induced learning on ELA exams may provide skills in other subjects, which is consistent with similar findings in Master, Loeb and Wyckoff (2017).

School value added has no effect on whether students become a STEM major, but high value added schools do reduce the likelihood that students are undecided in their first year of college. This likely focuses course enrollment and reduces frivolous classes. Elementary school has the strongest effect on whether a student persists to their second or third year year of college, with middle school also having a significant effect. A one standard deviation increase in elementary-school value added increases the likelihood of persisting to year three by 1.4 percentage points (1.9%). There is suggestive evidence that middle-school value added also increases degree receipt, although the estimates are noisy and insignificant[22]. Thus, the evidence suggests that while high school plays the largest role in whether students actually enroll in a postsecondary institution, as seen in section 4.2, elementary and middle schools develop the skills necessary for students to succeed in college.

The CCC outcomes are given in Table D.9. Elementary-school value added again reduces the need for remedial courses. A one standard deviation increase in elementary-school value added decreases the need for remedial math classes by 0.7 percentage points (3.1%). Persistence to year two at a community college is a somewhat complicated outcome because the failure to persist could be a good outcome if the student transferred to a four-year university or bad outcome if the student dropped out of college altogether. In order to avoid this issue, I code a student as persisting to year two if they persisted to year two at a community

[22]The sample size for degree receipt is small because I allow students six years to obtain a degree. This is also why I cannot examine the effect of elementary-school value added on degree receipt, because I don't have any elementary-school students who enrolled in college at least six years prior to my final year of data.

college or transferred to a four-year university. I recode degree receipt and associate's degree receipt in the same way. High school has the largest impact on both persistence to year two and transfer to a four-year university. The impact of high-school value added on transferring to a four-year university is particularly large, as a one standard deviation increase in high-school value added increases the likelihood of transferring to a four-year university after enrolling at a CCC by 3 percentage points (8.3%). Thus, high schools not only have the largest impact on initial four-year enrollment but also have the largest impact on students transferring into four-year universities. At all school levels attending a high value added school increases both degree receipt and associate's degree receipt, although this appears to be driven by transfer to a four-year university.

4.4 Value Added on Postsecondary Enrollment

The prior sections show that increases in school value added on test scores translate to college enrollment and readiness. Nevertheless, the results only show the effects of school quality that operate *through* test scores. There will likely be other factors within a school, however, that affect the likelihood that students enroll in a postsecondary institution but wouldn't affect how well students perform on standardized tests, such as college counselors or institutional knowledge on the college application process. In order to determine the contribution of these other factors within a school, I estimate a school's value added on postsecondary enrollment directly.

I do so by reestimating equation (3) with an indicator for postsecondary enrollment as the dependent variable instead of a student's test score.[23] It should be noted that the assumptions to obtain unbiased estimates of school value added on postsecondary enrollment are stronger than those for school value added on test scores. Value added on test scores relies upon the assumption that prior test scores and demographic characteristics are sufficient to predict how a student would perform on the current year's test, such that any differences in test scores after controlling for these variables are attributable to schools. Prior research shows that this is a valid assumption (Kane and Staiger, 2008; Deming, 2014).

Estimating value added on postsecondary enrollment, however, relies upon the assumption that prior test scores and demographic characteristics are sufficient to predict the likelihood that a student will attend a postsecondary institution. This assumption may not hold, especially for earlier grades where the prior test scores are many years removed from the time when a student decides whether to attend college. In fact, Abdulkadiroğlu et al. (2017) find that the bias of value-added estimates on postsecondary enrollment is larger than the bias of value-added estimates on test scores at the high-school level. Thus the results for

[23]Because each student's enrollment outcome is invariant across grades, I only use observations from 5th grade for elementary school, 8th grade for middle school, and 11th grade for high school. Results using 4th, 6th, and 9th grade are qualitatively similar.

school value added on postsecondary enrollment should be interpreted keeping these caveats in mind.

Figure D.7 shows the distributions of the estimated value added on postsecondary enrollment. As with the results using value added on test scores, we see that high school has the largest impact on postsecondary enrollment, as it has the highest variance in value added. A high school that is one standard deviation above average in the value added that it provides on postsecondary enrollment increases the postsecondary enrollment of its students by 8.7 percentage points on average. Middle school has the second largest variance in value added, while elementary school has the smallest. Thus, the closer a student gets to enrolling in college the bigger the impact the school they attend has on whether they actually end up enrolling. One notable difference between high school and the other school levels is the long, left tail of low-value added schools in high school.

5 Value Added Characteristics

5.1 School Characteristics

Finally, I explore what school characteristics are correlated with school value added. While these regressions are not causal, they provide a description of what high value added schools have in common. This analysis may therefore provide clues of some effective characteristics that could be explored in a causal framework in future studies.

I run regressions of school value added on school-level inputs as in equation (9). I cluster the standard errors at the school level. In the first regression the school characteristics included in $\boldsymbol{X}_{st}$ are the number of full-time equivalent (FTE) teachers per student, FTE pupil services staff[24] per student, English-learner staff per student, proportion teachers with three years or less experience, proportion teachers with full credentials, proportion male teachers, proportion male students, and the interaction between the two, and proportion minority[25] teachers, proportion minority students, and the interaction between the two. In the second regression I include district expenditure data on instruction, pupil services (counselors, nurses, food service, etc.), ancillary services (before- and after-school programs), and general administration expenditures. I also include total enrollment to account for fixed costs. In each regression I drop the top and bottom 2.5% of each independent variable in order to account for outliers and potential errors in the data that schools report.

$$\hat{\mu}_{st} = \boldsymbol{X}_{st}\boldsymbol{\beta} + \varepsilon_{st} \tag{9}$$

[24]This includes counselors, psychologists, librarian/library/media teachers, social workers, nurses, and speech/language/hearing specialists.

[25]Hispanic, black, Native American and two or more races.

Table D.10 shows the correlations between school value added and school characteristics. The left three columns give value added on test scores, while the right three columns give value added on postsecondary enrollment. There is no clear pattern of school characteristics that positively impact school value added at all levels. The coefficients often switch signs between school levels and rarely have similar magnitudes. There are a few examples of consistency between two adjacent school levels, however. English learner staff tend to increase value added in middle and high school, which may give evidence that these resources help Hispanic and Asian students who struggle with English. Fully-credentialed teachers increase elementary and middle-school value added, which contrasts with prior studies that show that teacher credentials have no effect on teacher value added (Kane, Rockoff and Staiger, 2008). Having more minority teachers appears to be beneficial when there are more minority students enrolled in a school, which suggests that minorities may benefit from having teachers similar to them.

Table D.11 gives correlations between school value added and district expenditures. As with other school characteristics, few patterns emerge. The results suggest that instruction expenditures have essentially no effect on the value added of the school, while pupil services expenditures may in fact have a negative effect on school value added. Expenditures on ancillary services in elementary school, however, are strongly correlated with value added on college-going, which suggests that after school programs in a student's earliest years may have long-lasting effects. In addition, general administration expenditures have a small, but consistently positive, effect on value added for both test scores and postsecondary enrollment.

5.2 Spatial Correlations

While I find that few school characteristics are consistently correlated with the value added that a school provides, it is possible that there is spatial correlation in school value added. Figure D.8 shows the average school test score value added within each zip code in California. While there is variation in school value added across the state, a broad pattern emerges in all school levels. In general, high value added schools tend to be clustered in the dense urban and suburban areas around Los Angeles, the Bay Area, Sacramento, and San Diego, while low value added schools tend to be located in the rural regions of the Central Valley and the Inland Empire. Exceptions to this include inner city Los Angeles in the areas around Compton.

Figure D.9 shows the same information for school value added on postsecondary enrollment. Here a similar pattern emerges, where the schools that increase the likelihood that their students attend a postsecondary institution the most are located near big cities while the low value added schools tend to be located in rural areas. The concentration of high value added schools in densely populated areas appears to be even stronger for value added on postsecondary enrollment than it does for value added on test scores. Interestingly, while

Los Angeles outperforms the Bay Area on test score value added, the Bay Area outperforms Los Angeles on postsecondary enrollment value added.

6 Conclusion

Human-capital formation is a lifelong process, but because later investments build off of earlier investments the human capital accrued during childhood and adolescence may be particularly important. This paper studies the impact of school quality on human-capital formation during these time periods. I estimate school quality in elementary, middle, and high school using individual-level data on the universe of public-school students in California. I measure school quality by extending the value added with drift methodology, as in Chetty, Friedman and Rockoff (2014*a*), to schools. I find that there is substantial variation in value added across schools, with the standard deviations of school value added ranging from 6.6% to 8.7% of a student test score standard deviation depending on the school level.

I then link these school value-added estimates to individual-level postsecondary enrollment data from the NSC and individual-level application, enrollment, and degree receipt data from the CSU and CCC systems in order to study the impact of school value added on postsecondary outcomes. I find that high-school value added has the largest effect on postsecondary enrollment, while elementary and middle-school value added have the largest effect on college readiness. All school levels therefore contribute to human-capital formation, but the different school levels contribute to different aspects of human-capital formation. Early education provides the skills necessary to succeed later in college, while high-school quality likely has a large impact on postsecondary enrollment due to its proximity to the college decision process. To my knowledge this paper is the first to compare the effect of school quality on long-run outcomes across elementary, middle, and high school.

There are numerous policy implications from my work. The first regards the measurement of school quality. I find that value-added estimates are a valuable tool for measuring school quality, as they predict long-run outcomes but are uncorrelated with prior student ability. Average test scores should be avoided when measuring school quality though, because differences in average test scores are largely due to the selection of students to schools. However, value added on long-run outcomes should also be used when possible, because value added on test scores may mask relevant differences in value added on important long-run outcomes.

Second, I find that early childhood education has important long-run consequences. Differences in school quality as early as elementary school affect students' college readiness. This is consistent with evidence that finds that early childhood education programs can improve the long-run outcomes of students.

Lastly, I find that high-school quality is unlikely to undo the effects of low-quality schools in prior years when it comes to college readiness. High-school quality has essentially no effect on the college readiness of students, while both elementary- and middle-school quality have substantial effects. This should be considered in the wake of college-going interventions that take place in high school. These efforts may need to be accompanied with academic support, as the students may not succeed in college after enrolling without additional assistance.

While this paper shows that high value added schools have long-term effects on postsecondary outcomes, the question remains as to what comprises a high value added school. Prior research on school and teacher characteristics has been largely inconclusive as to what makes an effective school or teacher, and the correlational results that I present in this paper do not shed much light on the issue. Further research is needed in order to identify the replicable characteristics of high value added schools, as the evidence shows that these schools can permanently improve the lives of their students.

References

Abdulkadiroğlu, Atila, Joshua D Angrist, Susan M Dynarski, Thomas J Kane, and Parag A Pathak. 2011. "Accountability and flexibility in public schools: Evidence from Boston's charters and pilots." *The Quarterly Journal of Economics*, 126(2): 699–748.

Abdulkadiroğlu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters. 2017. "Do Parents Value School Effectiveness?" National Bureau of Economic Research.

Altonji, Joseph G. 1995. "The Effects of High School Curriculum on Education and Labor Market Outcomes." *Journal of Human Resources*, 409–438.

Angrist, Joshua D, Peter D Hull, Parag A Pathak, and Christopher R Walters. 2017. "Leveraging lotteries for school value-added: Testing and estimation." *The Quarterly Journal of Economics*, 132(2): 871–919.

Angrist, Joshua D, Sarah R Cohodes, Susan M Dynarski, Parag A Pathak, and Christopher R Walters. 2016. "Stand and deliver: Effects of Boston's charter high schools on college preparation, entry, and choice." *Journal of Labor Economics*, 34(2): 275–318.

Arcidiacono, Peter. 2004. "Ability sorting and the returns to college major." *Journal of Econometrics*, 121(1-2): 343–375.

Bacher-Hicks, Andrew, Thomas J Kane, and Douglas O Staiger. 2014. "Validating teacher effect estimates using changes in teacher assignments in Los Angeles." National Bureau of Economic Research.

Barr, Andrew, and Ben Castleman. 2019. "Exploring Variation in College Counselor Effectiveness." Vol. 109, 227–31.

Béteille, Tara, Demetra Kalogrides, and Susanna Loeb. 2012. "Stepping stones: Principal career paths and school outcomes." *Social Science Research*, 41(4): 904–919.

Bitler, Marianne P, Hilary W Hoynes, and Thurston Domina. 2014. "Experimental evidence on distributional effects of Head Start." National Bureau of Economic Research.

Bitler, Marianne P, Sean P Corcoran, Thurston Domina, and Emily K Penner. 2014. "Teacher Effects on Student Achievement and Height: A Cautionary Tale." *Society for Research on Educational Effectiveness.*

Bitler, Marianne, Sean Corcoran, Thurston Domina, and Emily Penner. 2019. "Teacher effects on student achievement and height: A cautionary tale." National Bureau of Economic Research.

Branch, Gregory F, Eric A Hanushek, and Steven G Rivkin. 2012. "Estimating the effect of leaders on public sector productivity: The case of school principals." National Bureau of Economic Research.

Carrell, Scott E, and Bruce Sacerdote. 2013. *Late interventions matter too: The case of college coaching New Hampshire.* Citeseer.

Carrell, Scott E, and James E West. 2010. "Does professor quality matter? Evidence from random assignment of students to professors." *Journal of Political Economy*, 118(3): 409–432.

Carrell, Scott E, and Mark Hoekstra. 2014. "Are school counselors an effective education input?" *Economics Letters*, 125(1): 66–69.

Carrell, Scott E, and Susan A Carrell. 2006. "Do lower student to counselor ratios reduce school disciplinary problems?" *The BE Journal of Economic Analysis & Policy*, 5(1).

Carrell, Scott E, Marianne E Page, and James E West. 2010. "Sex and science: How professor gender perpetuates the gender gap." *The Quarterly Journal of Economics*, 125(3): 1101–1144.

Castleman, Benjamin, and Joshua Goodman. 2018. "Intensive College Counseling and the Enrollment and Persistence of Low-Income Students." *Education Finance and Policy*, 13(1): 19–41.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014*a*. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *The American Economic Review*, 104(9): 2593–2632.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014*b*. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *The American Economic Review*, 104(9): 2633–2679.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2017. "Measuring the Impacts of Teachers: Reply." *American Economic Review*, 107(6): 1685–1717.

Clark, Damon, Paco Martorell, and Jonah Rockoff. 2009. "School Principals and School Performance. Working Paper 38." *National Center for Analysis of longitudinal data in Education research.*

Corcoran, Sean P. 2010. "Can Teachers Be Evaluated by Their Students' Test Scores? Should They Be? The Use of Value-Added Measures of Teacher Effectiveness in Policy and Practice. Education Policy for Action Series." *Annenberg Institute for School Reform at Brown University (NJ1).*

Cunha, Flavio, James J Heckman, Lance Lochner, and Dimitriy V Masterov. 2006. "Interpreting the evidence on life cycle skill formation." *Handbook of the Economics of Education*, 1: 697–812.

Currie, Janet, and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *American Economic Review*, 85(3): 341–64.

Currie, Janet, and Duncan Thomas. 1999. "Does head start help Hispanic children?" *Journal of public Economics*, 74(2): 235–262.

Deming, David. 2009. "Early childhood intervention and life-cycle skill development: Evidence from Head Start." *American Economic Journal: Applied Economics*, 1(3): 111–34.

Deming, David J. 2014. "Using school choice lotteries to test measures of school effectiveness." *The American Economic Review*, 104(5): 406–411.

Deming, David J, Justine S Hastings, Thomas J Kane, and Douglas O Staiger. 2014. "School choice, school quality, and postsecondary attainment." *American Economic Review*, 104(3): 991–1013.

De Vlieger, Pieter, Brian Jacob, and Kevin Stange. 2017. "Measuring Instructor Effectiveness in Higher Education." In *Productivity in Higher Education.* University of Chicago Press.

Dobbie, Will, and Roland G Fryer Jr. 2011. "Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics*, 3(3): 158–87.

Dobbie, Will, and Roland G Fryer Jr. 2015. "The medium-term impacts of high-achieving charter schools." *Journal of Political Economy*, 123(5): 985–1037.

Dobbie, Will S, and Roland G Fryer Jr. 2016. "Charter schools and labor market outcomes." National Bureau of Economic Research.

Doyle, Orla, Colm P Harmon, James J Heckman, and Richard E Tremblay. 2009. "Investing in early human development: timing and economic efficiency." *Economics & Human Biology*, 7(1): 1–6.

Duncan, Greg J, and Katherine Magnuson. 2013. "Investing in preschool programs." *Journal of Economic Perspectives*, 27(2): 109–32.

Everson, Kimberlee C. 2017. "Value-Added Modeling and Educational Accountability: Are We Answering the Real Questions?" *Review of Educational Research*, 87(1): 35–70.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-term effects of Head Start." *American economic review*, 92(4): 999–1012.

Gates, Susan M, Laura S Hamilton, Paco Martorell, Susan Burkhauser, Paul Heaton, Ashley Pierson, Matthew D Baird, Mirka Vuollo, Jennifer J Li, Diana Lavery, et al. 2014. "Principal Preparation Matters: How Leadership Development Affects Student Achievement. Brief." *RAND Corporation*.

Gemici, Ahu, and Matthew Wiswall. 2014. "Evolution of gender differences in post-secondary human capital investments: College Majors." *International Economic Review*, 55(1): 23–56.

Grissom, Jason A, and Susanna Loeb. 2011. "Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills." *American Educational Research Journal*, 48(5): 1091–1123.

Grissom, Jason A, Richard SL Blissett, and Hajime Mitani. 2018. "Evaluating school principals: Supervisor ratings of principal practice and principal job performance." *Educational Evaluation and Policy Analysis*, 40(3): 446–472.

Guarino, Cassandra M, Michelle Maxfield, Mark D Reckase, Paul N Thompson, and Jeffrey M Wooldridge. 2015. "An evaluation of empirical Bayes's estimation of value-added teacher performance measures." *Journal of Educational and Behavioral Statistics*, 40(2): 190–222.

Hanushek, Eric A. 1989. "The impact of differential expenditures on school performance." *Educational researcher*, 18(4): 45–62.

Hanushek, Eric A. 1997. "Assessing the effects of school resources on student performance: An update." *Educational evaluation and policy analysis*, 19(2): 141–164.

Hanushek, Eric A. 2003. "The failure of input-based schooling policies." *The economic journal*, 113(485): F64–F98.

Hanushek, Eric A. 2011. "The economic value of higher teacher quality." *Economics of Education review*, 30(3): 466–479.

Hanushek, Eric A, and Steven G Rivkin. 2006. "Teacher quality." *Handbook of the Economics of Education*, 2: 1051–1078.

Hanushek, Eric A, and Steven G Rivkin. 2010. "Generalizations about using value-added measures of teacher quality." *American Economic Review*, 100(2): 267–71.

Hanushek, Eric A, John F Kain, Daniel M O'Brien, and Steven G Rivkin. 2005. "The market for teacher quality." National Bureau of Economic Research.

Hanushek, Eric A, Steven G Rivkin, and Lori L Taylor. 1996. "Aggregation and the Estimated Effects of School Resources." *The Review of Economics and Statistics*, 611–627.

Heckman, James J. 2006. "Skill formation and the economics of investing in disadvantaged children." *Science*, 312(5782): 1900–1902.

Heckman, J.J., A.B. Krueger, and B.M. Friedman. 2003. *Inequality in America: What Role for Human Capital Policies? Alvin Hansen Symposium Series on Public Policy Series*, MIT Press.

Hoekstra, Mark. 2009. "The effect of attending the flagship state university on earnings: A discontinuity-based approach." *The Review of Economics and Statistics*, 91(4): 717–724.

Horng, Eileen Lai, Daniel Klasik, and Susanna Loeb. 2010. "Principal's time use and school effectiveness." *American Journal of Education*, 116(4): 491–523.

Hubbard, Daniel. 2017. "More Gains Than Score Gains? High School Quality and College Success."

Ishii, Jun, and Steven G Rivkin. 2009. "Impediments to the estimation of teacher value added." *Education finance and Policy*, 4(4): 520–536.

Jacob, Brian A, and Lars Lefgren. 2005. "Principals as agents: Subjective performance measurement in education." National Bureau of Economic Research.

Jacob, Brian A, Lars Lefgren, and David P Sims. 2010. "The persistence of teacher-induced learning." *Journal of Human resources*, 45(4): 915–943.

Johnson, Rucker C, and C Kirabo Jackson. 2017. "Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending." National Bureau of Economic Research.

Kane, Thomas J, and Cecilia Elena Rouse. 1995. "Labor-market returns to two-and four-year college." *The American Economic Review*, 85(3): 600–614.

Kane, Thomas J, and Douglas O Staiger. 2008. "Estimating teacher impacts on student achievement: An experimental evaluation." National Bureau of Economic Research.

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(6): 615–631.

Kinsler, Josh. 2012. "Assessing Rothstein's critique of teacher value-added models." *Quantitative Economics*, 3(2): 333–362.

Kinsler, Josh, and Ronni Pavan. 2015. "The specificity of general human capital: Evidence from college major choice." *Journal of Labor Economics*, 33(4): 933–972.

Ladd, Helen F. 2011. "Teachers' perceptions of their working conditions: How predictive of planned and actual teacher movement?" *Educational Evaluation and Policy Analysis*, 33(2): 235–261.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach. 2018. "School finance reform and the distribution of student achievement." *American Economic Journal: Applied Economics*, 10(2): 1–26.

Lankford, Hamilton, Susanna Loeb, and James Wyckoff. 2002. "Teacher sorting and the plight of urban schools: A descriptive analysis." *Educational evaluation and policy analysis*, 24(1): 37–62.

Loeb, Susanna, Demetra Kalogrides, and Eileen Lai Horng. 2010. "Principal preferences and the uneven distribution of principals across schools." *Educational Evaluation and Policy Analysis*, 32(2): 205–229.

Loeb, Susanna, Demetra Kalogrides, and Tara Béteille. 2012. "Effective schools: Teacher hiring, assignment, development, and retention." *Education Finance and Policy*, 7(3): 269–304.

Ludwig, Jens, and Douglas L Miller. 2007. "Does Head Start improve children's life chances? Evidence from a regression discontinuity design." *The Quarterly journal of economics*, 122(1): 159–208.

Martorell, Paco, Kevin Stange, and Isaac McFarlin Jr. 2016. "Investing in schools: capital spending, facility conditions, and student achievement." *Journal of Public Economics*, 140: 13–29.

Master, Benjamin, Susanna Loeb, and James Wyckoff. 2017. "More Than Content: The Persistent Cross-Subject Effects of English Language Arts Teachers' Instruction." *Educational Evaluation and Policy Analysis*, 39(3): 429–447.

McClellan, Mark, and Douglas Staiger. 1999. "The quality of health care providers." National bureau of economic research.

McClellan, Mark, and Douglas Staiger. 2000. "Comparing the quality of health care providers." In *Frontiers in Health Policy Research, Volume 3*. 113–136. MIT Press.

Melguizo, Tatiana, and Gregory C Wolniak. 2012. "The earnings benefits of majoring in STEM fields among high achieving minority students." *Research in Higher Education*, 53(4): 383–405.

Meyer, Robert H. 1997. "Value-added indicators of school performance: A primer." *Economics of Education Review*, 16(3): 283–301.

Oreopoulos, Philip, and Uros Petronijevic. 2013. "Making college worth it: A review of research on the returns to higher education." National Bureau of Economic Research.

Pop-Eleches, Cristian, and Miguel Urquiola. 2013. "Going to a better school: Effects and behavioral responses." *American Economic Review*, 103(4): 1289–1324.

Reback, Randall. 2010. "Noninstructional spending improves noncognitive outcomes: Discontinuity evidence from a unique elementary school counselor financing system." *Education Finance and Policy*, 5(2): 105–137.

Rivkin, Steven G, Eric A Hanushek, and John F Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica*, 73(2): 417–458.

Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review*, 94(2): 247–252.

Rothstein, Jesse. 2009. "Student sorting and bias in value-added estimation: Selection on observables and unobservables." *Education*, 4(4): 537–571.

Rothstein, Jesse. 2010. "Teacher quality in educational production: Tracking, decay, and student achievement." *The Quarterly Journal of Economics*, 125(1): 175–214.

Rothstein, Jesse. 2017. "Measuring the Impacts of Teachers: Comment." *American Economic Review*, 107(6): 1656–84.

Schochet, Peter Z, John Burghardt, and Sheena McConnell. 2008. "Does job corps work? Impact findings from the national job corps study." *American Economic Review*, 98(5): 1864–86.

Staiger, Douglas O, and Thomas J Kane. 2014. "Making decisions with imprecise performance measures: The relationship between annual student achievement gains and a teacher's career value added." *Designing teacher evaluation systems: New guidance from the measures of effective teaching project.*

Willms, J Douglas, and Stephen W Raudenbush. 1989. "A longitudinal hierarchical linear model for estimating school effects and their stability." *Journal of educational measurement*, 26(3): 209–232.

Yu, Han, and Naci Mocan. 2018. "The Impact of High School Curriculum on Confidence, Academic Success, and Mental and Physical Well-Being of University Students." National Bureau of Economic Research.

Zimmerman, Seth D. 2014. "The returns to college admission for academically marginal students." *Journal of Labor Economics*, 32(4): 711–754.

A Data

Table D.12 gives the number of observations in the CST data conditional on a set of restrictions implemented in order to form the value-added sample. The rows are additive, such that the first row contains all observations, the second row imposes one restriction, the third row imposes two restrictions, etc. The first row denotes the total number of observations in the CST dataset. The second row keeps students who have information on test scores, as opposed to just demographic characteristics. The third row keeps only the first time that a student attempted a grade, and thus drops observations in which a student is repeating a grade. I impose this restriction because students repeating a grade are tested on material for which they have already been tested at least once. The fourth row keeps only students at "conventional" schools. This includes schools in the following categories defined by the CDE: Preschool, Elementary School (Public), Elementary School in 1 School District (Public), Intermediate/Middle Schools (Public), Junior High Schools (Public), K–12 Schools (Public), High Schools (Public), and High Schools in 1 School District (Public).[26]

[26] This drops students in the following categories: Special Education Schools (Public), County Community, Youth Authority Facilities (CEA), Opportunity Schools, Juvenile Court Schools, Other County or District Programs, State Special Schools, Alternative Schools of Choice, Continuation High Schools, District Community Day Schools, Adult Education Centers, and Regional Occupational Center/Program (ROC/P).

The fifth row drops any schools that enroll 10 students or fewer in a given year. The sixth row drops students who are missing a test score in the specific subject for which value added is calculated. The seventh row drops students who are missing any of the demographic controls. The eighth and ninth rows drop students who are missing test scores from one grade and two grades prior, respectively. This restriction is the cause of the vast majority of observations which are excluded from the value-added estimates. The tenth row drops observations for which peer averages of the control variables could not be calculated. The eleventh row drops students if fewer than seven observations can be used to estimate value added for their school by year cell, which insures that all value-added estimates are based on at least seven observations.

Table D.13 gives summary statistics for the students who are excluded from the value-added sample. For comparison, the table also includes the summary statistics of the value-added sample. Excluded students are more likely to be male and slightly more likely to be economically disadvantaged, black, two or more races, or American Indian. The most stark difference between included and excluded students, however, is in their likelihood of having a disability. Excluded students are over four times more likely to be disabled in elementary school and over twice as likely to be disabled in high school. This fact carries over to prior test scores, as excluded students are more likely to have lower prior achievement than students in the value-added sample.

B CST to Postsecondary Match

Because we lack a unique student identifier common to both the CST and CSU/CCC data, such as a social security number, I match the CST data to the CSU/CCC data based on students' name, birth date, and sex. While there is no common unique student identifier, each dataset does have a unique student identifier specific to that dataset, which I will call the CST ID, CSU ID, and CCC ID, respectively. The match is implemented as a sieve, with progressively less strict matches in each sieve level.

For the CCC data, I match on first name, last name, birth date, and sex. I start by dropping all students in the CST and CCC datasets that are not uniquely identified by these variables, as for all intents and purposes they constitute the student's unique identifier. Denote the remaining observations as the *master CST* and *master CCC* datasets, respectively. I then match the master CST and master CCC datasets on first name, last name, birth date, and sex. Of those matched observations, I drop any observations that were missing data on any of the match variables. I then drop any students for whom their CST ID matched to multiple CCC IDs or their CCC ID matched to multiple CST IDs. The remaining matched observations I denote *sieve level 1*, which contains one observation per CST ID and CCC ID.

Next I remove all of the sieve level 1 observations from the master CST and master CCC datasets and repeat the steps above matching on first name, last name, and birth date. I denote this sieve level 2. For sieve level 3 I match on first three letters of first name, first three letters of last name, birth date, and sex. This is due to the fact that from 1993 to 2011 the CCC data only contains the first three letters of a student's first and last name. Finally for sieve level 4 I match on first three letters of first name, first three letters of last name, and birth date.

For the CSU data I implement the same matching procedure. Sieve level 1 matches on first name, last name, birth date, sex, and middle name. Sieve level 2 matches on first name, last name, birth date, and sex. Sieve level 3 matches on first name, last name, and birth date.

The NSC data was matched by the NSC using their proprietary match process. This process relies on a student's first name, middle initial, last name, and birth date.

Because of data limitations, the number of cohorts than can be matched to the postsecondary data varies by school level. Given the years available for the CST and CSU/CCC data, 6 cohorts from the elementary-school value-added sample, 9 cohorts from the middle-school value-added sample, and 11 cohorts from the high-school value-added sample could potentially be matched to the CSU/CCC data. For the NSC data 6 cohorts from the elementary-school value-added sample, 8 cohorts from the middle-school value-added sample, and 7 cohorts from the high-school value-added sample could potentially be matched to the NSC data.

C School Value Added Methodology

I follow the methodology described in Chetty, Friedman and Rockoff (2014*a*), implementing a few modifications in order to estimate school-by-year value added instead of teacher-by-year value added. A school's value added is calculated by first removing the portion of each student's test score that is due to non-school factors. To do so, I regress student test scores z_{isgt} on a vector of prior test scores $\boldsymbol{z}_{ig}$, demographic characteristics $\boldsymbol{X}_{it}$, the number of students in a student's cohort $\boldsymbol{W}_{sgt}$, grade fixed effects γ_g, year fixed effects ψ_t, and a school fixed effect α_s as in equation (10).

The vector of prior test scores $\boldsymbol{z}_{ig}$ contains a cubic polynomial in one-grade-prior same-subject test score and a cubic polynomial in two-grade-prior same-subject test score. I allow the polynomials in prior scores to differ by grade by interacting the polynomials with grade fixed effects. The demographic characteristics $\boldsymbol{X}_{it}$ contain a linear term for age and fixed effects for sex, ethnicity, socioeconomic disadvantage, limited English proficiency, and disability status.

$$z_{isgt} = \boldsymbol{z}_{ig}\boldsymbol{\delta}_g + \boldsymbol{X}_{it}\boldsymbol{\beta}_X + \boldsymbol{W}_{sgt}\boldsymbol{\beta}_W + \gamma_g + \psi_t + \alpha_s \tag{10}$$

I then calculate residual test scores r_{isgt}, as shown in equation (11), that contain the component of student test scores that does not project onto observable student characteristics. Notice that while equation (10) contains a school fixed effect α_s, the residual test scores do not subtract off this predicted school fixed effect $\hat{\alpha}_s$.

Equation (10) contains a school fixed effect in order to account for potential correlation between school value added and student characteristics. If school value added is in fact correlated with the types of students that enroll in the school, then the regression coefficients in equation (10) would be biased in the absence of a school fixed effect because the omitted variable of school quality would be correlated with both the dependent and independent variables. The residual r_{isgt} does not subtract the predicted school fixed effect, however, because doing so would leave us with a residual test score that no longer contained school value added.

$$r_{isgt} = z_{isgt} - (\boldsymbol{z}_{ig}\hat{\boldsymbol{\delta}}_g + \boldsymbol{X}_{it}\hat{\boldsymbol{\beta}}_X + \boldsymbol{W}_{sgt}\hat{\boldsymbol{\beta}}_W + \hat{\gamma}_g + \hat{\psi}_t) \tag{11}$$

It is helpful to decompose r_{isgt} into its corresponding components. Equation (12) shows that residual test scores are composed of school value added μ_{st}, common shocks θ_{st}, and an individual level error term ε_{isgt}.

$$r_{isgt} = \mu_{st} + \theta_{st} + \varepsilon_{isgt} \tag{12}$$

I then average residual test scores r_{isgt} to the school-by-year level as in equation (13). Substituting (12) into equation (13) gives us equation (14), and equation (15) follows from the fact that μ_{st} and θ_{st} are constant conditional on school and year.

$$\begin{aligned} \bar{r}_{st} &= \frac{\sum_{i\in st} r_{isgt}}{N_{st}} && (13)\\ &= \frac{\sum_{i\in st}(\mu_{st} + \theta_{st} + \varepsilon_{isgt})}{N_{st}} && (14)\\ &= \mu_{st} + \theta_{st} + \frac{\sum_{i\in st} \varepsilon_{isgt}}{N_{st}} && (15) \end{aligned}$$

Under the assumptions that ε_{isgt} is a mean zero random error term and students do not sort to schools in each year on unobservable characteristics, the student-level error terms have expected value zero conditional on school and year, which gives us equation 17.[27] Thus the expected value of the average residual test score

[27] This is due to the fact that $\mathbf{E}[\varepsilon_{isgt}|s,t] = \mathbf{E}[\varepsilon_{isgt}]$ under the assumption that students do not sort to schools in each year based on ε_{isgt}, and $\mathbf{E}[\varepsilon_{isgt}] = 0$ under the assumption that ε_{isgt} is mean zero.

conditional on school and year is equal to school value added plus the common shock.

$$\mathbf{E}[\bar{r}_{st}|s,t] = \mu_{st} + \theta_{st} + \frac{\sum_{i \in st} \mathbf{E}[\varepsilon_{isgt}|s,t]}{N_{st}} \tag{16}$$

$$= \mu_{st} + \theta_{st} \tag{17}$$

The issue is that there is no variation in the common shocks θ_{st} within each school-by-year cell, so the average residual test score contains both true school value added as well as a common shock that is unrelated to school quality. Under the assumptions that school value added is correlated across years, the school-level common shocks are uncorrelated across years, and the school-level common shocks are not correlated with school value added across years, however, we can project the average residual in year t onto the average residuals in all other years from the same school in order to eliminate variation from the common shocks.

Formally, if the assumptions in (18) hold,

$$\mathrm{cov}(\mu_{st}\mu_{st'}) \neq 0, \qquad \mathrm{cov}(\theta_{st}\theta_{st'}) = 0, \qquad \mathrm{cov}(\mu_{st}\theta_{st'}) = 0 \qquad \forall t' \neq t \tag{18}$$

then we can project $\bar{r}_{st}$ onto $\bar{\boldsymbol{r}}_{st'}$ following equation (19), where $\bar{\boldsymbol{r}}_{st'}$ is the vector of average residuals $\bar{r}_{st'}$ $\forall t' \neq t$, to recover variation in μ_{st}.

$$\bar{r}_{st} = \bar{\boldsymbol{r}}_{st'} \boldsymbol{\beta}_{\bar{r}t'} + \epsilon_{st} \tag{19}$$

If we have years $t = 1, \ldots, T$ then

$$\bar{\boldsymbol{r}}_{st'} = \begin{bmatrix} \bar{r}_{s1} & \bar{r}_{s2} & \ldots & \bar{r}_{st-1} & \bar{r}_{st+1} & \ldots & \bar{r}_{sT-1} & \bar{r}_{sT} \end{bmatrix} \tag{20}$$

and

$$\boldsymbol{\beta}_{\bar{r}t'} = \begin{bmatrix} \beta_{\bar{r}1} & \beta_{\bar{r}2} & \ldots & \beta_{\bar{r}t-1} & \beta_{\bar{r}t+1} & \ldots & \beta_{\bar{r}T-1} & \beta_{\bar{r}T} \end{bmatrix}' \tag{21}$$

However, there is a fundamental tradeoff between the number of independent variables and the number of observations that can be included in equation (19). For example, if the average residual from 5 years prior to year t is used, then a regression following (19) can only include schools that have at least 6 consecutive years of data. Some schools that close or open during the span of the dataset won't have this many observations, so they will be dropped from the regression. Thus, while including the average residual from 5 years prior to year t will increase the information that can be used to identify μ_{st}, it will decrease the number of observations used to identify $\beta_{\bar{r}t-1}$ to the subset of schools that have been open for at least 6 consecutive years. This subset will contain fewer observations than the number of observations that have a valid $\bar{r}_{st-1}$, as this only requires having 2 consecutive years of data. Thus, including $\bar{r}_{st-5}$ would essentially discard useful information in identifying $\boldsymbol{\beta}_{\bar{r}t-1}$ in order to identify $\boldsymbol{\beta}_{\bar{r}t-5}$.

Examining the coefficient vector $\boldsymbol{\beta}_{\bar{r}t'}$ is helpful in dealing with this issue. Solving equation (19) using ordinary least squares (OLS) provides us with the solution vector in (22).

$$\hat{\boldsymbol{\beta}}_{\bar{r}t'} = (\bar{\boldsymbol{r}}'_{st'}\bar{\boldsymbol{r}}_{st'})^{-1}(\bar{\boldsymbol{r}}'_{st'}\bar{\boldsymbol{r}}_{st}) \tag{22}$$

First let's examine $(\bar{\boldsymbol{r}}'_{st'}\bar{\boldsymbol{r}}_{st'})^{-1}$. The resulting matrix is equal to (23).

$$\begin{bmatrix}
\bar{\boldsymbol{r}}'_{s1}\bar{\boldsymbol{r}}_{s1} & \bar{\boldsymbol{r}}'_{s1}\bar{\boldsymbol{r}}_{s2} & \cdots & \bar{\boldsymbol{r}}'_{s1}\bar{\boldsymbol{r}}_{st-1} & \bar{\boldsymbol{r}}'_{s1}\bar{\boldsymbol{r}}_{st+1} & \cdots & \bar{\boldsymbol{r}}'_{s1}\bar{\boldsymbol{r}}_{sT-1} & \bar{\boldsymbol{r}}'_{s1}\bar{\boldsymbol{r}}_{sT} \\
\bar{\boldsymbol{r}}'_{s2}\bar{\boldsymbol{r}}_{s1} & \bar{\boldsymbol{r}}'_{s2}\bar{\boldsymbol{r}}_{s2} & \cdots & \bar{\boldsymbol{r}}'_{s2}\bar{\boldsymbol{r}}_{st-1} & \bar{\boldsymbol{r}}'_{s2}\bar{\boldsymbol{r}}_{st+1} & \cdots & \bar{\boldsymbol{r}}'_{s2}\bar{\boldsymbol{r}}_{sT-1} & \bar{\boldsymbol{r}}'_{s2}\bar{\boldsymbol{r}}_{sT} \\
\vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\
\bar{\boldsymbol{r}}'_{st-1}\bar{\boldsymbol{r}}_{s1} & \bar{\boldsymbol{r}}'_{st-1}\bar{\boldsymbol{r}}_{s2} & \cdots & \ddots & & \cdots & \bar{\boldsymbol{r}}'_{st-1}\bar{\boldsymbol{r}}_{sT-1} & \bar{\boldsymbol{r}}'_{st-1}\bar{\boldsymbol{r}}_{sT} \\
\bar{\boldsymbol{r}}'_{st+1}\bar{\boldsymbol{r}}_{s1} & \bar{\boldsymbol{r}}'_{st+1}\bar{\boldsymbol{r}}_{s2} & \cdots & & \ddots & \cdots & \bar{\boldsymbol{r}}'_{st+1}\bar{\boldsymbol{r}}_{sT-1} & \bar{\boldsymbol{r}}'_{st+1}\bar{\boldsymbol{r}}_{sT} \\
\vdots & \vdots & & \vdots & \vdots & \ddots & \vdots & \vdots \\
\bar{\boldsymbol{r}}'_{sT-1}\bar{\boldsymbol{r}}_{s1} & \bar{\boldsymbol{r}}'_{sT-1}\bar{\boldsymbol{r}}_{s2} & \cdots & \bar{\boldsymbol{r}}'_{sT-1}\bar{\boldsymbol{r}}_{st-1} & \bar{\boldsymbol{r}}'_{sT-1}\bar{\boldsymbol{r}}_{st+1} & \cdots & \bar{\boldsymbol{r}}'_{sT-1}\bar{\boldsymbol{r}}_{sT-1} & \bar{\boldsymbol{r}}'_{sT-1}\bar{\boldsymbol{r}}_{sT} \\
\bar{\boldsymbol{r}}'_{sT}\bar{\boldsymbol{r}}_{s1} & \bar{\boldsymbol{r}}'_{sT}\bar{\boldsymbol{r}}_{s2} & \cdots & \bar{\boldsymbol{r}}'_{sT}\bar{\boldsymbol{r}}_{st-1} & \bar{\boldsymbol{r}}'_{sT}\bar{\boldsymbol{r}}_{st+1} & \cdots & \bar{\boldsymbol{r}}'_{sT}\bar{\boldsymbol{r}}_{sT-1} & \bar{\boldsymbol{r}}'_{sT}\bar{\boldsymbol{r}}_{sT}
\end{bmatrix}^{-1} \tag{23}$$

Under the stationarity assumptions in (24),

$$\mathbf{E}[\mu_{st}|t] = 0, \qquad \mathbf{E}[\theta_{st} + \varepsilon_{isgt}|t] = 0, \qquad \text{cov}(\mu_{st}\mu_{st+y}) = \sigma_{\mu y} \tag{24}$$

we have $\mathbf{E}[\bar{r}_{st}|t] = 0$, which therefore allows us to write the sums of squares $\bar{\boldsymbol{r}}'_{st}\bar{\boldsymbol{r}}_{st}$ as variances $\sigma^2_{\bar{r}_{st}}$ and the cross products $\bar{\boldsymbol{r}}'_{st}\bar{\boldsymbol{r}}_{st'}$ as covariances $cov(\bar{r}_{st}, \bar{r}_{st'})$. Furthermore, because we assume that the covariance between the average residual from any two years only depends on the number of years between them, we can rewrite $\sigma^2_{\bar{r}_{st}} = \sigma^2_{\bar{r}}$ and $cov(\bar{r}_{st}, \bar{r}_{st'}) = \sigma_{\bar{r}y}$, where $y \equiv |t - t'|$ indexes the number of years between t and

t'. This simplifies the matrix in (23) further to (25).

$$\begin{bmatrix} \sigma^2_{\bar{r}} & \sigma_{\bar{r}1} & \cdots & \sigma_{\bar{r}t-2} & \sigma_{\bar{r}t} & \cdots & \sigma_{\bar{r}T-2} & \sigma_{\bar{r}T-1} \\ \sigma_{\bar{r}1} & \sigma^2_{\bar{r}} & \cdots & \sigma_{\bar{r}t-3} & \sigma_{\bar{r}t-1} & \cdots & \sigma_{\bar{r}T-3} & \sigma_{\bar{r}T-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ \sigma_{\bar{r}t-2} & \sigma_{\bar{r}t-3} & \cdots & \ddots & & \cdots & \sigma_{\bar{r}T-t} & \sigma_{\bar{r}T-t+1} \\ \sigma_{\bar{r}t} & \sigma_{\bar{r}t-1} & \cdots & & \ddots & \cdots & \sigma_{\bar{r}T-t-2} & \sigma_{\bar{r}T-t-1} \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{\bar{r}T-2} & \sigma_{\bar{r}T-3} & \cdots & \sigma_{\bar{r}T-t} & \sigma_{\bar{r}T-t-2} & \cdots & \sigma^2_{\bar{r}} & \sigma_{\bar{r}1} \\ \sigma_{\bar{r}T-1} & \sigma_{\bar{r}T-2} & \cdots & \sigma_{\bar{r}T-t+1} & \sigma_{\bar{r}T-t-1} & \cdots & \sigma_{\bar{r}1} & \sigma^2_{\bar{r}} \end{bmatrix}^{-1} \tag{25}$$

Next, let's examine $(\bar{\boldsymbol{r}}'_{st'}\bar{\boldsymbol{r}}_{st})$. This matrix is equal to (26).

$$\begin{bmatrix} \bar{\boldsymbol{r}}'_{s1}\bar{r}_{st} \\ \bar{\boldsymbol{r}}'_{s2}\bar{r}_{st} \\ \vdots \\ \bar{\boldsymbol{r}}'_{st-1}\bar{r}_{st} \\ \bar{\boldsymbol{r}}'_{st+1}\bar{r}_{st} \\ \vdots \\ \bar{\boldsymbol{r}}'_{sT-1}\bar{r}_{st} \\ \bar{\boldsymbol{r}}'_{sT}\bar{r}_{st} \end{bmatrix} \tag{26}$$

Again, under the stationarity assumptions in (24), we can simplify this matrix to (27).

$$\begin{bmatrix} \sigma_{\bar{r}t-1} \\ \sigma_{\bar{r}t-2} \\ \vdots \\ \sigma_{\bar{r}1} \\ \sigma_{\bar{r}1} \\ \vdots \\ \sigma_{\bar{r}T-t-1} \\ \sigma_{\bar{r}T-t} \end{bmatrix} \tag{27}$$

In order to circumvent the tradeoff between the number of independent variables and number of observations, I calculate the variance $\sigma^2_{\bar{r}}$ and covariances $\sigma_{\bar{r}y}$ manually using all observations that can contribute to the calculation. I then plug these values back into the matrices in (25) and (27) and manually perform the matrix algebra necessary to obtain the coefficient vector $\boldsymbol{\beta}_{\bar{r}t'}$. Note that the matrices in (20) through (27) will look slightly different for schools that do not have data for all years. Specifically (20) will not contain values $\bar{r}_{st'}$ for years t' for which the school does not have data. This will then follow into the subsequent matrices. Again, this is the advantage of manually calculating the variances and covariances, as it allows us to use a flexible set of projection variables tailored to each school depending on data availability while identifying coefficients using the maximum amount of available variation.

Figure D.10 shows the autocorrelation values $\sigma_{\bar{r}y}$. For all school levels and subjects the correlation between residual values $\bar{r}_{st}$ and $\bar{r}_{st'}$ that are y years apart gradually fades out in an essentially linear fashion. The autocorrelation between years is highest for middle school and similar for elementary and high school.

It is important to note that due to the inclusion of a constant in equation (10) the residuals r_{isgt} will sum to zero by definition of the first order conditions of OLS. For this reason, school value added can only be estimated in relative terms. While equation (10) imposes the constraint that $\sum_{i=1}^{N} r_{isgt} = 0$, I rescale the value-added estimates such that $\sum_{t=1}^{T} \sum_{s=1}^{S} \hat{\mu}_{st} = 0$, effectively altering the constraint on the value-added estimates so that they have mean zero at the school-by-year level instead of at the student level. This rescaling has no impact on the analyses but simplifies the interpretation of the value-added estimates, as a school with positive value added is above average and a school with negative value added is below average.

D Elementary-School Math Value Added

Starting in the 7th grade, students have the option of taking different math assessments based on the math subject in which they are enrolled. This makes calculating school value added in math difficult, because scores are not directly comparable between the various math subjects within grades. In this section I present results for math value added in elementary school, where there is a common test.

Figure D.11 shows the distribution of elementary-school math value added. The standard deviation of elementary-school math value added is 0.134. This tells us that a one standard deviation increase in elementary-school math value added increases the average test score of its students by 13.4% of a student-level standard deviation. The standard deviation is about twice the size of that for ELA, which is consistent with prior studies of school and teacher value added.

Table D.14 explores whether elementary-school math value added impacts future K–12 performance. The outcomes I look at are ELA and math test scores one grade later, whether a student enrolled in a public school one grade later, and whether a student took the most advanced math subject in future grades. The results show that school value added persists to future test scores. A one standard deviation increase in elementary-school math value added increases ELA and math test scores in the next grade by 5.9% and 12.6% of a standard deviation in ELA and math respectively. Elementary-school math value added has an economically insignificant effect on remaining in the public-school system, which combines the effect of transferring to a private school, dropout, and moving to another state. Finally, elementary-school math value added has a positive impact on the math subject that students take, as students are more likely to take the hardest math subject when they first track to different math subjects in grade 7 as well as for their final math exam in grade 11.

Table D.15 shows the results from the regression of postsecondary enrollment on elementary-school math value added. Results show that a one standard deviation increase in elementary-school math value added increases overall enrollment by 0.5 percentage points (0.8%). The impact on 4-year enrollment is 1.6 percentage points (5.6%). Elementary schools that provide high value added on math appear to induce students to enroll in a 4-year university instead of a 2-year community college. While elementary-school math value added appears to induce a similar amount of 2-year to 4-year switches, elementary-school ELA value added has a larger overall effect on postsecondary enrollment that is over triple the size of the effect of elementary-school math value added.

The regressions for CSU outcomes are shown in Table D.16. Elementary schools that provide high value added on math decrease a student's likelihood of being accepted conditional on application, likely due to increases in CSU application on the extensive margin, where students have a low likelihood of acceptance, that dominate any increases in the probability of acceptance on the intensive margin. Elementary-school math value added has a much smaller, although still negative, effect on the need for remedial classes upon enrolling at a CSU than elementary-school ELA value added. Interestingly, this also applies to remedial math classes. An elementary school with math value added one standard deviation above average decreases the need for remedial ELA and math classes by 0.3 percentage points (1.3% and 1.2% respectively). Elementary-school math value added has no effect on whether students become a STEM major, but high value added schools do reduce the likelihood that students are undecided in their first year of college. Unlike elementary-school ELA value added, elementary-school math value added has essentially no effect on whether a student persists to their second or third year of college.

The CCC outcomes are given in Table D.17. Elementary-school math value added reduces the need for remedial CCC courses. A one standard deviation increase in elementary-school math value added decreases the need for remedial math and ELA classes by 0.6 percentage points (2.7% and 2.6%, respectively). The effect size is similar to that of elementary-school ELA value added. Elementary-school math value added has a smaller, although still positive, impact on both persistence to year two and transfer to a four-year university than elementary-school ELA value added. The same is true for degree receipt and associate's degree receipt.

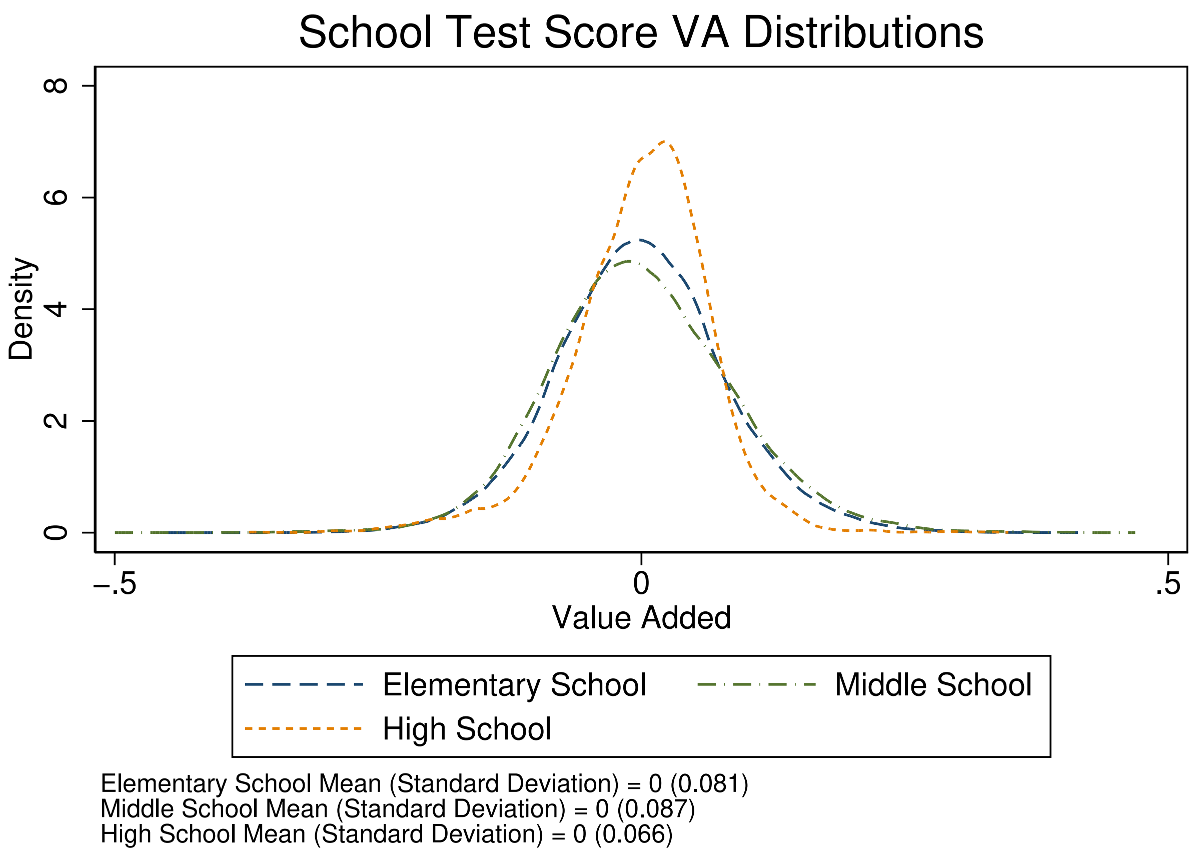


Figure D.1: School Test Score Value Added Distributions

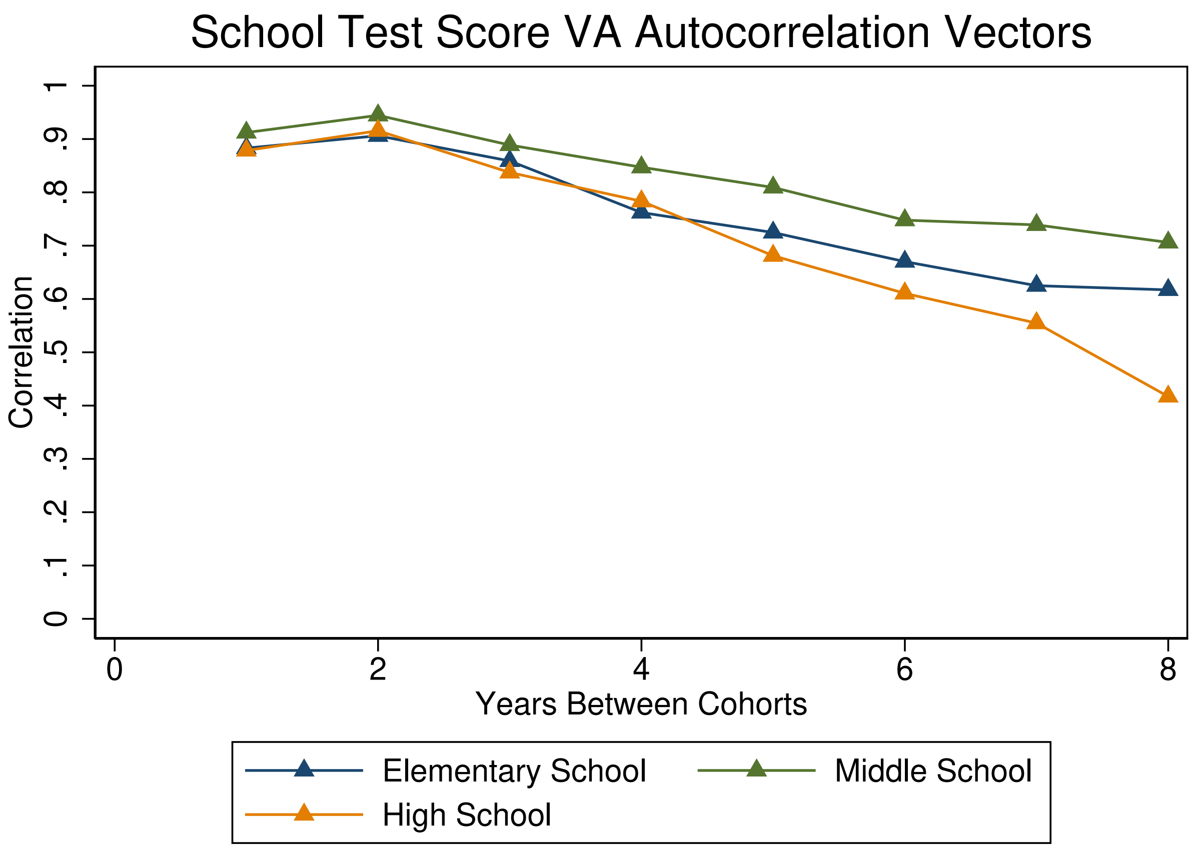


Figure D.2: School Test Score Value Added Autocorrelation Vectors

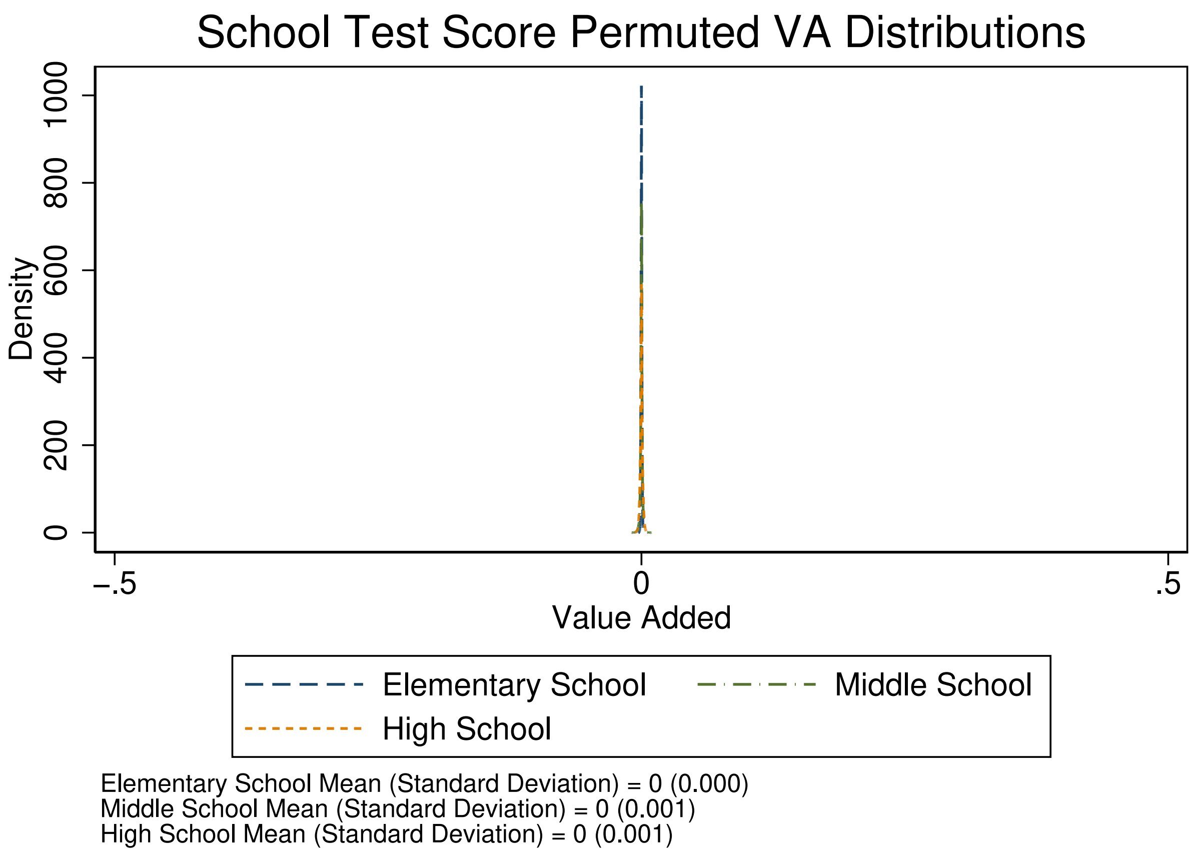
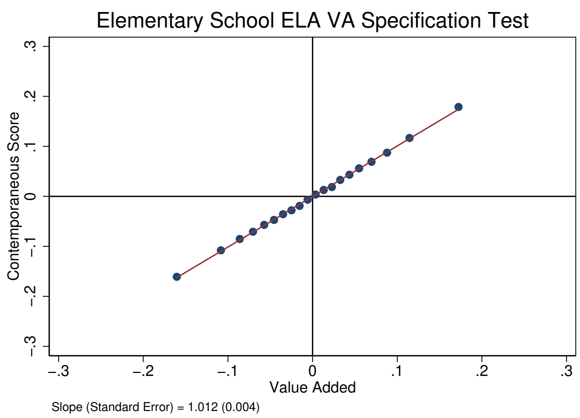
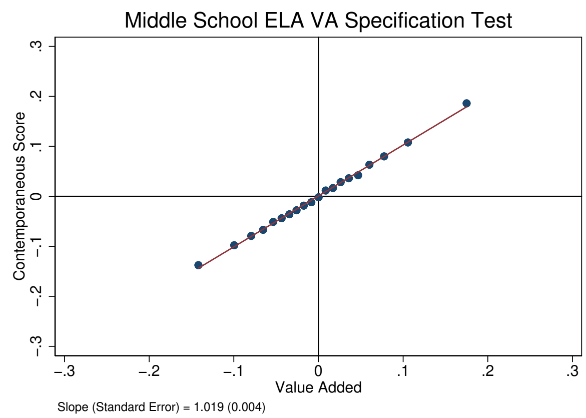


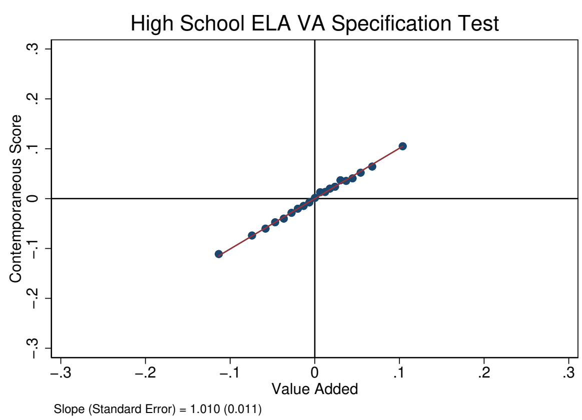
Figure D.3: School Test Score Permuted Value Added Distributions



(a) Elementary School

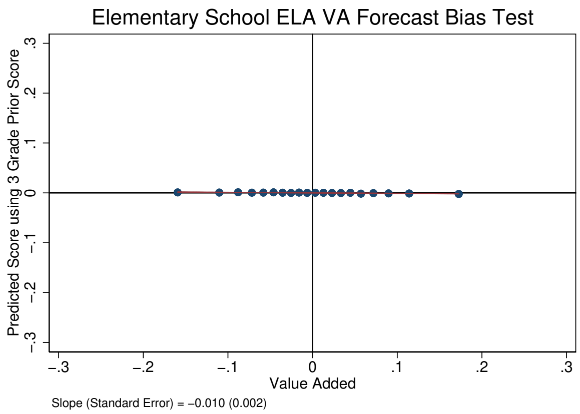


(b) Middle School

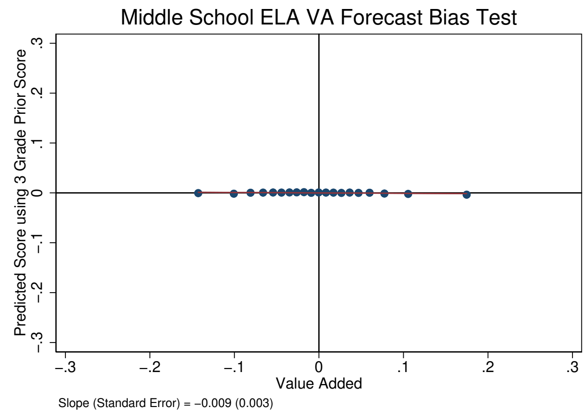


(c) High School

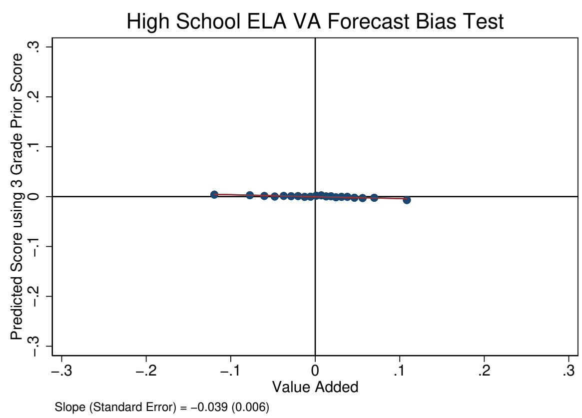
Figure D.4: School Test Score Value Added Specification Tests



(a) Elementary School

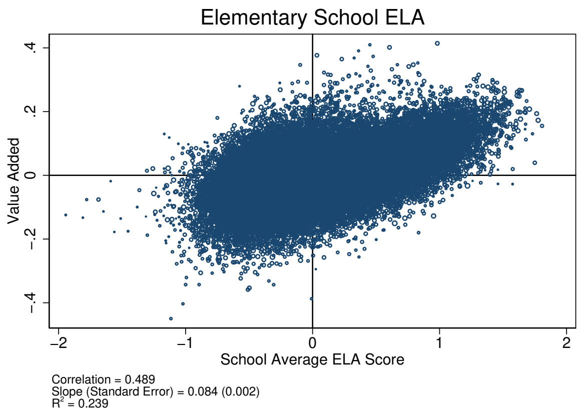


(b) Middle School

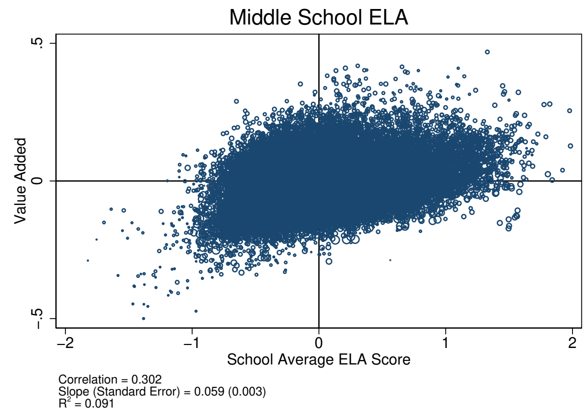


(c) High School

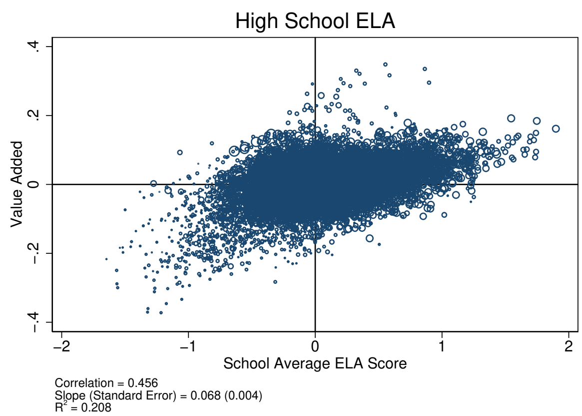
Figure D.5: School Test Score Value Added Forecast Bias Tests



(a) Elementary School



(b) Middle School



(c) High School

Figure D.6: School Test Score Value Added vs. School Average Test Score

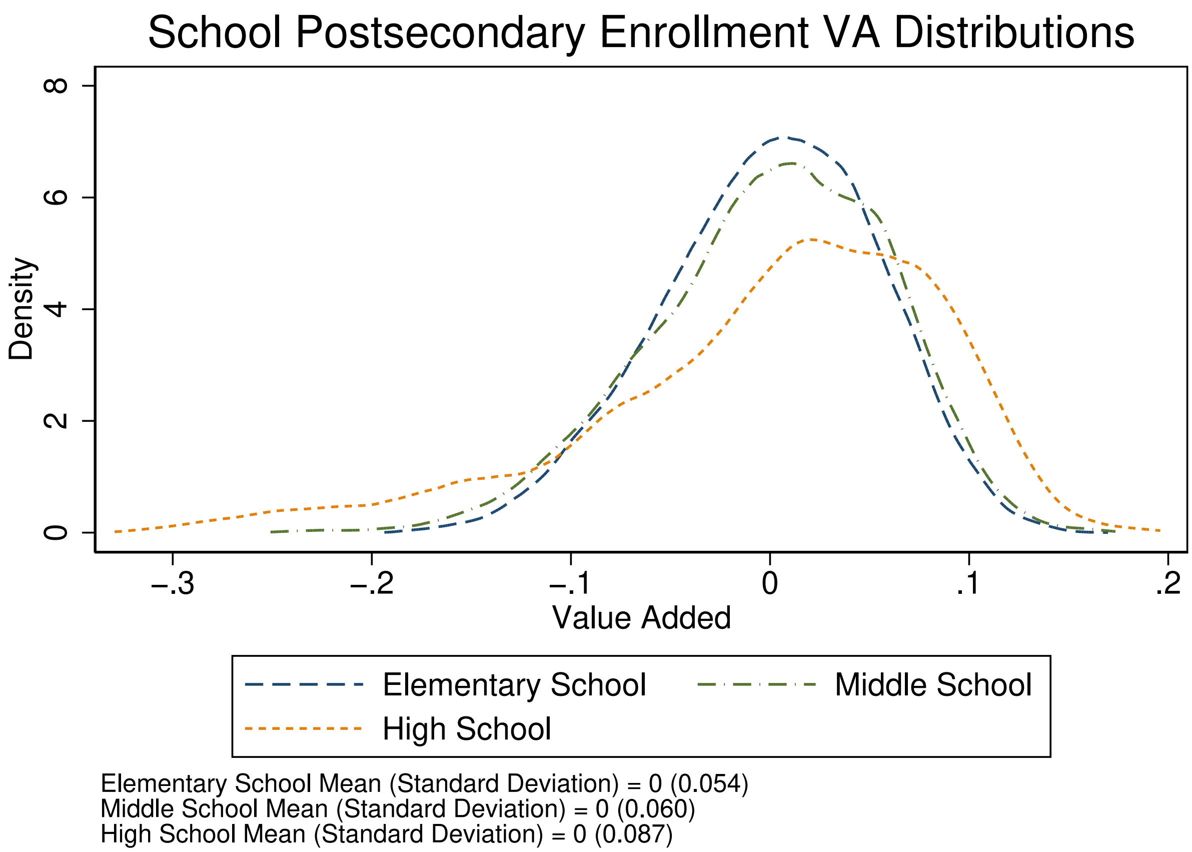
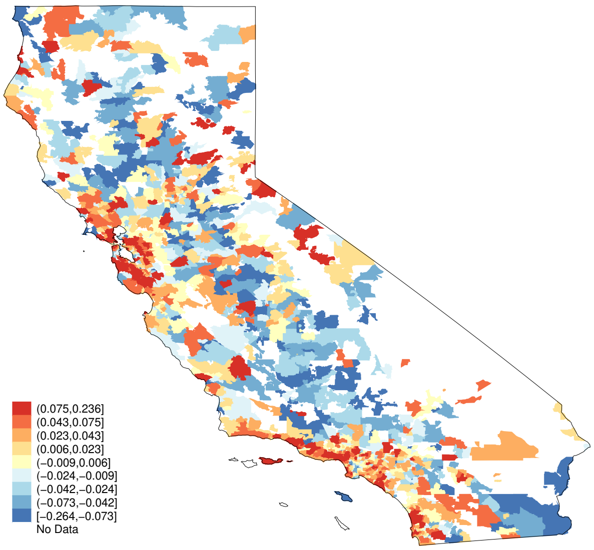
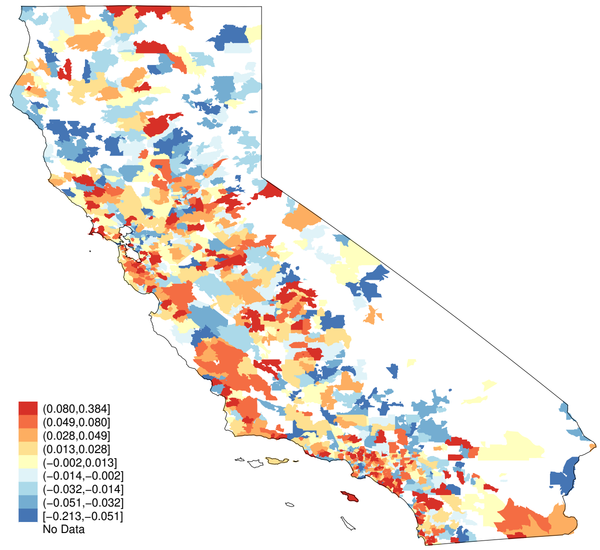


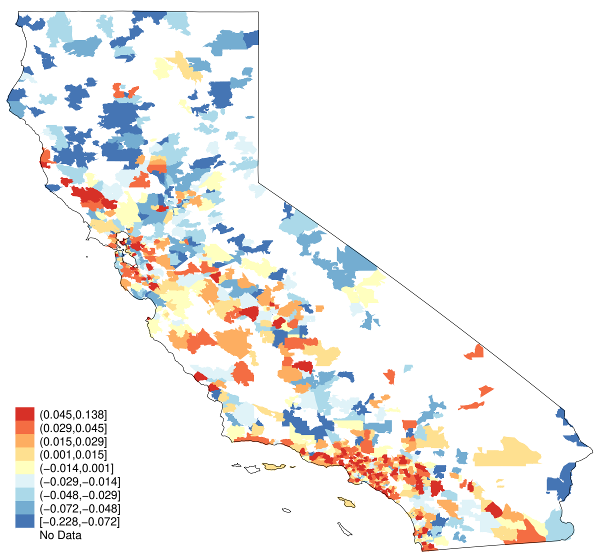
Figure D.7: School Postsecondary Enrollment Value Added Distributions



(a) Elementary School

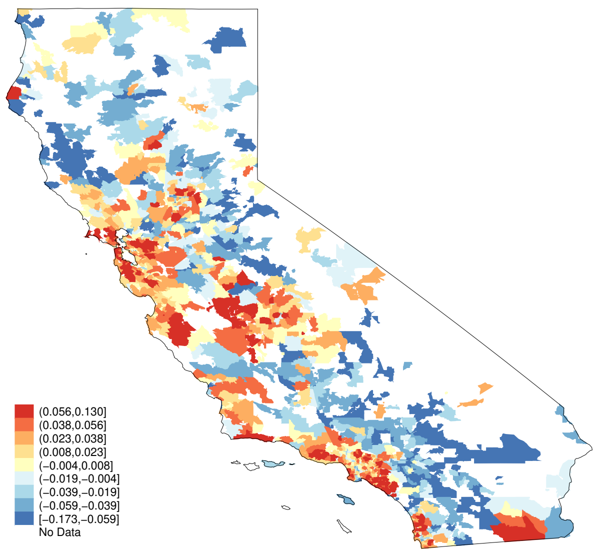


(b) Middle School

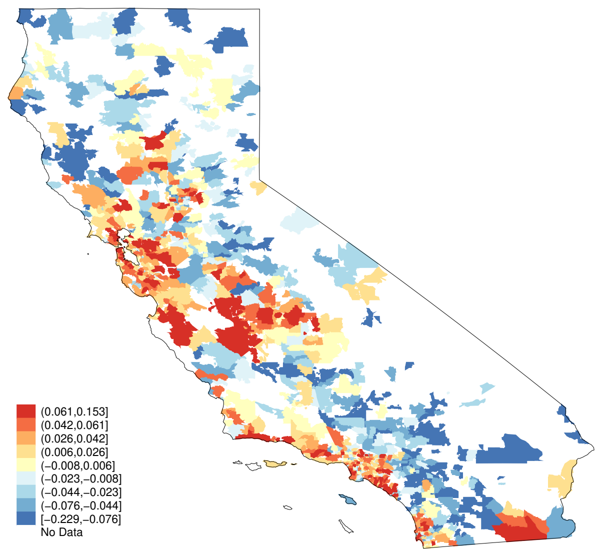


(c) High School

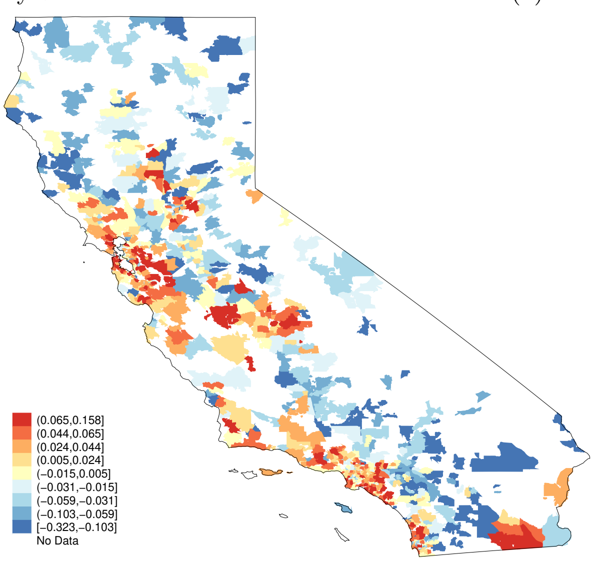
Figure D.8: Map of School Test Score Value Added by Zip Code



(a) Elementary School



(b) Middle School



(c) High School

Figure D.9: Map of School Postsecondary Enrollment Value Added by Zip Code

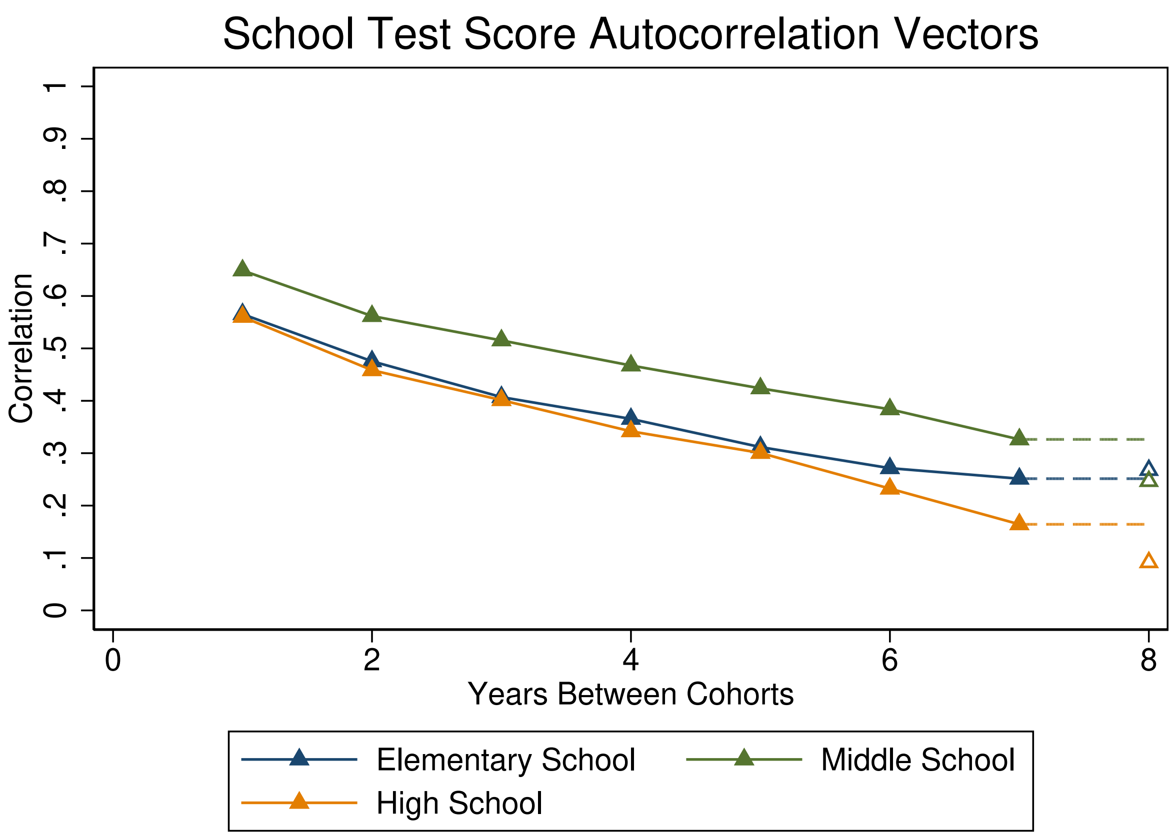


Figure D.10: School Test Score Autocorrelation Vectors

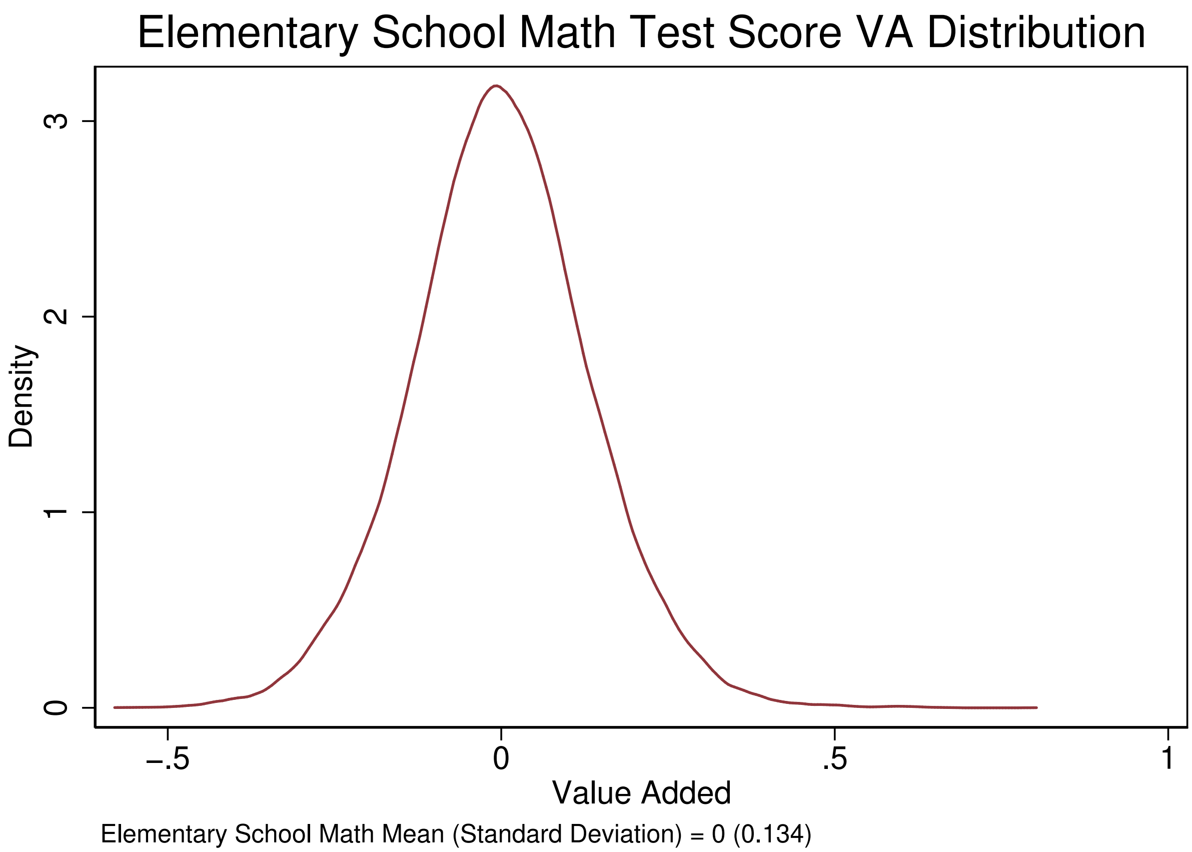


Figure D.11: Elementary-School Math Test Score Value Added Distribution

Table D.1: K–12 Summary Statistics

	Elementary	Middle	High
School Controls			
# of Students in School-Grade	104.1	326.2	549.1
	[43.58]	[193.1]	[233.8]
Demographic Controls			
Age in Years	10.22	12.74	15.71
	[0.644]	[0.921]	[0.922]
Male	0.498	0.497	0.495
	[0.500]	[0.500]	[0.500]
Hispanic or Latino	0.524	0.510	0.472
	[0.499]	[0.500]	[0.499]
White	0.273	0.284	0.310
	[0.445]	[0.451]	[0.463]
Asian	0.122	0.125	0.137
	[0.327]	[0.330]	[0.344]
Black or African American	0.0617	0.0632	0.0617
	[0.241]	[0.243]	[0.241]
Other Race	0.0197	0.0188	0.0189
	[0.139]	[0.136]	[0.136]
Economic Disadvantage	0.595	0.567	0.492
	[0.491]	[0.495]	[0.500]
Limited English Proficient Status	0.251	0.161	0.120
	[0.433]	[0.367]	[0.325]
Disabled	0.0410	0.0387	0.0434
	[0.198]	[0.193]	[0.204]
Test Scores			
Current Test Score	0.0568	0.0688	0.135
	[0.982]	[0.979]	[0.965]
1 Grade Prior Test Score	0.0801	0.0769	0.156
	[0.973]	[0.974]	[0.961]
2 Grade Prior Test Score	0.102	0.0794	0.157
	[0.968]	[0.975]	[0.964]
Schools	6,036	5,068	1,593
Students	3,407,230	3,903,559	3,819,155
Observations	5,785,167	8,541,805	7,911,067

Values are means and standard deviations [in brackets] of the dependent and independent variables used in the test score value added estimation. Only students included in the test-score value-added sample are included in this table. Data comes from public schools in the state of California between the 2004–2005 and 2012–2013 school years. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11.

Table D.2: Datasets

Dataset	Begin	End	Description
CST	Spring 2003	Spring 2013	CA public-school 2nd-11th graders
NSC	Spring 2010	Spring 2017	CA graduates linked to national postsecondary
CSU Application	Fall 2001	Spring 2017	Universe of applicants
CSU Enrollment	Fall 2001	Spring 2017	Universe of enrolled students
CSU Degree	Fall 2001	Spring 2016	Universe of degree recipients
CCC Enrollment	Fall 1992	Spring 2017	Universe of enrolled students
CCC Degree	Fall 1992	Spring 2016	Universe of degree recipients

Table D.3: School Test Score Value Added Specification/Forecast Bias Tests

	Elementary	Middle	High
VA Specification Test: Contemporaneous Score	1.012***	1.019***	1.010
	(0.004)	(0.004)	(0.011)
	[1.003,1.020]	[1.011,1.026]	[0.989,1.031]
VA Forecast Bias Test: Prior Score	-0.010***	-0.009***	-0.039***
	(0.002)	(0.003)	(0.006)
	[-0.015,-0.006]	[-0.014,-0.003]	[-0.050,-0.027]

Each cell represents a separate regression. The first row contains the coefficient from a bivariate regression of test score residuals r_{isgt} on school value added $\hat{\mu}_{st}$. Statistical inference is conducted under the null hypothesis that the coefficient equals 1. The second row contains the coefficient from a regression of the projection of test scores onto three grade prior test scores on school value added $\hat{\mu}_{st}$. Statistical inference is conducted under the null hypothesis that the coefficient equals 0. Standard errors cluster bootstrapped at the school level are presented in parentheses. The 95% confidence intervals are presented in brackets.

Table D.4: K–12 Outcomes on School Test Score Value Added

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Elementary School	1 Grade Later CST ELA Score	1 Grade Later CST Math Score	Stayed in Public School	Highest Math Subject in Grade 7		Highest Math Subject in Grade 11
Elementary School ELA Value Added	0.088***	0.093***	-0.001***	0.002		0.020***
	(0.001)	(0.002)	(0.000)	(0.001)		(0.002)
Y Mean	0.056	0.050	0.793	0.081		0.373
Observations	4,782,394	4,779,116	5,776,992	3,587,713		640,149
R^2	0.711	0.548	0.784	0.154		0.269
Panel B: Middle School	1 Grade Later CST ELA Score		Stayed in Public School		Highest Math Subject in Grade 9	Highest Math Subject in Grade 11
Middle School ELA Value Added	0.067***		-0.001		-0.001	0.005*
	(0.001)		(0.000)		(0.002)	(0.003)
Y Mean	0.072		0.771		0.062	0.344
Observations	7,010,235		8,528,085		5,758,625	2,900,331
R^2	0.741		0.753		0.137	0.310
Panel C: High School	1 Grade Later CST ELA Score		Stayed in Public School			Highest Math Subject in Grade 11
High School ELA Value Added	0.104***		0.003**			0.009***
	(0.001)		(0.001)			(0.004)
Y Mean	0.084		0.771			0.324
Observations	4,268,621		5,320,860			5,593,418
R^2	0.715		0.697			0.337

Each cell is a separate regression of the outcome listed in the column header on school value added. Panels A-C are differentiated by which school level value added is included as an independent variable. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11. Each regression also includes the controls included in the estimation of school value added. Standard errors cluster bootstrapped at the school level are presented in parentheses.

Table D.5: Postsecondary Summary Statistics

	Elementary	Middle	High
Enrolled at a Postsecondary Institution	0.629	0.630	0.620
	[0.483]	[0.483]	[0.485]
Enrolled at a 2-Year College	0.367	0.373	0.372
	[0.482]	[0.484]	[0.483]
Enrolled at a 4-Year University	0.262	0.258	0.248
	[0.440]	[0.437]	[0.432]
Enrolled at a Public Institution	0.574	0.573	0.560
	[0.495]	[0.495]	[0.496]
Enrolled at a Private Institution	0.055	0.057	0.060
	[0.229]	[0.232]	[0.237]
Enrolled at a CA Institution	0.569	0.571	0.561
	[0.495]	[0.495]	[0.496]
Enrolled at an Out-of-State Institution	0.060	0.059	0.058
	[0.238]	[0.236]	[0.234]
Observations	4,291,249	4,355,400	4,201,902

Values are means and standard deviations [in brackets] of the extensive margins of college. Only students included in the postsecondary-enrollment value-added sample are included this table. Data comes from public schools in the state of California between the 2004–2005 and 2012–2013 school years. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11.

Table D.6: Postsecondary Enrollment on School Test Score Value Added

Panel A: Elementary School	(1) Enrolled	(2) Enrolled 2-Year	(3) Enrolled 4-Year
Elementary School ELA Value Added	0.016***	-0.008***	0.024***
	(0.001)	(0.001)	(0.001)
Y Mean	0.652	0.369	0.283
Observations	2,701,070	2,701,070	2,701,070
R^2	0.114	0.027	0.192
Panel B: Middle School	Enrolled	Enrolled 2-Year	Enrolled 4-Year
Middle School ELA Value Added	0.018***	0.007***	0.011***
	(0.001)	(0.002)	(0.002)
Y Mean	0.649	0.373	0.276
Observations	6,358,028	6,358,028	6,358,028
R^2	0.119	0.033	0.207
Panel C: High School	Enrolled	Enrolled 2-Year	Enrolled 4-Year
High School ELA Value Added	0.022***	-0.006**	0.028***
	(0.002)	(0.003)	(0.003)
Y Mean	0.655	0.382	0.273
Observations	6,077,315	6,077,315	6,077,315
R^2	0.125	0.042	0.224

Each cell is a separate regression of the outcome listed in the column header on school value added. Panels A-C are differentiated by which school level value added is included as an independent variable. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11. Each regression also includes the controls included in the estimation of school value added. Standard errors cluster bootstrapped at the school level are presented in parentheses.

Table D.7: Postsecondary Enrollment on Horse Race of School Test Score Value Added

	(1) Enrolled	(2) Enrolled 2-Year	(3) Enrolled 4-Year
Elementary	0.002***	-0.008***	0.010***
	(0.001)	(0.001)	(0.001)
Middle	0.005***	0.020***	-0.015***
	(0.001)	(0.001)	(0.001)
High	0.021***	-0.004***	0.025***
	(0.001)	(0.001)	(0.001)
Elementary × Middle	0.002***	-0.003***	0.005***
	(0.001)	(0.001)	(0.001)
Middle × High	-0.004***	0.001	-0.005***
	(0.001)	(0.001)	(0.001)
Elementary × High	-0.003***	-0.004***	0.001*
	(0.001)	(0.001)	(0.001)
Elementary × Middle × High	-0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)
Y Mean	0.680	0.382	0.299
Observations	1,068,706	1,068,706	1,068,706
R^2	0.133	0.051	0.246

Each column is a separate regression of the outcome listed in the column header on school value added. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11. Each regression also includes the controls included in the estimation of school value added, averaged across grades. Heteroskedasticity-robust standard errors are presented in parentheses.

Table D.8: CSU Outcomes on School Value Added

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Elementary School	Accepted	Eng. Rem.	Math Rem.	STEM Major	Undecided	Persist to Year 2	Persist to Year 3		
Elementary School ELA Value Added	-0.006***	-0.021***	-0.023***	0.000	-0.019***	0.012***	0.014***		
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		
Y Mean	0.778	0.227	0.243	0.344	0.206	0.836	0.759		
Observations	705,594	287,420	287,420	288,616	288,616	224,291	171,644		
R^2	0.068	0.195	0.156	0.017	0.019	0.030	0.032		
Panel B: Middle School	Accepted	Eng. Rem.	Math Rem.	STEM Major	Undecided	Persist to Year 2	Persist to Year 3	Degree	STEM Degree
Middle School ELA Value Added	-0.005***	-0.022***	-0.032***	-0.001	-0.018***	0.012***	0.014***	0.008	-0.004
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.006)	(0.005)
Y Mean	0.805	0.229	0.226	0.333	0.195	0.853	0.771	0.603	0.194
Observations	1,878,369	831,691	831,688	834,821	834,821	698,690	564,558	52,144	52,144
R^2	0.070	0.169	0.127	0.017	0.016	0.026	0.029	0.028	0.014
Panel C: High School	Accepted	Eng. Rem.	Math Rem.	STEM Major	Undecided	Persist to Year 2	Persist to Year 3	Degree	STEM Degree
High School ELA Value Added	-0.006***	-0.000	-0.008**	-0.001	-0.005	0.002	0.002	0.002	-0.000
	(0.002)	(0.004)	(0.004)	(0.002)	(0.005)	(0.002)	(0.003)	(0.003)	(0.002)
Y Mean	0.841	0.233	0.215	0.317	0.187	0.860	0.775	0.627	0.196
Observations	2,410,027	1,207,064	1,207,065	1,212,166	1,212,166	1,114,516	940,304	407,399	407,399
R^2	0.059	0.122	0.089	0.018	0.015	0.020	0.024	0.033	0.014

Each cell is a separate regression of the outcome listed in the column header on school value added. Panels A-C are differentiated by which school level value added is included as an independent variable. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11. Each regression also includes the controls included in the estimation of school value added. Standard errors cluster bootstrapped at the school level are presented in parentheses.

Table D.9: CCC Outcomes on School Value Added

Panel A: Elementary School	(1) Eng. Rem.	(2) Math Rem.	(3) Persist to Year 2	(4) Transfered to 4-Year University	(5) Degree	(6) Associate's
Elementary School ELA Value Added	-0.007***	-0.007***	0.012***	0.026***	0.022***	0.024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Y Mean	0.232	0.223	0.733	0.303	0.389	0.374
Observations	1,226,401	1,226,401	892,405	1,155,703	207,098	207,098
R^2	0.139	0.093	0.052	0.161	0.137	0.144
Panel B: Middle School	Eng. Rem.	Math Rem.	Persist to Year 2	Transfered to 4-Year University	Degree	Associate's
Middle School ELA Value Added	0.003	-0.000	0.014***	0.010***	0.014***	0.014***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Y Mean	0.253	0.243	0.729	0.343	0.372	0.354
Observations	3,342,562	3,342,562	2,765,280	3,041,407	1,493,125	1,493,125
R^2	0.156	0.099	0.054	0.181	0.142	0.152
Panel C: High School	Eng. Rem.	Math Rem.	Persist to Year 2	Transfered to 4-Year University	Degree	Associate's
High School ELA Value Added	-0.003	-0.004	0.013***	0.030***	0.016***	0.015***
	(0.004)	(0.003)	(0.001)	(0.002)	(0.002)	(0.002)
Y Mean	0.269	0.259	0.705	0.362	0.308	0.284
Observations	4,517,170	4,517,170	4,225,586	3,458,398	3,408,606	3,408,606
R^2	0.163	0.105	0.056	0.199	0.145	0.166

Each cell is a separate regression of the outcome listed in the column header on school value added. Panels A-C are differentiated by which school level value added is included as an independent variable. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11. Each regression also includes the controls included in the estimation of school value added. Standard errors cluster bootstrapped at the school level are presented in parentheses.

Table D.10: School Value Added Characteristics

	Test Scores			Enrollment		
	Elementary	Middle	High	Elementary	Middle	High
FTE Teachers per Student	-0.04	1.51	-21.47***	-19.19***	-13.03***	-9.61**
	(2.15)	(2.17)	(4.10)	(2.32)	(3.50)	(3.97)
FTE Pupil Services per Student	5.62	-65.10***	9.59	-1.09	2.69	4.84
	(4.56)	(5.09)	(13.42)	(4.65)	(8.84)	(12.85)
English Learner Staff Per Student	-0.31	7.05***	4.45***	1.61***	4.44**	-2.57*
	(0.53)	(0.60)	(1.54)	(0.54)	(1.84)	(1.40)
Proportion $\leq$ 3 Years Experience Teachers	0.17	-0.39***	-0.33	-0.50***	-0.57***	0.47*
	(0.10)	(0.11)	(0.23)	(0.12)	(0.19)	(0.26)
Proportion Full Credential Teachers	0.82***	1.06***	-0.43*	-0.36*	-0.01	1.05***
	(0.17)	(0.14)	(0.25)	(0.21)	(0.21)	(0.36)
Proportion Male Teachers	1.70	-1.32	7.36*	2.77	0.10	9.75**
	(2.31)	(1.52)	(3.93)	(2.56)	(3.73)	(3.97)
Enrollment Proportion Male	0.37	0.70	5.59	0.56	-0.51	7.80**
	(0.74)	(0.92)	(3.66)	(0.81)	(2.26)	(3.85)
Proportion Male Teachers×Enrollment Proportion Male	-5.56	-2.27	-15.15*	-7.02	-1.69	-17.49**
	(4.48)	(2.96)	(7.75)	(4.94)	(7.26)	(7.64)
Proportion Minority Teachers	-0.66**	-0.23	-0.92	0.53*	0.30	2.43***
	(0.30)	(0.36)	(0.66)	(0.31)	(0.53)	(0.57)
Enrollment Proportion Minority	-1.20***	-0.08	-0.33*	-2.04***	-1.97***	-1.41***
	(0.07)	(0.09)	(0.18)	(0.08)	(0.16)	(0.17)
Proportion Minority Teachers×Enrollment Proportion Minority	1.60***	0.43	1.64**	0.51	1.00*	-1.13*
	(0.33)	(0.44)	(0.77)	(0.35)	(0.59)	(0.67)
Constant	-0.22	-1.01**	-1.11	2.18***	2.10*	-4.20**
	(0.44)	(0.50)	(1.89)	(0.48)	(1.18)	(2.07)
Observations	22,054	15,734	4,573	20,075	6,105	1,078
R^2	.0847	.287	.0348	.219	.235	.201

Each column represents a separate regression of value added on school characteristics. Standard errors clustered at the school level are presented in parentheses.

Table D.11: School Value Added Characteristics

	Test Scores			Enrollment		
	Elementary	Middle	High	Elementary	Middle	High
Instruction Expenditures ($1,000s) per Student	0.01	0.01***	0.01	0.01**	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Pupil Services Expenditures ($1,000s) per Student	-0.33***	-0.17***	-0.09*	-0.51***	-0.52***	-0.36***
	(0.03)	(0.03)	(0.04)	(0.05)	(0.06)	(0.05)
Ancillary Services Expenditures ($1,000s) per Student	0.39***	-0.05	0.24*	1.15***	0.62***	0.47***
	(0.12)	(0.08)	(0.13)	(0.43)	(0.17)	(0.17)
Other Expenditures ($1,000s) per Student	0.03***	-0.01**	0.01	0.01*	0.02**	0.02***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)
General Administration Expenditures ($1,000s) per Student	0.02*	0.02**	-0.01	0.06***	0.07***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Total Enrollment (Thousands)	-0.01	-0.45***	0.18***	-0.23***	-0.06	0.16***
	(0.04)	(0.04)	(0.03)	(0.06)	(0.05)	(0.03)
Constant	0.19***	0.09*	-0.21***	0.53***	0.50***	0.16**
	(0.05)	(0.05)	(0.08)	(0.06)	(0.07)	(0.08)
Observations	36,688	26,901	9,708	21,008	12,405	5,858
R^2	.024	.0886	.0485	.0491	.0474	.0791

Each column represents a separate regression of value added on school characteristics. Standard errors clustered at the school level are presented in parentheses.

Table D.12: K–12 Counts

	Elementary		Middle	High
	ELA	Math	ELA	ELA
All Students	8,533,348	8,533,348	12,976,007	13,232,134
+ Nonmissing Test Score	8,104,810	8,104,810	12,395,493	12,770,573
+ First Test Score for Grade	7,857,137	7,857,137	11,930,098	12,089,459
+ Conventional School	7,720,543	7,720,543	11,657,560	11,164,330
+ School Size $>$ 10	7,719,862	7,719,862	11,656,772	11,163,636
+ Nonmissing Subject Test Score	7,688,015	7,681,991	11,599,425	11,021,578
+ Nonmissing Demographic Controls	7,380,519	7,375,086	11,184,642	10,664,530
+ Nonmissing 1 Grade Prior Test Score	6,574,347	6,563,929	9,848,334	9,202,521
+ Nonmissing 2 Grade Prior Test Score	5,794,367	5,779,309	8,549,077	7,913,160
+ Nonmissing Peer Controls	5,793,109	5,778,066	8,547,750	7,912,702
+ School VA Sample Size $\geq$ 7	5,785,167	5,770,100	8,541,805	7,911,067

Values are counts of the number of observations in each sample. Each row is additive, so the restrictions from all prior rows are also present in the current row. Data comes from public schools in the state of California between the 2004–2005 and 2012–2013 school years. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11.

Table D.13: K–12 Summary Statistics

	Elementary				Middle		High	
	ELA		Math		ELA		ELA	
	Included	Excluded	Included	Excluded	Included	Excluded	Included	Excluded
School Controls								
# of Students in School-Grade	104	107	104	107	326	328	549	492
	[43.6]	[54.1]	[43.6]	[54]	[193]	[220]	[234]	[332]
Demographic Controls								
Age in Years	10.2	10.3	10.2	10.3	12.7	12.8	15.7	15.9
	[.644]	[.71]	[.644]	[.71]	[.921]	[.948]	[.922]	[.991]
Male	.498	.544	.498	.544	.497	.544	.495	.538
	[.5]	[.498]	[.5]	[.498]	[.5]	[.498]	[.5]	[.499]
Hispanic or Latino	.524	.51	.524	.51	.51	.488	.472	.481
	[.499]	[.5]	[.499]	[.5]	[.5]	[.5]	[.499]	[.5]
White	.273	.272	.273	.272	.284	.289	.31	.296
	[.445]	[.445]	[.445]	[.445]	[.451]	[.453]	[.463]	[.456]
Asian	.122	.105	.122	.105	.125	.106	.137	.104
	[.327]	[.307]	[.327]	[.306]	[.33]	[.308]	[.344]	[.306]
Black or African American	.0617	.0938	.0615	.094	.0632	.0983	.0617	.0981
	[.241]	[.292]	[.24]	[.292]	[.243]	[.298]	[.241]	[.297]
Other Race	.0197	.0395	.0196	.0394	.0188	.0372	.0189	.037
	[.139]	[.195]	[.139]	[.195]	[.136]	[.189]	[.136]	[.189]
Economic Disadvantage	.595	.619	.595	.62	.567	.579	.492	.508
	[.491]	[.486]	[.491]	[.485]	[.495]	[.494]	[.5]	[.5]
Limited English Proficient Status	.251	.337	.251	.336	.161	.259	.12	.212
	[.433]	[.473]	[.433]	[.472]	[.367]	[.438]	[.325]	[.409]
Disabled	.041	.163	.0409	.162	.0387	.147	.0434	.101
	[.198]	[.369]	[.198]	[.369]	[.193]	[.354]	[.204]	[.301]
Test Scores								
Current Test Score	.0568	-.144	.0541	-.136	.0688	-.155	.135	-.23
	[.982]	[1.03]	[.991]	[1.01]	[.979]	[1.03]	[.965]	[1.02]
1 Grade Prior Test Score	.0801	-.145	.0779	-.141	.0769	-.161	.156	-.182
	[.973]	[1.03]	[.978]	[1.02]	[.974]	[1.03]	[.961]	[1.02]
2 Grade Prior Test Score	.102	-.172	.1	-.171	.0794	-.348	.157	-.374
	[.968]	[1.02]	[.966]	[1.02]	[.975]	[1]	[.964]	[.962]
Schools	6,036	7,543	6,035	7,543	5,068	7,292	1,593	3,674
Students	3,407,230	1,844,000	3,400,121	1,853,176	3,903,559	2,448,072	3,819,155	2,901,514
Observations	5,785,167	2,856,176	5,770,100	2,871,243	8,541,805	4,527,422	7,911,067	5,405,132

Values are means and standard deviations [in brackets] of the dependent and independent variables used in the value added estimation. The included column contains students included in the value added estimation, while the excluded column contains students who were excluded from the value added estimation. Data comes from public schools in the state of California between the 2004–2005 and 2012–2013 school years. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11.

Table D.14: K–12 Outcomes on School Test Score Value Added

Panel A: Elementary-School Math	(1) 1 Grade Later CST ELA Score	(2) 1 Grade Later CST Math Score	(3) Stayed in Public School	(4) Highest Math Subject in Grade 7	(5)	(6) Highest Math Subject in Grade 11
Elementary School Math Value Added	0.059***	0.126***	0.000*	0.004***		0.012***
	(0.002)	(0.001)	(0.000)	(0.001)		(0.002)
Y Mean	0.057	0.052	0.793	0.081		0.374
Observations	4,770,271	4,767,070	5,761,968	3,578,272		639,038
R^2	0.585	0.648	0.784	0.185		0.318

Each cell is a separate regression of the outcome listed in the column header on school value added. Elementary school includes grades 4–5. Each regression also includes the controls included in the estimation of school value added. Standard errors cluster bootstrapped at the school level are presented in parentheses.

Table D.15: Postsecondary Enrollment on School Test Score Value Added

Panel A: Elementary-School Math	(1) Enrolled	(2) Enrolled 2-Year	(3) Enrolled 4-Year
Elementary School Math Value Added	0.005***	-0.011***	0.016***
	(0.001)	(0.001)	(0.001)
Y Mean	0.652	0.369	0.284
Observations	2,693,741	2,693,741	2,693,741
R^2	0.116	0.026	0.195

Each cell is a separate regression of the outcome listed in the column header on school value added. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11. Each regression also includes the controls included in the estimation of school value added. Standard errors cluster bootstrapped at the school level are presented in parentheses.

Table D.16: CSU Outcomes on School Value Added

Panel A: Elementary-School Math	(1) Accepted	(2) Eng. Rem.	(3) Math Rem.	(4) STEM Major	(5) Undecided	(6) Persist to Year 2	(7) Persist to Year 3	(8)	(9)
Elementary School Math Value Added	-0.006***	-0.003**	-0.003**	0.000	-0.011***	0.002	0.004**		
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		
Y Mean	0.778	0.227	0.243	0.344	0.206	0.836	0.759		
Observations	704,501	287,001	287,001	288,196	288,196	223,982	171,423		
R^2	0.071	0.153	0.191	0.023	0.018	0.030	0.032		

Each cell is a separate regression of the outcome listed in the column header on school value added. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11. Each regression also includes the controls included in the estimation of school value added. Standard errors cluster bootstrapped at the school level are presented in parentheses.

Table D.17: CCC Outcomes on School Value Added

Panel A: Elementary-School Math	(1) Eng. Rem.	(2) Math Rem.	(3) Persist to Year 2	(4) Transfered to 4-Year University	(5) Degree	(6) Associate's
Elementary School Math Value Added	-0.006***	-0.006***	0.004***	0.015***	0.010***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Y Mean	0.232	0.223	0.733	0.304	0.389	0.375
Observations	1,223,384	1,223,384	890,338	1,152,925	206,767	206,767
R^2	0.124	0.110	0.054	0.165	0.141	0.149

Each cell is a separate regression of the outcome listed in the column header on school value added. Elementary school includes grades 4–5, middle school includes grades 6–8, and high school includes grades 9–11. Each regression also includes the controls included in the estimation of school value added. Standard errors cluster bootstrapped at the school level are presented in parentheses.