Easy A's, Less Pay: The Long-Term Effects of Grade Inflation

Jeffrey T. Denning[*] Rachel Nesbit[†] Nolan Pope[‡] Merrill Warnick[§]

October 13, 2025

Abstract

Average grades continue to rise in the United States, prompting discussion about the possibility of grade inflation. However, we know very little about the consequences of grading standards for students. We study how grade inflating teachers affect students. First, we extend, develop, and validate teacher-level measures of grade inflation. We construct two measures of grade inflation, one that measures average grade inflation and another that measures whether teachers are more likely to give a passing grade. These measures pass forecast bias tests common in the teacher effects literature. We show that grade inflation is distinct from teacher value-added, with grade inflating teachers having moderately lower cognitive value-added and slightly higher noncognitive value-added. Next, we consider the effects of grade inflation on future outcomes. Mean grade inflation reduces future test scores, reduces the likelihood of graduating from high school, reduces college enrollment, and ultimately reduces earnings. However, passing grade inflation reduces the likelihood of being held back, increases high school graduation, and increases initial enrollment in two-year colleges.

[*]University of Texas at Austin, NBER, IZA, and CESifo, Email: jeffdenning@utexas.edu
[†]RAND, Email: rnesbit@rand.org
[‡]University of Maryland, NBER Email: npope@umd.edu
[§]University of Georgia, Email: merrill.warnick@uga.edu

1 Introduction

Average grades have risen substantially in recent years (Gershenson, 2018). Although this might, in principle, reflect real improvements in student learning, standardized test scores have not increased at the same rate (Sanchez and Moore, 2022). This divergence suggests that the rise in average grades likely reflects a shift toward lower grading standards. Understanding how grading standards impact students – both in their academic performance and in their subsequent educational and professional outcomes – constitutes a central question for researchers and policymakers. Indeed, school districts throughout the country are in the throes of contentious debates about whether and how to change grading standards, with many teachers reporting an increase in practices that lower grading standards and raise grade inflation (Alex, 2022; Randazzo, 2023; Graham, 2017; Las Vegas Review-Journal Editorial, 2023; Griffith and Tyner, 2025).

Grading standards can arise from a variety of sources, but teachers interact most closely with students and generally have discretion over assigning grades. This paper therefore focuses on teachers' individual grading standards. Effective teachers benefit students across numerous dimensions in both the short- and long-run, including test scores, suspensions, absences, effort, and earnings in adulthood (Koedel and Rockoff, 2015; Petek and Pope, 2023; Jackson, 2018; Chetty et al., 2014a,b). However, the practices that make some teachers more effective than others are poorly understood, which prevents school leaders and policymakers from making optimal personnel, training, and policy decisions (Staiger and Rockoff, 2010). As a notable exception to this critique, a small literature shows that teachers' grading practices, biases, and expectations affect student outcomes (Carlana, 2019; Figlio and Lucas, 2004; Gershenson et al., 2022; Papageorge et al., 2020). Those findings suggest that teachers, schools, and districts could actively change or adopt specific grading policies to benefit students.

Debates on grading hinge on the theoretical ambiguity about whether high standards boost performance by eliciting effort from students or hamper performance by discouraging students. One camp suggests that grade inflation[1] is harmful to students, leaving them unprepared for future educational or vocational endeavors and decreasing their overall effort (Wright, 2019). Others

[1]Much of the prior literature has referred to this concept as "lower grading standards." We use the terms grade inflation, grading leniency, and reduced/heightened grading standards interchangeably; we do not take a stand on the correct name for this phenomenon.

suggest that grade inflation does not negatively affect students and that high grading standards might even discourage students (Kohn, 2002). The arguments put forth by these two camps are not necessarily mutually exclusive. Minimum grades are required to pass classes and continue in high school, so grade inflation may help students by allowing them to continue to accumulate human capital and to graduate when they otherwise would drop out. However, high standards for awarding high grades may provide an incentive for students to study and learn the material. If this incentive is reduced, students may study less and learn less. These two effects are in tension: inflated grades can reduce early failure and dropping out, but blunt incentives to study and learn material.[2] This is a key insight in Costrell's model of grade inflation; changing the standards for graduation may increase effort and learning for some students, while reducing effort and learning for others (Costrell, 1994).[3]

The effects of grade inflation also likely differ in importance for different students. Students on the margin of failing a class are more likely to benefit from a small increase in their assigned grade. Students who are significantly above the passing margin may be harmed by the reduced incentive to study. Given the theoretical ambiguity and clear policy importance, empirical evidence can help settle the question of what grading practices are best for students.

In this paper, we study the effects of grade inflation on high school students. To study the effects of different types of grading standards, we construct two teacher-specific measures of grade inflation. The first is "mean grade inflation," which measures how much higher *on average* grades are than expected, given student standardized test scores and other characteristics. This concept is related to the one measured in Figlio and Lucas (2004), Mozenter (2019), and Gershenson et al. (2022). The second is a novel measure, which we call "passing grade inflation," and which measures inflation at the margin of passing a class – that is, inflating the student's grade from an F to a D or above. We find that these two measures of grading leniency are correlated, but distinct, and highlight the tension between higher standards either improving student performance by eliciting more effort or hampering performance by discouraging students. Consistent with the literature, mean grade inflation decreases contemporaneous test scores. Additionally, we show new evidence

[2] Landaud et al. (2024) show that randomly higher GPAs have long-term benefits for students. This is likely one source of the incentive effects of grades.

[3] Costrell (1994) nicely illustrates this key tension but makes simplifying assumptions such as a single credential and a pooling equilibrium for wages for all students with the same credential level.

that mean grade inflation has long-term negative effects on students. In contrast, we find that lower grading standards may help students at the margin: passing grade inflation increases five-year high school graduation rates and enrollment in postsecondary education following high school.

To construct our measures of grade inflation in high school, we rely on several administrative data sets of high school students. We use administrative data on students and teachers from the nation's second-largest school district, Los Angeles Unified School District (LAUSD). We also use administrative data on students in the universe of public high schools in Maryland linked to administrative college and earnings records. These data offer different strengths. More frequent testing in LAUSD allows us to evaluate the impacts of grading standards on future test scores and include more grades of high school. The Maryland data allow us to explore additional longer-term outcomes such as college enrollment, graduation, and earnings. Despite the disparate settings, the effects of grade inflation on student outcomes are remarkably similar across datasets.

Using these administrative data sets, we confirm that grades are increasing over time and that these grade increases represent grade inflation and not increases in human capital. Figure 1 shows that average GPAs increased over the 2004-2014 school cohorts. Over these 10 years, the mean GPA increased from 2.21 to 2.49, roughly a quarter of a letter grade. Over those same 10 years, for students who score within one-tenth of a standard deviation of the average on the state standardized test, the average GPA rose by 0.41. Students scoring over 3 standard deviations above the mean saw their average GPA increase by 0.17.[4] Average GPA increases are also occurring within subjects, as shown in Appendix Figure A.1. These trends in grade inflation are corroborated by the literature. Several studies that have shown that grading has become more lenient over time, both in high school and in college (Zhang and Sanchez, 2013; Gershenson, 2018; Hurwitz and Lee, 2018; Denning et al., 2022; Sanchez and Moore, 2022).

This paper makes several contributions to our understanding of grade inflation and grading practice. Our first is the conceptual contribution of distinguishing mean grade inflation, which is similar to measures constructed by others in the literature, from passing grade inflation, which is novel. These measures allow us to understand the trade offs of greater grading leniency along the dimensions of effort elicitation and discouragement or persistence. These two measures of grading

[4] The achievement of these top students is unlikely to have increased substantially over time, so these trends suggest that the change in grades may reflect grade inflation rather than increases in human capital.

leniency are correlated (correlation coefficient of .85 in LAUSD and .3 in Maryland) but distinct.[5] Hence we conclude that mean grade inflation and passing grade inflation represent related but sometimes different grading practices of teachers. Because mean grade inflation and passing grade inflation empirically often have opposite effects on student outcomes, the distinction is important. We validate these measures of grading leniency in several ways. We use two tests to show that our measures are forecast unbiased, and perform as well or better than cognitive value-added in these tests. Additionally, we construct grade inflation measures with extra demographic controls available in Maryland and find very similar results. Finally, we estimate the effect of grade inflation on student outcomes separately for math and English courses, and again find very similar results.

Second, we ask how grade inflation relates to other teacher characteristics. One advantage of our data is that we can consider several value-added measures, including those related to student motivation and learning as in Petek and Pope (2023) as well as test score value-added. We document that grading leniency is somewhat correlated with other well-established teacher characteristics, such as cognitive value-added (correlation -.40) and noncognitive value-added (correlation 0.30). If these correlations are causal, they suggest that teachers may face tradeoffs in classroom practices. This motivates our next series of exercises, which test whether grade inflation is on net good or bad for students.[6]

Third, we evaluate the effect of both measures of grade inflation on student academic outcomes such as high school graduation, future test scores, and SAT test taking. Previous work has documented that more lenient teachers reduce performance on tests in subsequent grades (Betts and Grogger, 2003; Figlio and Lucas, 2004; Mozenter, 2019; Gershenson et al., 2022).[7] Consistent with prior evidence, we show that having higher mean grade inflating teachers reduces performance

[5] Failure is uncommon in Maryland, which means there is less variation in passing grade inflation.

[6] While we report correlations for measures of teacher value-added and grade inflation, conceptually they are different in important ways. Teacher value-added can be thought of as a black box, and measures the effect of all teacher attributes and behaviors on student success. Teacher value-added is hard to manipulate via policy, because it is less clear what contributes to high teacher value-added. In contrast, grading practices are a policy choice of teachers (or other school administrators) on how to map student performance into grades.

[7] Our paper is related to work that uses a change in grading policy in North Carolina and finds that inflating grades reduced student attendance and reduced ACT scores for some students (Bowden et al., 2023). This paper is very useful for looking at statewide policy, but does not focus on teachers' role in affecting grades. Our paper is also related to work that uses random assignment of college instructors at the US Naval academy and find that teachers that give higher grades tend to harm performance in follow on courses (Insler et al., 2021). We extend this study by focusing on high shool students, considering longer term outcomes, and developing and validating measures that can be used in settings without random assignment.

on future tests.[8] Having math and English teachers who are on average one standard deviation higher in mean grade inflation reduces test scores in the next year by approximately 0.03 standard deviations. We build on previous work by documenting the persistent effects of grading leniency. We find that having a one standard deviation higher mean grade inflating teacher decreases the likelihood of graduating high school by 0.8 percentage points and of taking the SAT by 0.8 percentage points.

We further contribute to the existing literature by documenting important heterogeneity in the effects of different types of grade inflation. While higher mean grade inflation is detrimental to students' future academic achievement, having a teacher with a higher passing grade inflation measure is beneficial for students. We show that having higher passing grade inflating teachers improves performance on future tests and increases the likelihood of graduating from high school and taking the SAT. Having math and English teachers who are on average one standard deviation higher in passing grade inflation increases students' next year test scores by 0.01 standard deviations and increases the likelihood of graduating high school and taking the SAT by 0.6 and 0.7 percentage points, respectively. In addition, having a higher passing grade inflating teacher decreases the likelihood of being held back in the next year by 1.1 percentage points. Hence, the nature of the grade inflation is critical for understanding the longer-term effects on student outcomes, yet has previously been unexplored.

Fourth, we present novel evidence that grade inflation has longer-term impacts on students, influencing their postsecondary enrollment, graduation, and labor market outcomes. We find that exposure to higher mean-grade-inflating teachers reduces enrollment in postsecondary education programs, particularly Associate's degrees, and also reduces graduation from Associate's degree programs. Furthermore, mean grade inflation reduces the likelihood that a student will be employed up to six years after expected high school graduation and their earnings up to seven years after expected high school graduation. On the other hand, passing grade inflation increases enrollment in Associate's programs after high school graduation, but reduces graduation from Bachelor's programs. We also find suggestions that passing grade inflation may have a small positive effect on students' labor market outcomes via increased earnings and employment, though

[8] Mozenter (2019) finds no effect on longer-term outcomes, but the author only considers the effects of one of our grade inflation measures. He also focuses on middle school students while we focus on high school students. Betts and Grogger (2003) consider the consequences of grade inflation at the school level.

those results are sensitive to the definition of employment and the time period measured.

Fifth, we show that these two measures of grade inflation have heterogeneous effects among different groups of students, especially along the academic ability distribution. Having a higher mean-grade-inflating teacher has similar negative effects on students' future test scores for students of different academic abilities, but reduces grade retention more (in Maryland and Los Angeles) and reduces SAT-test-taking less (in Los Angeles) among students in the bottom of the 8th grade GPA distribution. In addition, the positive effects of passing grade inflation are concentrated among lower-performing students: having a teacher who engages in passing grade inflation increases graduation rates more and reduces grade retention more among students in the bottom of the 8th grade GPA distribution. These results suggest that both mean and passing grade inflation are slightly more beneficial for low-achieving students than for high-achieving students.

The paper proceeds as follows. Section 2 discusses the data and describes grades in the LAUSD. Section 3 discusses our construction and estimation of the two different types of grading leniency. Section 4 discusses the results. Section 5 concludes.

2 Data

We study grading practices in two settings, the Los Angeles Unified School District (LAUSD) and the state of Maryland, using two datasets with complementary strengths. In Los Angeles, students took end-of-year standardized tests through 11th grade. This allows us to create measures of grade inflation and value-added, which rely on current and lagged test scores, and evaluate the impact of those measures on future test scores through 11th grade. In addition, LAUSD teachers graded their students on noncognitive dimensions (effort and cooperation) along with the traditional GPA, allowing us to explore how grade inflation relates to both cognitive and noncognitive measures of value-added. Finally, students in the LAUSD frequently receive failing grades for classes, which means that the passing margin is meaningful for a large part of this dataset. However, the LAUSD data lacks information on student demographic characteristics such as race, gender, and measures of socioeconomic status.

We augment this analysis with linked K-12, college, and employment data from Maryland, which allows us to explore longer-term outcomes in a sample that covers more recent years with more demographic characteristics. Because in Maryland only approximately two years of math and English exams are administered in high school,[9] our analysis in Maryland is based on only a subset of high school students that skew toward grades 9 and 10. In addition, our measure of passing grade inflation in Maryland has less variation than the measure in Los Angeles, as few students in Maryland fail their classes in high school. We therefore view the two data sources as complements which together tell a more complete story about the effects of grade inflation. Wherever possible, we report estimates from both Los Angeles and Maryland in order to provide additional validation of the results. We describe both data sources in detail below.

Los Angeles Unified School District Data

The LAUSD data contain student-year observations from 2004 to 2013 for high school students.[10] In 2004, the school district was 72.5 percent Hispanic, 11.8 percent Black, and 9.1 percent white.[11] In addition, over 70 percent of students received free and reduced price lunch in 2010.[12] Historically, the LAUSD's academic performance has been lower than the national average. In the early 2000s, the LAUSD had graduation rates that were below 50 percent but rose to 70 percent by 2014.

During this time period, students in grades 8 through 11 took the end of the year California state test (CST) in math and English. The CST is a high-stakes, multiple-choice test administered to all California students each spring. The English and math portions each consist of two 90-minute parts. We standardize test scores at the grade-year level. In addition to yearly test scores, the data include information on students' grades in each course and their overall GPA. Students are given a grade of A, B, C, D, or F for each class and GPA is measured on a 0 to 4 scale (e.g., A = 4.0, B = 3.0). These two variables, GPA and end of year test score, are the main components of our measures of grade inflation. The data also provide information about student behavior which we use to construct measures of teacher value-added, as controls in our empirical analysis, and as next-year outcomes. In particular, we use information about the number of days a student was suspended, the number of days a student was absent, whether a student did not progress on time to the next

[9]Specifically, Algebra I, Algebra II, Geometry, and either 9th or 10th grade English are the standardized tests that might be administered to high school students in Maryland.

[10]The data actually begin in 2003, but our main analysis sample starts in 2004 so that we have lagged test scores and behavioral measures for use as controls.

[11]These statistics were constructed using the downloadable files at https://www.cde.ca.gov/ds/ad/fileshistenr8122.asp

[12]These statistics can be found at https://dq.cde.ca.gov/dataquest/

grade (i.e., held back), and whether the student is an English Language Learner (ELL).

We evaluate the impact of grade inflation on a variety of outcome variables. These include whether a student took the SAT and their score conditional on taking it, whether a student took the PSAT and their score conditional on taking it, and indicators for high school graduation within four or five years of starting ninth grade. We also look at next-year outcomes including test scores, absences, suspensions, and whether the student is held back a grade.

Table 1 presents student-level summary statistics for the 985,020 high school students enrolled between 2004-2013 for which we have information about test scores, grades, and behavior. In this table we average over years, so that each observation represents one student.[13] Several characteristics are notable in the LAUSD. First, measures of student success are low. Only 51 percent of ninth graders graduate from the LAUSD within 4 years and 62 percent graduate within 5 years. 12 percent of students are held back at some point, which we define as their administrative grade level being the same as their grade level in the previous year. The average student is absent for 7 percent of school days. While the LA data has very few demographic characteristics, 20 percent of students are English language learners. In the LAUSD, our best measure of college intention is taking the SAT, which is taken by 36 percent of students in our data. Among those who take the SAT, the average score is a 1334 on a 2400 scale, which is approximately the 30th percentile of test takers. Notably, grades are low, with an average GPA of 2.29. Course specific GPAs are even lower for math and English, with averages of 1.76 and 2.19, respectively. During this time period in the LA school district, teachers gave students grades based on their effort and their cooperation in addition to the usual academic course grades. These grades facilitate our estimation of noncognitive teacher value-added measures. Students perform similarly in these noncognitive dimensions, with an average effort GPA of 2.16 and an average cooperation GPA of 2.39.

[13]We require that students have the following information to be included in this sample: school and district identifiers, grade level, math or English grade inflation measures, math or English teacher value-added measures, noncognitive teacher value-added measures (GPA total value-added, fraction days absent value-added, suspension value-added, and held back value-added), lagged math and English test scores, lagged total GPA, lagged fraction days absent, lagged suspended, lagged held back, and an indicator for being ELL. This is the same restriction we require in our empirical analysis; we make the restriction in this table to aid in interpreting our regression results.

Maryland Data

The Maryland data contain student-year observations from 2013 to 2023 for high school students. They are provided via the Maryland Longitudinal Data System.[14] These include educational records from the Maryland State Department of Education linked to Maryland Higher Education Commission records and Maryland Department of Labor records. For high school students, we observe student test scores in Algebra, Geometry, and either 9th or 10th grade English, as well as student grades, demographic characteristics, and teacher identifiers. Notably, Maryland has information about gender, race, ethnicity, and a proxy for low family income (Free and Reduced Price Meals), which the LA data does not provide. We construct test scores and grades in a similar fashion to those in the LA sample. We standardize test scores at the grade-year level, and grades in each course are measured on a 0 to 4 scale. In addition to only administering end of grade standardized tests for a subset of courses, Maryland's exams also changed several times throughout the sample period. Originally the standardized tests were the High School Assessments (HSA), which switched to the Partnership for the Assessment of Readiness for College and Careers (PARCC) around 2014, which then switched to the Maryland Comprehensive Assessment Program (MCAP) around 2022.[15]

Most of the higher education records are sourced from the National Student Clearinghouse, which allows us to track students' enrollment and degrees achieved at any school in the United States, with additional information for those students who stay in Maryland. Our main higher education outcomes of interest are enrollment in any higher education institution one and two years after high school and graduating from any higher education institution four and six years after high school. We are also able to explore enrollment and graduation separately for different degree types as defined by the Maryland Higher Education Commission (Bachelor's, Associate's, Master's, Doctorate, and Professional or Certificate).

The workforce records from the Department of Labor are based primarily on unemployment insurance and provide earnings at the quarterly level. They are therefore subject to the usual caveat that we are only able to see employment with employers who have mandatory reporting to the state unemployment insurance system; students who are self-employed, for example, are not captured in

[14] See https://mldscenter.maryland.gov/

[15] The switch to the MCAP was originally intended for the 2019-2020 school year by was delayed by the pandemic.

our employment and earnings data. Our main workforce outcomes of interest are employment one and six years after high school and annual earnings one and six years after high school. Our main measure of employment is an indicator for having nonzero earnings in any quarter of the year.[16] Our main measure of earnings is winsorized at the 99th percentile. For both higher education and workforce outcomes, we define the time relative to their expected high school graduation – that is, based on when the student was first in 9th grade – to avoid basing the timing on a potentially endogenous outcome such as high school graduation. For example, "enrollment one year after high school" is measured as whether the student was enrolled in a higher education institution five years after 9th grade.

Table 2 presents summary statistics for the Maryland sample of high school students enrolled between 2013 to 2023 for which we have information about test scores, grades, and behavior. As in LA, in this table we average over years so that each observation represents one student. The average math and English test scores are skewed slightly positive, consistent with a small positive selection into the sample given that we require students to have lagged values of test scores and course grades. The average GPA is 2.59 in math and 2.76 in English and 66 percent of students took the SAT. Graduation rates are slightly higher than national averages with 92 percent of 9th graders graduating within 5 years. Thus the students in Maryland appear to have better educational outcomes than those in LAUSD, where many fewer students took the SAT and graduated within 5 years. The Maryland sample is 41 percent white, 34 percent Black, 14 percent Hispanic, and 7 percent Asian. In terms of longer-term outcomes, 65 percent of students enrolled in postsecondary education a year after high school and 27 percent graduated within 4 years. Lastly, 62 percent of students are employed in a job covered by unemployment insurance six years after their expected high school graduation.[17]

3 Estimation

In this section, we construct two measures of teacher-specific grade inflation, mean and passing grade inflation, and show that these measures do not suffer from forecast bias. We also construct

[16] We also explore a more restrictive definition that requires earnings in more than two quarters, which yields qualitatively similar results.

[17] This is slightly lower than is seen for public high school graduates in Texas which have an employment rate of 76 percent 9-11 years after high school graduation Black et al. (2023).

measures of teacher value-added, and describe the estimation of correlation between different grade inflation and value-added measures. Finally, we outline how we estimate the effects of grade inflation on future outcomes.

3.1 Constructing Measures of Grade Inflation

In practice, teachers have a lot of discretion over how they assign grades. For a group of students with the same underlying performance, a teacher can map that performance into different grades. Those different mappings of performance to grades can change the number of students who receive a grade of A, the average GPA of a class, or the number of students who fail. Different mappings of performance to grades are likely to have different effects. For instance, a teacher who does not fail students very often may reduce the probability of a student repeating a grade or failing to graduate because they did not pass a required class. Alternatively, a teacher who gives many students A grades may reduce the incentive for top students to study, which could hurt their performance in future classes. From these two examples, it is clear that grade inflation theoretically could improve or damage a students' future academic performance. Whether grade inflation helps or hurts depends critically on both the type of grade inflation in which the teacher engages and the characteristics of the student.

With this motivation in mind, we explore two types of grade inflation. First, we are interested in characterizing "mean grade inflation" which measures how much the average GPA is inflated. Second, we are interested in characterizing "passing grade inflation" which measures how likely a teacher is to pass a student.

To construct our first measure of mean grading leniency, we follow a method somewhat related to Figlio and Lucas (2004), Gershenson et al. (2022), and Mozenter (2019).[18] We model the

[18] Our method differs in one important way from Figlio and Lucas (2004) and Gershenson et al. (2022): we predict grade as a function of test score and teacher, whereas they predict test score as a function of grade and teacher. Our approach follows from a structural model in which a student's measured performance in a course (their grade) is a function of their underlying academic performance (their test score and prior performance in that subject) combined with whatever discretion the teacher has in assigning the grade (we call that discretion "grade inflation" or "grading leniency").

student's grade as in Equation 1:

$$\begin{aligned} Grade_{ijst} = \beta_1 TestScore_{ijst} + \beta_2 Grade_{ist-1} + \beta_3 MathTest_{ijst-1} \\ + \beta_4 EnglishTest_{ijst-1} + GI^{mean}_{jt} + X_{it}\beta + \varepsilon_{ijst} \end{aligned} \quad (1)$$

where i indexes student, j indexes teacher, s indexes school and t indexes year. The object of interest is GI^{mean}_{jt} which is the year-specific teacher fixed effect. This is the teacher's contribution to grades after controlling for several important factors. First we account for a student's performance in the subject as measured by the corresponding subject test score, $TestScore_{ijst}$. Second, we also control for $Grade_{ijst-1}$ which is the student's grade in the focal subject from the prior year. Third, we control for prior test scores in both math and English, Additionally, we control for student characteristics. The controls include school, grade, and year fixed effects, an indicator for English language learner (ELL), previous year math test score, previous English test score, the total academic GPA, the fraction of days absent, indicators for suspension, and held back status.[19]

For each teacher we calculate $\widehat{GI}^{mean}_{jt}$, which is our measure of grading leniency. We will call this "mean grading leniency" or "mean grade inflation" to distinguish it from passing grade inflation which we discuss below. Mean grading leniency represents how much a teacher raises (or lowers) their students' average grades relative to their academic performance. We make several adjustments to our measures of grade inflation for use in estimating the effect of grade inflation on future outcomes. Following Chetty et al. (2014a), we estimate $\widehat{GI}^{mean}_{jt}$ and $\widehat{GI}^{pass}_{jt}$ using a jackknife empirical Bayes estimator. This approach uses data from surrounding years to estimate a teacher's propensity to grade inflate in year t, which avoids biasing estimates of the long-term effects of teacher grade inflation on student outcomes (Jacob et al., 2010). Including year t in the prediction would likely bias the estimates because unobservables in year t that are related to any dimension of student performance would be captured in both the measure of grade inflation in year t and the outcome of interest.

The estimate $\widehat{GI}^{mean}_{jt}$ is fundamentally defined as a residual. It is worth considering what $\widehat{GI}^{mean}_{jt}$ could be capturing aside from grading leniency. We rule out some potential alternative explanations by controlling for student characteristics. For example, student performance is ac-

[19]For the Maryland sample, because students do not necessarily take math and English exams every year, we replace the previous year test score with the most recent previous test score.

counted for in two ways. First, Equation 1 includes the student's contemporaneous and past standardized test score. Second, we include the student's grades from the previous year. This accounts for students who might perform poorly on subject tests, but demonstrate their understanding of the subject through their performance on non-test assessments.[20]

However, $\widehat{GI}_{jt}^{mean}$ could represent something a teacher does to improve their students' grades in a way not captured by contemporaneous test scores. That could be a skill the teacher conveys to their students that improves grades but not test scores, such as helping students learn to work in groups. In our setting, a teacher who is very good at conveying skills not captured by contemporaneous test scores would have a high $\widehat{GI}_{jt}^{mean}$ measure. As a result, we would expect higher $\widehat{GI}_{jt}^{mean}$ to *improve* future performance in the next year. Instead, we will show in our results that high $\widehat{GI}_{jt}^{mean}$ teachers reduce future performance, which lends more support for our interpretation of this residual as a measure of grading leniency.

Our second measure of grading leniency replaces $Grade_{ijst}$ in Equation 1 with an indicator for passing the class. This indicator is equal to 0 if a student received an F in the course and equal to 1 if the student received a grade of D or better. As discussed, we create this alternative measure of grading leniency because some teachers may raise the grades of their students generally, whereas others may only raise grades when students are on the margin of failing the class. We expect that these two measures of leniency may have different effects on students' future performance.[21] For our second measure, we are still interested in estimating the teacher effect in the modified Equation 1, which we refer to as GI_{jt}^{pass}.

The ultimate goal is for our measures to characterize how much grade inflation a student experiences in a given year.[22] To accomplish this, we calculate $\widehat{GI}_{jt}^{mean}$ and $\widehat{GI}_{jt}^{pass}$ for each teacher-year-subject observation and standardize this to be mean zero, standard deviation one within year and subject.[23] We then characterize the grade inflation that a student experiences in a year by

[20]Insofar as a student has a surge in performance in year t (above that expected by their contemporaneous test scores and prior performance in that subject) our measure will not account for this.

[21]We study two measures of how teachers map student performance into grades, but there are many potential alternative measures of this mapping. For example, one teacher may be more likely to give Bs (and fewer Cs) than another teacher, but not to inflate As or Ds. We focus on mean grade inflation to capture general leniency. We focus on passing grade inflation because of the institutional importance of a student passing a class, and because some schools may have formal or informal policies that pressure teachers not to fail their students.

[22]Note that since we are using high school data, students have more than one math and English class per year, and could even have multiple in the same semester.

[23]Calculating at the subject, rather than course, level means that if a teacher teaches two different math classes in a year, our measure of grade inflation for that teacher will be the same for those two classes.

averaging over all the student's teachers in a given subject, weighted by the number of classes a student takes with that same teacher.[24] We then standardize the students' subject-specific grade inflation measures to be mean zero, standard deviation one within year. In most of our analysis, we use a measure of grade inflation that combines math and English. We generate this by summing the standardized grade inflation measures for each subject. We then standardize this sum so that it is mean zero, standard deviation one within year. We do all of the above separately for mean and passing grade inflation. As a result, in our estimates of the effect of grade inflation on future performance, the coefficient on these measures represents a one standard deviation increase in the grade inflation that a student experiences in a given year.

3.2 Validating Measures of Teacher Grade Inflation

Our measures of grade inflation face some of the same challenges faced by value-added measures. First, these measures are estimated with noise and may not be predictive of actual student outcomes out of sample. We refer to this lack of predictive power as estimation error. For example, if student grades were very noisy year after year due to random variation in unobservables, the variation could drown out the teacher signal, and teacher measures would not be very predictive of student outcomes. Second, these measures may be biased due to student selection on some observable omitted by our econometric model. We refer to this type of bias as selection. For example, if certain students who earned high grades through some means not captured by our model consistently are assigned to teachers who give high grades, other past measures of student ability not included in our model would predict that those students would have higher grades.

Both of these concerns would result in forecast-biased measures of grade inflation. As in Chetty et al. (2014a), forecast bias is measured by one minus the correlation between true theoretical teacher measures and the estimated teacher measure:

$$Bias = \left(1 - \frac{cov(GI_{jt}, \widehat{GI}_{jt})}{var(\widehat{GI}_{jt})}\right).$$

Estimation error and selection would push forecast bias in different directions. Estimation error can cause forecast bias because very noisy estimates of grade inflation would lead to a zero correlation

[24] Weighting means that if a student takes three math courses and two are with the same teacher, that teacher's grade inflation measure will be used twice in the average.

on the right hand side of the expression and bias of 1. Selection can cause forecast bias because it will misattribute some of the effect of the omitted variable on the outcome to the teacher effect, leading to a lower covariance between true and estimated measures.

To address these concerns, we conduct two tests for forecast bias. Both tests use jackknifed estimates of mean and passing grade inflation, which use all other years of data to estimate value-added $\widehat{G}I_{jt}$ this year. The first, which we call the "individual test," addresses the first concern, by checking whether these teacher measures are predictive out of sample. To implement this test, we regress student grades and passing indicators on grade inflation measures, controlling for the same observables as in the estimating equations:

$$
\begin{aligned}
Grade_{ijst} &= \delta^{mean}\widehat{GI}_{jt}^{mean} + \alpha_1 TestScore_{ijst} + \alpha_2 Grade_{ist-1} + X_{it}\alpha + \eta_{ijst} \\
Pass_{ijst} &= \delta^{pass}\widehat{GI}_{jt}^{pass} + \gamma_1 TestScore_{ijst} + \gamma_2 Grade_{ist-1} + X_{it}\gamma + \xi_{ijst}.
\end{aligned}
$$

In these regressions, $1-\widehat{\delta}$ will be our estimates of forecast bias. In this test, an estimate of δ close to 1 shows that grade inflation measures, though subject to estimation error, are actually predictive of student outcomes out of sample. This predictive power results in lower forecast bias.

Results from the individual forecast bias tests are found in Table 3. The first two columns show that our grade inflation measures show small levels of forecast bias across both subjects and samples. Across both subjects, samples, and measures, we see estimates of forecast bias ranging from 0.03 to 0.1, suggesting our measures are quite predictive. These estimates of forecast bias are similar to estimates from the value-added literature.[25]

The second exercise, which we call the "prediction test," addresses the second concern of omitted observables. Specifically, we test whether selection on twice-lagged math and English test scores leads to bias in our grade inflation measures.[26] To conduct this test, we begin by residualizing the predictors (twice-lagged test scores) and the outcomes (grade and the passing indicator) on the same set of controls that were used in estimation of our teacher measures, which include

[25]For example, Kane and Staiger (2008) obtain forecast bias estimates of 0.06 to 0.17 when testing measures of value-added of Kindergarten students. Chetty et al. (2014a) obtain forecast bias estimates of 0.02 to 0.05 for children in grades 3-8.

[26]We only conduct this exercise in LAUSD because students take fewer standardized tests in Maryland.

school, grade and year fixed effects, English Language Learner status, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension and held back. We then construct predicted outcomes by regressing the residualized outcome on the residualized predictors and construct fitted outcomes $\widehat{Grade}$ and $\widehat{Pass}$. Finally, we conduct the forecast bias test by regressing the fitted outcomes on our teacher measures:

$$\widehat{Grade}_{ijst} = \nu^{mean}\widehat{GI}_{jt}^{mean} + \eta'_{ijst},$$

with a similar equation for passing grade inflation. The estimates of ν are then estimates of forecast bias. The closer $\widehat{\nu}$ is to zero, the less sorting due to the predictors, and the less forecast bias due to sorting on the omitted observables.

Panel C from Table 3 shows the results of the prediction tests. Again, across both subjects and outcomes, our grade inflation measures show low levels of forecast bias. The forecast bias estimates range from 0.00 to 0.02 in absolute value, providing even stronger evidence than the estimates from the individual tests.

Together, these tests suggest that our measures of grade inflation are predictive of instructor grading behavior and do not suffer from unobserved selection. Establishing these patterns is important for our use of this measure in our paper, but also for future work that use this measure.[27]

3.3 Constructing Other Measures of Teacher Quality

In order to understand how related our measures of grade inflation are to more common indicators of teacher quality, we construct both test score and noncognitive value-added measures. We estimate cognitive value-added using a jackknife empirical Bayes estimator following Chetty et al. (2014a). We also estimate noncognitive valued-added using a jackknife empirical Bayes estimator following Petek and Pope (2023). We create six different noncognitive value-added measures: absences, suspension, grade retention, total academic GPA, cooperation GPA, and effort GPA. Due to concerns about teachers affecting these noncognitive measures directly, we follow Petek and Pope (2023) and use outcomes measured in the year after the student and teacher interact.

We then create student-year level measures of value-added in a similar way to our student-

[27] Figlio and Lucas (2004); Mozenter (2019); Gershenson (2018) use related measures but we are unaware of prior work showing that they are forecast unbiased.

year level measures of grade inflation. We first construct test score and noncognitive teacher value-added measures for each teacher-subject-year observation, standardized to have mean zero, standard deviation one within year. We then construct a student-level value-added measure that averages the value-added from all the student's teachers in a given subject/year, standardized to be mean zero and standard deviation one. We combine math and English test score value-added into a single "cognitive value-added" measure by summing these measures across subjects and then standardizing to be mean zero, standard deviation one. Similarly, following Petek and Pope (2023) we combine math and English course measures of days absent, suspension, grade retention, cooperation GPA, total academic GPA, and effort GPA[28] into a single "noncognitive" value-added measure by summing the components and standardizing to be mean zero, standard deviation one. We also create subject-specific "noncognitive" value-added measures by summing the components from only that subject's teachers and standardizing to be mean zero, standard deviation one.

3.4 Estimating Correlations Between Teacher Measures

In order to understand whether or not grade inflation is simply another measure of value-added we construct correlations between both grade inflation measures and both value-added measures. As described in Jackson et al. (2024), raw correlations between estimates of value-added may be too small or too large relative to the cross-measure correlation between true value-added. First, the correlations may be attenuated because of measurement error-induced noise in grade inflation and value-added measures. Second, the correlations may be too large because if these measurement errors are correlated across outcomes within the same year.

In order to address these challenges, we estimate the correlation between value-added measures using a split-sample, bootstrapping approach. Splitting the sample before estimation resolves the concern of within-year correlations in measurement error across outcomes. We randomly split the sample within classroom,[29] then estimate teacher measures μ_A and μ_B, with the subscripts denoting different outcome measures (e.g. grade inflation or cognitive value-added), separately on both samples. This produces two estimates of value-added for each measure: μ_A^0, μ_A^1, μ_B^0, and μ_B^1, with the superscripts denoting the separate samples. Next, we estimate the across-sample

[28] The cooperation and effort GPAs are only available in LAUSD.
[29] This within-classroom splitting approach was inspired by the correlation estimation of student-level measures in Ayllón et al. (2025).

correlation between the measures as

$$\widehat{\rho}_{AB}^{01} = corr\left(\mu_A^0, \mu_B^1\right).$$

We then adjust this correlation for measurement error attenuation by dividing by the product of the within-measure, across-sample correlations $\widehat{\rho}_A^{01} = corr\left(\mu_A^0, \mu_A^1\right)$ and $\widehat{\rho}_B^{01} = corr\left(\mu_B^0, \mu_B^1\right)$. This adjustment addresses the attenuation concern from measurement error in teacher measures. The adjustment gives us $\widehat{r}_{AB}$ as a unattenuated estimate of the correlation between the teacher measures:

$$\widehat{r}_{AB} = \frac{\widehat{\rho}_{AB}^{01}}{\sqrt{\widehat{\rho}_A^{01}\widehat{\rho}_B^{01}}}.$$

Note that each sample split will give us two estimates of the correlation $\widehat{r}_{AB}$ and $\widehat{r}_{BA}$, which need not be numerically identical. To obtain our actual correlation estimates, we bootstrap this quantity by splitting the sample in half 100 times, then averaging across the 200 resulting estimates.

3.5 Estimating the Effects of Grade Inflation

To explore the effects of grade inflation on future outcomes, we estimate specifications similar to Chetty et al. (2014b), Petek and Pope (2023), Gershenson et al. (2022), Figlio and Lucas (2004), and Mozenter (2019) where an observation is a student-year and we regress a longer-term outcome on the two grade inflation measures, the test score and noncognitive value-added measures, and the same set of controls used to construct these measures. In particular we estimate the following equation:

$$Y_{it} = \alpha_{mean}\widehat{GI}_{it}^{mean} + \theta_{pass}\widehat{GI}_{it}^{pass} + \delta_{cogVA}\widehat{VA}_{it}^{test} + \psi_{noncogVA}\widehat{VA}_{it}^{noncog} + X_{it}\beta + \eta_{it} \tag{2}$$

where Y_{it} is a future outcome, like graduation from high school or test score performance in the following year, $\widehat{GI}_{it}^{mean}$ is the average measure of mean grade inflation a student experiences in year t, $\widehat{GI}_{it}^{pass}$ is the average measure of passing grade inflation a student experiences in year t, $\widehat{VA}_{it}^{test}$ and $\widehat{VA}_{it}^{noncog}$ are the average measures of test score and noncognitive value-added a student experiences in year t, and X_{it} is the same vector of individual level controls we use when estimating grade inflation in Equation 1.[30]

In our main results we use measures of grade inflation and value-added which are averages of the subject-specific measures, as described in Sections 3.1 and 3.3. We also explore how grade inflation might be different in math and English classes and might have different effects on future performance. In those specifications, we use the subject-specific versions of our grade inflation and value-added measures. In both cases, the grade inflation and value-added measures have been standardized to mean zero, standard deviation one.

In regressions based on Equation 2, an observation is a student-year. We cluster our standard errors at the school level to account for within-school correlation of outcomes. The coefficients of interest are α_{mean} and θ_{pass} which estimate the effect of mean grade inflation and passing grade inflation after accounting for other student and teacher characteristics. We also report δ_{cogVA} and $\psi_{noncogVA}$ to verify that our measures of teacher value-added have the expected estimated effects and to compare magnitudes. In all regressions, we limit our sample as described in Section 2 for Table 1, requiring students to have the necessary information to construct grade inflation and value-added measures. In addition, we implement sample restrictions that vary by the availability of the outcome. For example, the CSTs are only administered through 11th grade, so we do not have future test scores for 11th and 12th graders, and we exclude them from the analysis when future test score is the outcome variable.

4 Results

In this section, we find that grade inflation is a separate characteristic from value-added, and that our two measures of grade inflation are related but distinct. We then find that grade inflation does matter for students' future outcomes. Mean grade inflation decreases students' future standardized test scores, probability of graduation, college enrollment, employment and earnings. Passing grade inflation reduces the probability of being held back, increases the probability of graduation, and increases postsecondary enrollment. We also find that the effects of grade inflation vary across the ability distribution.

[30]These controls include school, grade, and year fixed effects; an indicator for English language learner (ELL); and previous year math test score, english test score, total academic GPA, fraction of days absent, an indicator for being suspended, and an indicator for being heldback.

4.1 How is Grade Inflation Related to Other Teacher Characteristics?

Before exploring the relationship between grade inflation and students' future academic and career outcomes, we first characterize the relationship between our measures of teacher grade inflation and other common measures of teacher quality. We show that the two measures we define in Section 3.1 represent related, but distinct aspects of teacher grading standards. Furthermore, we show that these grade inflation measures capture a separate aspect of teacher quality than that captured by traditional cognitive and noncognitive value-added measures.

Table 4 presents correlations of our two measures of grade inflation with two measures of teacher value-added: cognitive value-added and noncognitive value-added. All correlations are calculated using the method described in Section 3.4 to avoid bias due to measurement error. These correlations are inherently not causal, but rather reflect the extent to which grade inflation and other teacher attributes covary in the observed in the data.

We would expect our two measures of grade inflation to be positively correlated, since a teacher who raises the grades of all students will also increase the probability of a student passing. This is confirmed in panels A and B of Table 4, which show that mean grade inflation and passing grade inflation are positively correlated with each other in both the LAUSD and Maryland settings. In Los Angeles, mean and passing grade inflation have a correlation of 0.86, while in Maryland that correlation is notably lower, around 0.36. This lower estimated correlation in Maryland is likely explained by the low fail rates in Maryland,[31] which translates into less variation in the passing grade inflation measure. Nevertheless, in both settings, the two measures are highly but not perfectly correlated – passing grade inflation and mean grade inflation appear to be distinct attributes. Therefore, we will be able to use both measures in Equation 2 to isolate the effects of different types of grade inflation.

Table 4 also presents estimates of the correlation between teachers' grade inflation and value-added measures. Teachers who inflate grades tend to have lower cognitive (test score) value-added. Mean grade inflation is negatively correlated with our cognitive value-added measure (a combination of math and english test score value-added) at -0.41 for Los Angeles and -0.31 for Maryland. This negative correlation is consistent with Betts and Grogger (2003), Figlio and Lucas

[31] Only 3.1% of Math courses and 2.3% of English courses are failed by Maryland students, compared to 32% and 21% in LAUSD.

(2004), Mozenter (2019), and Gershenson et al. (2022), who find that grade-inflating teachers reduce test scores. Passing grade inflation is also negatively correlated with cognitive value-added, though the correlation is weaker, with estimated correlations of -0.30 and -0.07 in Los Angeles and Maryland respectively. Grade inflation tends to be weakly positively correlated with noncognitive value-added (a combination of measures of value-added on absences, suspensions, grade retention, academic GPA, cooperation GPA, and effort GPA[32]). The correlation between mean grade inflation and noncognitive value-added is estimated to be 0.16 in Los Angeles and 0.12 in Maryland. For passing grade inflation, the estimated correlation is 0.18 in Los Angeles and 0.05 in Maryland.

The correlations between grade inflation and value-added measures suggest that grade inflation represents a distinct teacher characteristic from typical value-added metrics. Mean grade inflation and passing grade inflation are correlated, but are separate attributes. Grade inflation is moderately negatively correlated with cognitive value-added and weakly positively correlated with noncognitive value-added. If these correlations contain a causal component, a teacher's decision to inflate grades could carry a tradeoff because grade inflation is both positively and negatively correlated with other desirable teacher characteristics. For instance, a grade-inflating teacher might induce students to attend class more but reduce student learning as measured by test scores. Inflating grades could even be the optimal choice to maximize certain student outcomes. As a result, in the next section we explore whether grade inflation has a positive or negative impact on students' cognitive and noncognitive performance in high school, college, and their careers, while holding other teacher characteristics constant.

4.2 Does Grade Inflation Matter for Future Outcomes?

Table 5 presents the results from a simple regression of the relevant student outcome on just one of our measures of teacher characteristics (mean or passing grade inflation, or cognitive or noncognitive value-added) at a time. Unless otherwise stated, all regression analysis discussed here and in later sections controls for the same student characteristics used to create the grade inflation and value-added measures (school, grade, year, ELL, previous math test score, previous English test score, previous total academic GPA, previous fraction of days absent, previous suspension, and previous grade retention). Standard errors are clustered at the school level.

[32]The cooperation and effort GPAs are only available in LAUSD.

Table 5 verifies that, when considered in isolation, the estimated associations of teacher value-added measures with student outcomes have the expected sign. Having higher cognitive value-added teachers is associated with increases in test scores, SAT scores, and the likelihood of graduating. Higher noncognitive value-added is associated with reductions in the likelihood a student is held back and increases in the likelihood of graduating. The direction of these associations is consistent with the existing literature. In addition, we estimate that grade-inflating teachers are associated with reductions in future test scores, SAT scores, and the likelihood that a student is held back. However, these results do not separately identify the impact of grade inflation while holding other teacher characteristics (i.e., value-added) fixed. That is, the negative relationship between grade inflation and student outcomes estimated in Table 5 could simply reflect the negative correlation between grade inflation and cognitive value-added rather than any causal effect of grade inflation on outcomes.

To disentangle this, in our main results we explore whether grade inflation matters for future outcomes *conditional on other teacher characteristics* by estimating Equation 2. Throughout we include mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added in each regression. This allows us to identify the effect of each measure of teacher quality on student outcomes while holding constant the other teacher characteristics.

In Table 6 we first focus on cognitive and noncognitive value-added to verify that the estimated coefficients have the expected magnitude and direction. In Los Angeles, having teachers with higher cognitive value-added improves future end-of-course standardized test scores, SAT scores, the likelihood of graduating, and the likelihood of taking the SAT. In Maryland, the estimated effects of cognitive value-added are similar, though the relationships with graduation rates and with SAT scores are not statistically significant. One exception is that in Maryland, cognitive value-added increases the chance of being held back, although the effect size is very small (one tenth of a percentage point). In both Los Angeles and Maryland, having teachers with higher noncognitive value-added improves future English test scores, reduces the likelihood of grade retention, increases graduation rates, and increases the likelihood of taking the SAT. In addition, in Maryland, noncognitive value-added slightly reduces future SAT scores. These results closely mirror those of Petek and Pope (2023).

High School Outcomes

Having determined that Equation 2 yields sensible estimates for well-established value-added measures, we turn to estimating the effects of mean and passing grade inflation on high school outcomes conditional on teacher value-added. Table 6 presents estimates for our main high school outcomes (future math and English test scores, the likelihood of being held back, the likelihood of graduating in five years, the likelihood of taking the SAT, and SAT score); Tables A.11 and A.12 presents estimates for additional high school outcomes such as absences, suspensions, and PSAT scores.

When evaluating the effect of grade inflation on future test scores, we focus on the Los Angeles sample, where standardized testing is more frequent. LAUSD administers standardized tests yearly through 11th grade, whereas Maryland high school students take standardized math and English tests in only about two years of high school.[33] In Los Angeles, we find that mean grade inflation reduces future test scores. A one standard deviation increase in the amount of mean grade inflation a student is exposed to in one year reduces future math test scores by 0.024 standard deviations and future English test scores by 0.020 standard deviations. In comparison, being exposed to higher cognitive value-added teachers increases future test scores by .113 for math and .054 for English. Comparing the grade inflation and value-added estimates, a teacher's grading leniency has a meaningful impact on future test score performance – about 30 percent as large as the effect of traditional value-added measures in math and 44 percent as large in English. This finding closely mirrors that of Figlio and Lucas (2004) and Gershenson et al. (2022).

Beyond future test scores, mean grade inflation also impacts some other high school outcomes. We estimate that mean grade inflation negligibly reduces the chance that a student takes the SAT, which allows us to interpret the effect of grade inflation on an alternative measure of student achievement, SAT scores, without worrying about selection into taking the test. However, in both Los Angeles and Maryland, we find no effect of mean grade inflation on SAT test scores. Mean grade inflation does have a small negative effect on PSAT scores (see Table A.11). The PSAT is designed for a different purpose and test different skills than end-of-course standardized tests. Thus the negative effects of mean grade inflation on both PSAT score and end-of-course standardized

[33] Specifically, Algebra I, Algebra II, Geometry, and either 9th or 10th grade English are the standardized tests administered in Maryland.

test scores suggest a more general reduction in human capital. Finally, we find that mean grade inflation reduces the likelihood of graduating within 5 years by 0.8 percentage points in LAUSD and 0.1 percentage points in Maryland. Five-year graduation rates are much higher in Maryland than in Los Angeles (90 percent and 58 percent, respectively); the smaller effect of mean grade inflation on graduation in Maryland could be explained by differences in the marginal student at risk of not graduating. In Tables A.11 and A.12 we also show that grade inflation increases absences and increases suspensions in both the LAUSD and in Maryland.

Unlike mean grade inflation, we find that passing grade inflation has no effect on future test scores in math or English. However, passing grade inflation appears to be beneficial for some other outcomes. We estimate that being exposed to more passing grade-inflating teachers reduces the likelihood of being held back by 1.2 percentage points in Los Angeles and 0.4 percentage points in Maryland. Being held back is defined as having the same administrative grade code in the current and following year, and is a rare event, with 13 percent of students in our Los Angeles sample and 4 percent of students in our Maryland sample being held back. Thus this is a meaningful effect size, representing about a nine to ten percent decline in the likelihood of repeating a grade. We also estimate that passing grade inflation increases the likelihood of graduating within five years by 0.6 percentage points in Los Angeles and 0.2 percentage points in Los Angeles. We find no effect of passing grade inflation on taking the SAT, but a small reduction in SAT scores of about 3 points in both Los Angeles and Maryland.

In summary, for high school outcomes we find that our more traditional measure of mean grade inflation is reduces educational outcomes, evidenced by lower test scores in the following year and a lower probability of graduating high school. This is consistent with reductions in learning associated with muted incentives that arise from easier grading. In contrast, passing grade inflation offers potential benefits for students by decreasing grade retention and increasing the probability of graduating high school. However, from this evidence alone it is not clear whether these higher rates of high school completion brought about by passing grade inflation are beneficial to students in the longer term. We turn to our analysis of college and early career outcomes to explore this further.

Longer Term Outcomes

Using Maryland's linked K-12, college, and employment data, we are able to explore the impact of mean and passing grade inflation experienced during high school on college enrollment and graduation, earnings, and employment. In this analysis, we estimate the effect of grade inflation on college and career outcomes at different times relative to the students' expected high school graduation in order to capture any dynamic effects of the two grade inflation measures. This means that the sample is composed of different cohorts for different outcome measures. For example, when the outcome is being enrolled in any postsecondary education one year after expected high school graduation, our data includes students who were in 12th grade in 2013 through 2022, whereas when the outcome is measured eight years after expected high school graduation, only 12th-grade cohorts 2013 through 2015 are included.

We find that the effects of grade inflation extend into early adulthood.[34] Table 7 displays these results. We find that mean grade inflation reduces enrollment in any postsecondary schooling by about 0.5 percentage points, but has no measurable effect on the likelihood of graduating from postsecondary schooling. This suggests that the students marginally induced to avoid college enrollment would not have graduated if they had enrolled. We investigate this result further by evaluating the effect of grade inflation on enrollment in two-year and four-year college separately in Table 8. We find that mean grade inflation reduces enrollment in Associate's degree programs from one to seven years past expected high school graduation. For Bachelor's degrees, we estimate reductions in enrollment from three to six years after expected high school graduation, with negative but not statistically significant reductions in the first two years as well. Looking at postsecondary graduation in Table 9, mean grade inflation reduces Associate's degree completion but has no effect on Bachelor's degree completion. Passing grade inflation increases enrollment in two-year schools one to five years after expected high school graduation. The point estimate for passing grade inflation on enrollment in (and graduation from) four-year schools is negative, perhaps because it reduces preparation as measured by the SAT, though it is only (marginally) significant in the fourth year after expected high school graduation. The net effect is that passing grade inflation brings about no change in any postsecondary enrollment, but as Table 7 shows,

[34] This is similar to value-added. For example, Chetty et al. (2014b) showed that test-score value-added in elementary and middle school affected student's college-going and earnings.

passing grade inflation does decrease the likelihood of postsecondary degree completion by about 0.3 percentage points.

Grade inflation also affects students' labor market outcomes, as Tables 7 and 10 show. Mean grade inflation reduces employment one to four and six years after expected high school graduation, whereas passing grade inflation has no measureable effect on employment. Our main measure of employment is having any positive wages in at least one quarter of the year, but in 10 we also impose a more strict criterion of working at least two quarters of the year to exclude those who only work a summer job while in college. We estimate very similar effects of mean grade inflation on employment; exposure to one standard deviation higher mean grade inflation in high school reduces the probability of employment by 0.2 to 0.3 percentage points in the first four years after expected high school graduation, or 0.3 to 0.4 percentage points when our measure is employment for two or more quarters. Passing grade inflation has no statistically significant effect on any employment, but does have a statistically significant positive effect (0.2 to 0.3 percentage points) on employment for two or more quarters in the first four years after expected high school graduation. This is consistent with the evidence we see on enrollment, which suggests passing grade inflation contributes to a small shift away from four-year to two-year programs, where working throughout the school year is more common.

Given the changes in employment due to mean grade inflation, we focus on unconditional earnings (winsorized at the 99th percentile) as our measure of earnings. We find that mean grade inflation reduces earnings by about \$42 to \$133 a year from one through six years after expected high school graduation. We estimate a small positive relationship between passing grade inflation and earnings of similar magnitude to the negative relationship between mean grade inflation and earnings, though it is only statistically significant three and four years after expected high school graduation.

Our exploration of the longer-term impact of mean and passing grade inflation suggests that the negative effects of mean grade-inflating teachers persist into young adulthood. However, passing grade inflation shows more muted long-term effects, with some suggestion of a shift away from four-year college enrollment and toward two-year. Despite the increased probability of graduating high school, passing grade inflation does not result in an appreciable increase in earnings. These results are consistent with the hypothesis that mean grade inflation reduces human capital accumulation. They also suggests that the signaling value of a high school diploma may be minimal, or offset by reductions in human capital as measured by SAT scores.

4.3 Heterogeneity

In the previous sections, we have focused on two forms of grade inflation and have shown they have different average effects on students. However, these effects may be different for different students. For instance, passing grade inflation is more relevant to students who are likely on the margin of passing a class (holding mean grade inflation fixed). Additionally, mean grade inflation might reduce human capital formation to a larger degree for top students who can reduce their effort and still receive an A. On the other hand, the incentive effects induced by mean grade inflation likely apply to all students, since students now need to study less to earn a B as well.

We test the hypothesis that mean and passing grade inflation impact different students differently by estimating Equation 2 on separate samples split by a variety of student and school characteristics. In Tables 11 through 13, we explore whether lower-achieving students are more or less impacted by grade inflation than higher-achieving students. We use 8th grade GPA as a rough proxy for achievement that is, importantly, measured before students enter high school and therefore not impacted by high school teacher grade inflation. Panel C of the table presents p-values from tests of the difference between the below-median and above-median estimates for each measure and outcome.

First, we find that that mean grade inflation reduces future test scores for both above- and below-median students. Our tests, with p-values between 0.38 to 0.97 in Los Angeles and 0.17 to 0.77 in Maryland, give no evidence that the estimated effect of grade inflation on future test scores differs by student achievement. We conclude, therefore, that the reduction in human capital due to mean grade inflation occurs at all parts of the student achievement distribution. Mean grade inflation also does not have a differential effect on the SAT score, but does for the likelihood of taking the SAT. In Los Angeles we estimate that mean grade inflation is associated with larger declines in SAT test-taking for above-median students (p-value: 0.06) while in Maryland we estimate that the declines in SAT test-taking are larger for below-median students (p-value: 0.0001).

Second, we find heterogeneity in the effects of grade inflation on graduation and persistence among students of different abilities. Our tests suggest that the estimated effects of passing grade inflation on grade retention, with p-values of 0.2 in Los Angeles and 0.0001 in Maryland, and graduation, with p-values of 0.01 in Los Angeles and 0.02 in Maryland, are slightly larger for below-median students than for above-median students. This lines up with intuition since above-median students are not close to the passing margin. We also estimate that mean grade inflation reduces grade retention more among below-median students, with p-values of 0.017 in Los Angeles and 0.0002 in Maryland. Looking at longer-term outcomes, the main difference in the effect across the academic achievement distribution is in enrollment in postsecondary education: we find that mean grade inflation reduces enrollment more among below-median students, with p-values from 0.007 to 0.010. However, the effect does not persist into graduation or employment, where we see no statistically different effects of grade inflation for below- and above-median students.

We examine several other dimensions of heterogeneity. First, we explore heterogeneous effects by school-level grading patterns, proxied by school average GPA, to account for different grading contexts. Second, we explore heterogeneity by student characteristics in Maryland.[35] We split the sample by English Language Learner status, Free and Reduced Price Meal status, gender, and race. We do not find strong evidence of heterogeneity across either the school or student characteristics[36].

4.4 Robustness

Carrying out our analysis in both Los Angeles and Maryland provides an effective robustness check for our main results. These two settings differ in many ways, including the demographic makeup of students, curriculum, and the amount of funding, and yet we find similar patterns. Between the two, we have evidence from a large urban school district on the West Coast with a large Hispanic population, and a populous East Coast state with a range of school districts from small rural areas to larger urban areas. We find that mean and passing grade inflation behave similarly in both LAUSD and Maryland, with mean grade inflation reducing high school achievement and passing grade inflation showing evidence of benefits for high school grade retention and graduation. We view this as the strongest evidence that our results are not the result of a particular institutional context.

As a supplement to carrying out our analysis in two settings, we explore the robustness of

[35]The LAUSD data does not contain the rich student demographics necessary for this analysis.

[36]The results of these heterogeneity analyses are all available upon request.

our specification in two additional ways. First, we explore whether our choice of controls impacts our results. In Maryland, we have rich demographic data on students, which we use to augment our controls in both the construction of the value-added and grade inflation measures as well as the estimation of the effects of those measures on high school, college, and career outcomes. As Tables A.15 and A.16 show, the estimates remain very similar when we control for these additional student characteristics.

Second, we explore whether our results are consistent across math and English courses. Our main analysis is based on a combined sample of math and English teachers. We create the value-added and grade inflation measures separately for math and English teachers, then combine them to form a single measure of the average cognitive value-added, noncognitive value-added, and grade inflation experienced by a given student over all math and English courses in a given year (see Section 2 for more details). Because in high school teachers specialize in one subject and not all students take math and English courses in all years, combining the measures across math and English greatly increases our sample size. However, as a robustness check, we also carry out analysis with measures of grade inflation and value-added that are specific to math or English. The main results from these analysis are found in Tables A.8, A.9, and A.10. We find generally similar patterns, but with slightly less precision. The effects of grade inflation do not appear to be driven primarily by either math or English teachers.

By testing the robustness of our results to different school districts, different data, different model specifications, and different subjects, we have shown that the relationship between grade inflation and student outcomes in the short and long term is unlikely to be driven by any particular choice made by the research team.

5 Conclusion

In this paper, we demonstrate that teachers' grading practices affect students in important and heterogeneous ways. A teacher who generally inflates grades has a negative effect on students' future performance as measured by test scores, high school graduation, college enrollment, and earnings. In contrast, a teacher who inflates grades so that students are less likely to fail improves high school graduation rates. The magnitude of these grade inflation effects varies depending on student characteristics and whether a student is on the margin of passing a course.

Our results align with the concept that grades serve as a potent motivator for educational engagement. In light of increasing grade inflation in the United States, our findings suggest that students may be learning less as a result. While our paper does not directly address the change in grading practice over time, our results highlight the potential detrimental impact of increasing grades over time.

Overall, our results demonstrate an additional channel by which teachers affect their students' future outcomes–grading practices. While most teacher training programs that improve teacher quality are high-cost and time-consuming (Taylor and Tyler, 2012), changes in grading policies are a potential low-cost strategy that could be implemented by teachers, schools, and school districts to improve teacher quality and students' long-term outcomes.

References

Alex, P. (2022). Time to pull the plug on traditional grading? *Education Next*, 22(4).

Ayllón, S., Lefgren, L. J., Patterson, R. W., Stoddard, O. B., and Urdaneta, N. (2025). 'sorting' out gender discrimination and disadvantage: Evidence from student evaluations of teaching. Working Paper 33911, National Bureau of Economic Research.

Betts, J. R. and Grogger, J. (2003). The impact of grading standards on student achievement, educational attainment, and entry-level earnings. *Economics of Education Review*, 22(4):343–352.

Black, S. E., Denning, J. T., and Rothstein, J. (2023). Winners and losers? the effect of gaining and losing access to selective colleges on education and labor market outcomes. *American Economic Journal: Applied Economics*, 15(1):26–67.

Bowden, A. B., Rodriguez, V., and Weingarten, Z. (2023). The unintended consequences of academic leniency. *EdWorkingPaper*, pages 23–836.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, 134(3):1163–1224.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American economic review*, 104(9):2593–2632.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American economic review*, 104(9):2633–79.

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

Denning, J. T., Eide, E. R., Mumford, K. J., Patterson, R. W., and Warnick, M. (2022). Why have college completion rates increased? *American Economic Journal: Applied Economics*, 14(3):1–29.

Figlio, D. N. and Lucas, M. E. (2004). Do high grading standards affect student performance? *Journal of Public Economics*, 88(9-10):1815–1834.

Gershenson, S. (2018). Grade inflation in high schools (2005-2016). *Thomas B. Fordham Institute*.

Gershenson, S., Holt, S. B., and Tyner, A. (2022). Making the grade: The effect of teacher grading standards on student outcomes. *Contemporary Economic Policy*.

Graham, K. A. (2017). Does Philly's new grading policy level the playing field or lower standards? — inquirer.com. https://www.inquirer.com/philly/education/

does-phillys-new-grading-policy-level-the-playing-field-or-lower-standards-20170908.html. [Accessed 03-10-2023].

Griffith, D. and Tyner, A. (2025). 'Equitable' Grading Through the Eyes of Teachers. Technical report, Thomas B. Fordham Institute, Washington, DC.

Hurwitz, M. and Lee, J. (2018). Grade inflation and the role of standardized testing. *Measuring success: Testing, grades, and the future of college admissions*, pages 64–93.

Insler, M., McQuoid, A. F., Rahman, A. S., and Smith, K. (2021). Fear and loathing in the classroom: Why does teacher quality matter? Technical report, IZA Discussion Papers.

Jackson, C. K. (2018). What do test scores miss? the importance of teacher effects on non–test score outcomes. *Journal of Political Economy*, 126(5):2072–2107.

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

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

Kane, T. J. and Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. Working Paper 14607, National Bureau of Economic Research.

Koedel, C. and Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, 47:180–195.

Kohn, A. (2002). The dangerous myth of grade inflation. *The chronicle of higher education*, 49(11):B7.

Landaud, F., Maurin, É., Willage, B., and Willén, A. (2024). The value of a high school gpa. *Review of Economics and Statistics*, pages 1–24.

Las Vegas Review-Journal Editorial (2023). EDITORIAL: 'Too dumb to fail' grading policies aren't working — reviewjournal.com. https://www.reviewjournal.com/opinion/editorials/editorial-too-dumb-to-fail-grading-policies-arent-working-2773256/. [Accessed 03-10-2023].

Mozenter, Z. D. (2019). *Essays on the Effects of Teacher Grading Standards and Other Teaching Practices*. PhD thesis, The University of North Carolina at Chapel Hill.

Papageorge, N. W., Gershenson, S., and Kang, K. M. (2020). Teacher expectations matter. *Review of Economics and Statistics*, 102(2):234–251.

Petek, N. and Pope, N. G. (2023). The multidimensional impact of teachers on students. *Journal of*

Political Economy.

Randazzo, S. (2023). Schools Are Ditching Homework, Deadlines in Favor of 'Equitable Grading' — wsj.com. https://www.wsj.com/articles/schools-are-ditching-homework-deadlines-in-favor-of-equitable-grading-dcef7c3e. [Accessed 03-10-2023].

Sanchez, E. I. and Moore, R. (2022). Grade inflation continues to grow in the past decade. *ACT, Inc.*

Staiger, D. O. and Rockoff, J. E. (2010). Searching for effective teachers with imperfect information. *Journal of Economic Perspectives*, 24(3):97–118.

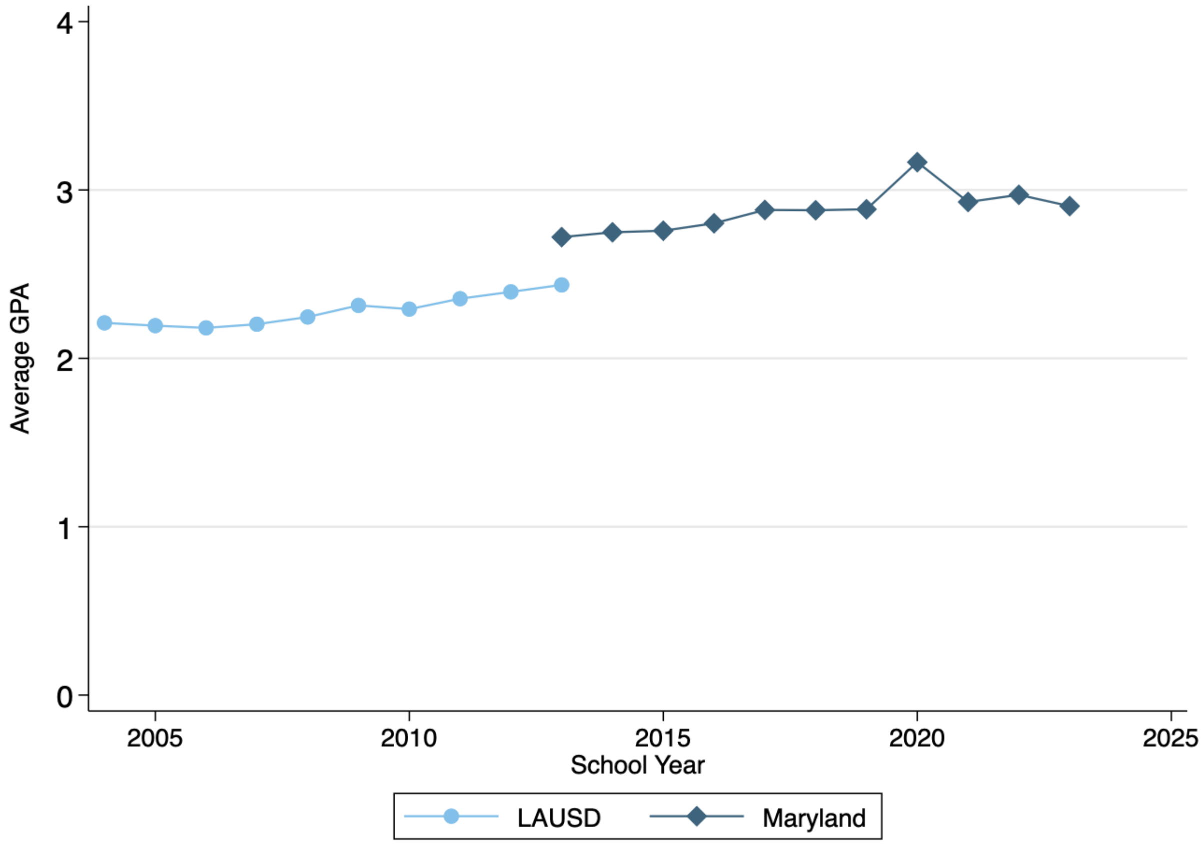
Taylor, E. S. and Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review*, 102(7):3628–3651.

Wright, B. (2019). The dangerous myth of grade inflation. *Thomas Fordham B. Institute*.

Zhang, Q. and Sanchez, E. I. (2013). High school grade inflation from 2004 to 2011. act research report series, 2013 (3). *ACT, Inc.*

6 Tables and Figures

Figure 1: Average GPA over Time



Notes: This figure plots average total GPA over time for students in our Los Angeles and Maryland samples. The data from the LAUSD span 2004 to 2013 and the data from Maryland span 2013 to 2023. GPA is measured on a 0 to 4 scale (e.g., A = 4.0, B = 3.0).

Table 1: Student-level Summary Statistics in Los Angeles

	Mean	Std. Dev.	N
Math CST Score	0.06	1.02	706,922
English CST Score	0.10	1.00	752,431
Math GPA	1.75	1.24	910,194
English GPA	2.18	1.20	985,020
Math Pass Rate	0.68	0.46	986,396
English Pass Rate	0.79	0.41	1,086,116
GPA	2.29	0.98	985,020
Effort GPA	2.16	0.53	985,020
Cooperation GPA	2.39	0.45	985,020
Fraction of Days Absent	0.07	0.09	985,020
Ever Suspended	0.06	0.24	985,020
Held Back	0.12	0.32	985,020
English Learner	0.20	0.40	985,020
Average Teacher Experience	6.48	2.83	985,020
Don't Graduate in LAUSD	0.31	0.46	733,949
Leave Dataset	0.27	0.45	733,949
Graduate on Time	0.49	0.50	833,266
Graduate within 5 Years	0.58	0.49	733,949
Number of AP Courses	1.11	2.04	985,020
Ever Took SAT	0.36	0.48	833,266
Ever Took PSAT	0.43	0.49	646,242
SAT Score	1335.38	299.80	363,022
PSAT Score	1114.21	258.12	393,366
English CAHSEE Score	0.13	0.99	854,094
Math CAHSEE Score	0.12	1.01	856,998
10th Grade Science CST Score	0.12	1.01	656,977
11th Grade Social CST Score	0.11	1.01	674,673

Notes: This table reports summary statistics (mean, standard deviation, and number of observations) for all high school students enrolled in LAUSD between 2004-2013. The sample is restricted to students with non-missing test scores, grades, and behavioral outcomes. CST scores (end of course standardized tests) are standardized before making sample restrictions. "Held back" is measured as whether we see the student with the same administrative grade code in the following year. "Average teacher experience" is the average number of years a student's teachers have taught for. "Don't graduate in LAUSD" and "Leave Dataset" capture the degree to which the graduation outcomes are censored by the sample time period ending or students moving outside of LAUSD. "CAHSEE" refers to the California High School Exit Examination administered from 2006 to 2014.

Table 2: Student-level Summary Statistics in Maryland

	Mean	Std. Dev.	N
Math Test Score	0.03	0.98	322,700
English Test Score	0.04	0.97	356,336
Math GPA	2.59	1.08	1,542,447
English GPA	2.76	1.05	1,647,988
Math Pass Rate	0.97	0.17	1,692,638
English Pass Rate	0.98	0.14	1,786,392
GPA	2.89	0.85	1,647,988
Frac. Days Absent	0.07	0.10	1,647,988
Suspended	0.05	0.21	1,647,988
Held Back	0.04	0.19	1,541,341
English Learner	0.03	0.16	1,647,988
Free and Reduced Price Meals	0.34	0.47	1,647,988
Black	0.34	0.47	1,647,988
White	0.41	0.49	1,647,988
Asian	0.07	0.26	1,647,988
Hispanic	0.14	0.34	1,647,988
Female or Non-Binary	0.50	0.50	1,647,988
AP Classes	3.51	5.04	1,647,988
Honors Classes	8.34	8.16	1,647,988
Ever Took SAT	0.66	0.47	1,587,134
Ever Took SAT or ACT	0.67	0.47	1,647,988
Ever Took PSAT	0.29	0.45	1,596,066
SAT Score	1174.54	315.67	410,547
PSAT Score	128.38	32.43	156,007
Graduate On Time	0.86	0.35	1,587,134
Graduate within 5 Years	0.92	0.27	1,403,078
Enrolled in Higher Ed 1 year Post HS	0.65	0.48	1,403,078
Enrolled in Higher Ed 2 years Post HS	0.58	0.49	1,216,803
Graduated from Higher Ed 4 years Post HS	0.27	0.45	837,086
Graduated from Higher Ed 6 years Post HS	0.41	0.49	461,864
Employed 6 years Post HS	0.68	0.47	1,403,078
Employed 6 years after Expected HS Graduation	0.62	0.48	461,864
Earnings 1 year Post HS	7901.18	7755.60	954,844
Earnings 6 years Post HS	27789.25	21293.59	288,268

Notes: This table reports summary statistics (mean, standard deviation, and number of observations) for all high school students enrolled in Maryland between 2013-2023. The sample is restricted to students with non-missing test scores, grades, and behavioral outcomes. Math tests are from Algebra I, Algebra II, and Geometry courses; English tests are from grades 9 and 10. Test scores are standardized before making sample restrictions. “Held back” is measured as whether we see the student with the same administrative grade code in the following year. The timing of the college and career outcomes are measured relative to expected high school graduation as of 9th grade, so “1 year post HS” refers to 5 years after we see the student in 9th grade. Earnings are conditional on having observed any earnings and are winsorized at the 99th percentile.

Table 3: Forecast Bias Tests of Grade Inflation

	Math		ELA	
Panel A: Individual Test, LAUSD	Course Grade	Pass Indicator	Course Grade	Pass Indicator
Corresponding GI Measure	1.07	1.07	1.03	1.04
	(0.01)	(0.02)	(0.01)	(0.01)
N	474,686	474,839	545,631	545,773
Panel B: Individual Test, Maryland	Course Grade	Pass Indicator	Course Grade	Pass Indicator
Corresponding GI Measure	1.08	1.11	1.03	0.96
	(0.02)	(0.07)	(0.02)	(0.21)
N	204,994	204,994	250,734	250,734
Panel C: Prediction Test, LAUSD	Course Grade	Pass Indicator	Course Grade	Pass Indicator
Corresponding GI Measure	-0.02	-0.01	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
N	538,525	538,542	603,581	603,585

Notes: This table presents results from two different forecast bias tests (the second test is only practical in LAUSD where there are more frequent standardized tests). For the "individual" test, an estimate close to 1 shows that these measures, though subject to estimation error, are actually predictive of student outcomes out of sample. For the 'prediction' test, an estimate close to 0 shows that these measures are not biased by sorting on omitted variables. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Correlations between Grade Inflation and Value-Added Measures

Panel A: LAUSD	Mean GI	Passing GI	Cog. VA	Noncog. VA
Mean GI	1.0000	.	.	.
Passing GI	0.8599	1.0000	.	.
Cog. VA	-0.4070	-0.3047	1.0000	.
Noncog. VA	0.1550	0.1787	0.0879	1.0000
Panel B: Maryland	Mean GI	Passing GI	Cog. VA	Noncog. VA
Mean GI	1.0000	.	.	.
Passing GI	0.3601	1.0000	.	.
Cog. VA	-0.3108	-0.0743	1.0000	.
Noncog. VA	0.1163	0.0542	-0.1241	1.0000

Notes: Within-teacher-year correlations are estimated using an error correction procedure based on randomly splitting the data within classroom. Reported correlations are the average across 200 bootstrapped estimates. The grade inflation and value-added measures used in estimating these correlations are constructed without shrinkage.

Table 5: Regressions of Single Teacher Measures on High School Outcomes

Panel A: LAUSD	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	-0.052***	-0.029***	-0.007***	-0.004**	-0.009***	-11.234***
	(0.007)	(0.003)	(0.001)	(0.002)	(0.001)	(2.629)
Passing GI	-0.043***	-0.022***	-0.009***	-0.001	-0.007***	-10.919***
	(0.007)	(0.003)	(0.001)	(0.001)	(0.001)	(2.824)
Cog. VA	0.114***	0.058***	0.001	0.005*	0.030***	25.004***
	(0.011)	(0.002)	(0.001)	(0.003)	(0.003)	(3.257)
Noncog. VA	0.008	0.016***	-0.006***	0.012***	0.013***	0.594
	(0.007)	(0.003)	(0.001)	(0.002)	(0.002)	(1.508)
Outcome Mean	0.115	0.235	0.106	0.645	0.405	1,368.568
Observations	199,038.	215,559.	411,275.	327,528.	323,635.	173,195.
Panel B: Maryland	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	-0.003	-0.011***	-0.003***	-0.000	-0.004***	-1.115
	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)	(1.244)
Passing GI	-0.003	-0.003	-0.004***	0.001+	-0.000	-3.051**
	(0.004)	(0.003)	(0.001)	(0.001)	(0.002)	(1.335)
Cog. VA	0.016***	0.033***	0.001**	0.001+	0.005***	-0.521
	(0.004)	(0.003)	(0.000)	(0.000)	(0.001)	(1.300)
Noncog. VA	0.007+	0.015***	-0.004***	0.007***	0.003+	-6.423***
	(0.004)	(0.003)	(0.001)	(0.001)	(0.002)	(1.517)
Outcome Mean	-0.20	0.03	0.04	0.90	0.65	1158.08
Observations	123,205	279,802	1,117,227	979,425	1,117,945	252,950

Notes: This table presents estimated coefficients from regressions of the relevant outcome (column) on the relevant teacher measure (row) and our main vector of controls. Thus, each *cell* reports an estimate from a different regression. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Effects of Teacher Grade Inflation and Value-Added on High School Outcomes

Panel A: LAUSD	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	-0.024***	-0.020***	0.002	-0.008***	-0.005+	-1.546
	(0.006)	(0.003)	(0.002)	(0.003)	(0.003)	(1.790)
Passing GI	-0.001	0.003	-0.012***	0.006**	0.003	-3.315+
	(0.006)	(0.003)	(0.002)	(0.003)	(0.002)	(1.896)
Cog. VA	0.113***	0.054***	0.001	-0.000	0.029***	24.748***
	(0.010)	(0.002)	(0.001)	(0.002)	(0.003)	(2.919)
Noncog. VA	-0.004	0.011***	-0.005***	0.011***	0.010***	1.625+
	(0.004)	(0.002)	(0.001)	(0.002)	(0.001)	(0.938)
Outcome Mean	-0.02	0.04	0.13	0.58	0.32	1327.79
Observations	391,782	432,534	832,002	733,946	680,305	186,350
R^2	0.522	0.659	0.167	0.296	0.364	0.738
Panel B: Maryland	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	0.001	-0.006+	-0.001***	-0.001**	-0.004***	-0.427
	(0.004)	(0.003)	(0.000)	(0.000)	(0.001)	(1.262)
Passing GI	-0.003	0.002	-0.004***	0.002**	0.001	-2.953**
	(0.004)	(0.003)	(0.001)	(0.001)	(0.002)	(1.416)
Cog. VA	0.015***	0.031***	0.001**	0.000	0.005***	-0.509
	(0.004)	(0.003)	(0.000)	(0.000)	(0.001)	(1.270)
Noncog. VA	0.006	0.013***	-0.004***	0.007***	0.003	-6.432***
	(0.004)	(0.003)	(0.001)	(0.001)	(0.002)	(1.512)
Outcome Mean	-0.20	0.03	0.04	0.90	0.65	1158.08
Observations	123,205	279,802	1,117,227	979,425	1,117,945	252,950
R^2	0.356	0.592	0.154	0.239	0.242	0.772

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Effects of Teacher Grade Inflation and Value-Added on College and Career Outcomes in Maryland

	Enrolled in Postsecondary		Graduated from Postsecondary		Employed		Unconditional Winz. Earnings	
	1 year	2 years	4 years	6 years	1 year	6 years	1 year	6 years
Mean GI	-0.006***	-0.005***	0.000	-0.002	-0.004***	-0.003**	-49.052***	-165.851***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(12.232)	(51.212)
Passing GI	0.002**	0.001	-0.002**	-0.003**	0.001	0.000	12.637	-10.113
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(11.698)	(41.285)
Cog. VA	0.001	0.001	-0.003***	-0.001	-0.001	-0.001	-33.437***	-47.525
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(12.618)	(41.879)
Noncog. VA	0.007***	0.006***	-0.001	0.001	0.002***	0.002**	32.501***	40.429
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(9.593)	(44.426)
Outcome Mean	0.64	0.56	0.27	0.40	0.68	0.62	5444.70	17801.59
Observations	979,425	838,616	552,566	274,354	979,425	274,354	979,425	274,354
R^2	0.287	0.279	0.237	0.338	0.037	0.050	0.080	0.041

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. The outcomes in this table are only available in the Maryland data, not in LAUSD. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Dynamic Effects of Grade Inflation on College Enrollment in Maryland

Panel A: Enrolled in Associate's X Years Post Expected HS Graduation								
	$X=1$	$X=2$	$X=3$	$X=4$	$X=5$	$X=6$	$X=7$	$X=8$
Mean GI	-0.004***	-0.004***	-0.003***	-0.002***	-0.001***	-0.001***	-0.001**	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Passing GI	0.003***	0.002***	0.001**	0.001**	0.001+	0.000	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Outcome Mean	0.24	0.21	0.15	0.10	0.07	0.05	0.04	0.03
Observations	1,403,076	1,216,801	1,026,593	837,085	652,118	461,860	276,317	135,233
R^2	0.032	0.031	0.022	0.015	0.010	0.009	0.007	0.008
Panel B: Enrolled in Bachelor's X Years Post Expected HS Graduation								
	$X=1$	$X=2$	$X=3$	$X=4$	$X=5$	$X=6$	$X=7$	$X=8$
Mean GI	-0.001	-0.001+	-0.002**	-0.002**	-0.003***	-0.002***	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Passing GI	-0.001	-0.001	-0.001	-0.001+	0.000	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
Outcome Mean	0.22	0.22	0.24	0.25	0.13	0.07	0.05	0.03
Observations	1,403,076	1,216,801	1,026,593	837,085	652,118	461,860	276,317	135,233
R^2	0.137	0.144	0.155	0.145	0.032	0.012	0.008	0.007
Panel C: Enrolled in Any College X Years Post Expected HS Graduation								
	$X=1$	$X=2$	$X=3$	$X=4$	$X=5$	$X=6$	$X=7$	$X=8$
Mean GI	-0.005***	-0.004***	-0.005***	-0.005***	-0.005***	-0.003***	-0.002***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Passing GI	0.001	0.000	0.000	-0.000	0.001	0.000	-0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Outcome Mean	0.65	0.58	0.50	0.49	0.29	0.16	0.12	0.10
Observations	1,403,076	1,216,801	1,026,593	837,085	652,118	461,860	276,317	135,233
R^2	0.285	0.278	0.254	0.268	0.069	0.025	0.017	0.014

Notes: This table presents estimates from regressions of being enrolled in the relevant postsecondary degree program (panel), measured X years (column) after expected high school graduation, on mean grade inflation and passing grade inflation as well as cognitive value-added, noncognitive value-added, and our main vector of controls. The outcomes in this table are only available in the Maryland data, not in LAUSD. The timing of enrollment is measured relative to expected high school graduation as of 9th grade; for example, "1 year post HS" refers to 5 years after we see the student in 9th grade. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Dynamic Effects of Grade Inflation on College Graduation in Maryland

Panel A: Graduated with Associate's X Years Post Expected HS Graduation			
	$X = 4$	$X = 5$	$X = 6$
Mean GI	-0.001***	-0.002***	-0.002***
	(0.000)	(0.001)	(0.001)
Passing GI	0.001	0.001+	0.001
	(0.000)	(0.000)	(0.001)
Outcome Mean	0.08	0.09	0.10
Observations	837,085	652,118	461,860
R^2	0.044	0.041	0.041
Panel B: Graduated with Bachelor's X Years Post Expected HS Graduation			
	$X = 4$	$X = 5$	$X = 6$
Mean GI	0.001	0.000	-0.000
	(0.001)	(0.001)	(0.001)
Passing GI	-0.002***	-0.003**	-0.003**
	(0.001)	(0.001)	(0.001)
Outcome Mean	0.20	0.31	0.35
Observations	837,085	652,118	461,860
R^2	0.242	0.325	0.341
Panel C: Graduated from Any College X Years Post Expected HS Graduation			
	$X = 4$	$X = 5$	$X = 6$
Mean GI	-0.000	-0.001	-0.002
	(0.001)	(0.001)	(0.001)
Passing GI	-0.002+	-0.002	-0.002+
	(0.001)	(0.001)	(0.001)
Outcome Mean	0.27	0.37	0.41
Observations	837,085	652,118	461,860
R^2	0.239	0.324	0.339

Notes: This table presents estimates from regressions of graduating from the relevant postsecondary degree program (panel), measured X years (column) after expected high school graduation, on mean grade inflation and passing grade inflation as well as cognitive value-added, noncognitive value-added, and our main vector of controls. The outcomes in this table are only available in the Maryland data, not in LAUSD. The timing of college graduation is measured relative to expected high school graduation as of 9th grade; for example, "1 year post HS" refers to 5 years after we see the student in 9th grade. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Dynamic Effects of Grade Inflation on Labor Market Outcomes in Maryland

Panel A: Unconditional Winz. Earnings X Years Post Expected HS Graduation								
	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$	$X = 7$	$X = 8$
Mean GI	-42.286***	-47.760***	-71.814***	-74.685***	-50.079+	-133.192***	-89.455+	-32.829
	(10.820)	(15.545)	(19.821)	(23.099)	(27.227)	(40.489)	(46.940)	(79.257)
Passing GI	13.452	22.284	47.988**	54.321**	19.158	4.432	-7.119	63.311
	(12.857)	(16.902)	(19.031)	(24.755)	(27.346)	(38.167)	(45.439)	(89.074)
Outcome Mean	5377.04	7486.49	8765.06	10121.78	14222.43	17344.43	19799.54	22105.47
Observations	1,403,076	1,216,801	1,026,593	837,085	652,118	461,860	276,317	135,233
R^2	0.083	0.084	0.079	0.069	0.038	0.043	0.045	0.039
Panel B: Employed 2 or More Quarters X Years Post Expected HS Graduation								
	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$	$X = 7$	$X = 8$
Mean GI	-0.003***	-0.003***	-0.004***	-0.003***	-0.001	-0.002**	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Passing GI	0.002**	0.002**	0.003***	0.002**	0.001	0.001	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Outcome Mean	0.40	0.45	0.46	0.45	0.48	0.49	0.49	0.48
Observations	1,403,076	1,216,801	1,026,593	837,085	652,118	461,860	276,317	135,233
R^2	0.060	0.066	0.070	0.066	0.046	0.048	0.051	0.049
Panel C: Any Employment X Years Post Expected HS Graduation								
	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$	$X = 7$	$X = 8$
Mean GI	-0.003***	-0.002***	-0.002***	-0.002***	-0.001	-0.002**	-0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Passing GI	0.001	0.001	0.001	0.001	0.000	0.000	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Outcome Mean	0.68	0.70	0.68	0.65	0.65	0.62	0.61	0.59
Observations	1,403,076	1,216,801	1,026,593	837,085	652,118	461,860	276,317	135,233
R^2	0.039	0.037	0.041	0.045	0.044	0.051	0.055	0.056

Notes: This table presents estimates from regressions of labor market outcomes (panel), measured X years (column) after expected high school graduation, on mean grade inflation and passing grade inflation as well as cognitive value-added, noncognitive value-added, and our main vector of controls. The outcomes in this table are only available in the Maryland data, not in LAUSD. Earnings are winsorized at the 99th percentile. The timing of college graduation is measured relative to expected high school graduation as of 9th grade; for example, "1 year post HS" refers to 5 years after we see the student in 9th grade. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Effects of Teacher Measures Among Low- and High-Achieving Students in LAUSD

Panel A: Below Median	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	-0.026***	-0.022***	-0.004	-0.007***	-0.000	-0.535
	(0.006)	(0.005)	(0.003)	(0.003)	(0.002)	(1.535)
Passing GI	0.001	0.003	-0.012***	0.007***	0.001	-4.821***
	(0.005)	(0.005)	(0.003)	(0.002)	(0.002)	(1.705)
Outcome Mean	-0.33	-0.31	0.23	0.47	0.14	1213.85
Observations	161,249	181,382	343,437	263,059	272,601	55,051
R^2	0.324	0.522	0.164	0.289	0.223	0.651
Panel B: Above Median	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	-0.026***	-0.018***	0.002+	-0.005**	-0.005+	-0.955
	(0.007)	(0.003)	(0.001)	(0.002)	(0.003)	(1.854)
Passing GI	0.001	0.004	-0.009***	0.001	0.003	-3.341+
	(0.007)	(0.003)	(0.001)	(0.002)	(0.003)	(1.929)
Outcome Mean	0.27	0.40	0.05	0.81	0.51	1371.51
Observations	189,020	202,392	337,968	283,415	268,546	178,682
R^2	0.568	0.683	0.082	0.181	0.287	0.752
Panel C: P-values from Tests of Coefficient Equality	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	0.9677	0.3786	0.0178	0.3330	0.0562	0.8006
Passing GI	0.9658	0.8216	0.2015	0.0103	0.3594	0.3417

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. The sample is split into students whose 8th grade GPA is below and above median as a proxy for low- and high-achieving students. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. Panel C shows p-values of tests for coefficient equality that were constructed using seemingly unrelated regressions. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: Effects of Teacher Measures Among Low- and High-Achieving Students in Maryland

Panel A: Below Median						
	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	-0.017***	-0.007**	-0.003***	-0.000	-0.004**	-2.751***
	(0.005)	(0.003)	(0.001)	(0.001)	(0.002)	(0.963)
Passing GI	-0.003	0.002	-0.005***	0.002**	0.001	-1.162
	(0.003)	(0.004)	(0.002)	(0.001)	(0.002)	(0.865)
Outcome Mean	-0.33	-0.36	0.08	0.87	0.51	958.20
Observations	130,241	128,690	461,942	527,128	462,508	71,299
R^2	0.524	0.481	0.160	0.249	0.203	0.645
Panel B: Above Median						
	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	-0.016***	-0.005	-0.001**	-0.000	-0.003**	1.190
	(0.005)	(0.004)	(0.000)	(0.000)	(0.001)	(1.118)
Passing GI	0.001	0.001	-0.002+	0.001	0.002	-1.250
	(0.005)	(0.004)	(0.001)	(0.001)	(0.002)	(1.027)
Outcome Mean	0.45	0.41	0.01	0.96	0.77	1146.36
Observations	134,573	137,366	497,777	557,599	497,894	116,390
R^2	0.624	0.604	0.125	0.147	0.156	0.707
Panel C: P-values from Tests of Coefficient Equality						
	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	0.1690	0.5521	0.0002	0.8490	0.0001	0.7023
Passing GI	0.6052	0.7727	0.0001	0.0220	0.9266	0.2160

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. The sample is split into students whose 8th grade GPA is below and above median as a proxy for low- and high-achieving students. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. Panel C shows p-values of tests for coefficient equality that were constructed using seemingly unrelated regressions. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: Effects of Teacher Measures Among Low- and High-Achieving Students in Maryland

Panel A: Below Median

	Enrolled in Postsecondary		Graduated from Postsecondary		Employed		Unconditional Winz. Earnings	
	1 year	2 years	4 years	6 years	1 year	6 years	1 year	6 years
Mean GI	-0.005***	-0.005***	-0.001	-0.002	-0.004***	-0.001	-57.674***	-92.938
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(16.326)	(84.959)
Passing GI	0.000	-0.001	-0.001+	-0.004***	0.001	0.003+	15.162	85.255
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(15.274)	(66.475)
Outcome Mean	0.50	0.42	0.13	0.24	0.69	0.66	6468.54	18343.79
Observations	527,128	436,928	261,427	85,800	394,905	64,182	394,905	64,182
R^2	0.255	0.240	0.152	0.256	0.035	0.041	0.060	0.045

Panel B: Above Median

	Enrolled in Postsecondary		Graduated from Postsecondary		Employed		Unconditional Winz. Earnings	
	1 year	2 years	4 years	6 years	1 year	6 years	1 year	6 years
Mean GI	-0.002**	-0.002	-0.001	-0.001	-0.003***	-0.002	-33.472**	-150.227
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(15.764)	(103.565)
Passing GI	0.002+	0.000	-0.001	-0.002	0.000	-0.001	-2.636	-27.763
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(15.197)	(76.457)
Outcome Mean	0.78	0.69	0.39	0.58	0.67	0.59	4813.85	19676.29
Observations	557,599	461,734	270,874	89,935	427,243	68,924	427,243	68,924
R^2	0.205	0.217	0.196	0.306	0.044	0.076	0.096	0.044

Panel C: P-values from Tests of Coefficient Equality

	Enrolled in Postsecondary		Graduated from Postsecondary		Employed		Unconditional Winz. Earnings	
	1 year	2 years	4 years	6 years	1 year	6 years	1 year	6 years
Mean GI	0.0066	0.0095	0.5658	0.9195	0.3386	0.5940	0.2302	0.6388
Passing GI	0.1059	0.2680	0.8109	0.4088	0.4917	0.0892	0.3562	0.2403

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. The sample is split into students whose 8th grade GPA is below and above median as a proxy for low- and high-achieving students. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. Panel C shows p-values of tests for coefficient equality that were constructed using seemingly unrelated regressions. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1: Student-level Summary Statistics (Sample for Separate Math and English Measures)

	Mean	Std. Dev.	N
Math CST Score	0.18	1.08	364,658
English CST Score	0.30	1.00	376,540
Math GPA	1.93	1.27	468,907
English GPA	2.33	1.22	468,907
GPA	2.44	0.98	468,907
Effort GPA	2.24	0.53	468,907
Cooperation GPA	2.45	0.44	468,907
Fraction of Days Absent	0.06	0.09	468,907
Ever Suspended	0.05	0.22	468,907
Held Back	0.10	0.29	468,907
English Learner	0.14	0.35	468,907
Average Teacher Experience	6.79	2.81	468,907
Don't Graduate in LAUSD	0.27	0.44	327,535
Leave Dataset	0.24	0.43	327,535
Graduate on Time	0.54	0.50	381,249
Graduate within 5 Years	0.65	0.48	327,535
Number of AP Courses	1.54	2.38	468,907
Ever Took SAT	0.45	0.50	381,249
Ever Took PSAT	0.45	0.50	328,059
SAT Score	1372.40	300.98	212,460
PSAT Score	1161.57	263.83	218,950
English CAHSEE Score	0.32	0.97	414,660
Math CAHSEE Score	0.32	1.02	415,862
10th Grade Science CST Score	0.30	1.04	334,061
11th Grade Social CST Score	0.28	1.02	329,346

Notes: This table reports summary statistics (mean, standard deviation, and number of observations) for all high school students enrolled in LAUSD between 2004-2013. The sample is restricted to students with non-missing test scores, grades, and behavioral outcomes and requires students to have taken both math and English in a given year. CST scores (end of course standardized tests) are standardized before making sample restrictions. "Held back" is measured as whether we see the student with the same administrative grade code in the following year. "Average teacher experience" is the average number of years a student's teachers have taught for. "Don't graduate in LAUSD" and "Leave Dataset" capture the degree to which the graduation outcomes are censored by the sample time period ending or students moving outside of LAUSD. "CAHSEE" refers to the California High School Exit Examination administered from 2006 to 2014.

Table A.2: Student-level Summary Statistics, Maryland (Sample for Separate Math and English Measures)

	Mean	Std. Dev.	N
Math Test Score	0.06	0.98	151,929
English Test Score	0.07	0.94	174,079
Math GPA	2.54	1.09	631,853
English GPA	2.74	1.06	667,585
GPA	2.90	0.84	667,585
Frac. Days Absent	0.07	0.09	667,585
Suspended	0.05	0.21	667,585
Held Back	0.03	0.16	636,207
English Learner	0.02	0.13	667,585
Free and Reduced Price Meals	0.34	0.47	667,585
Black	0.36	0.48	667,585
White	0.42	0.49	667,585
Asian	0.07	0.25	667,585
Hispanic	0.12	0.32	667,585
Female or Non-Binary	0.52	0.50	667,585
AP Classes	3.07	4.36	667,585
Honors Classes	6.90	6.76	667,585
Ever Took SAT	0.68	0.46	652,155
Ever Took SAT or ACT	0.69	0.46	667,585
Ever Took PSAT	0.29	0.45	656,237
SAT Score	1163.84	302.15	166,923
PSAT Score	127.18	30.99	73,049
Graduate On Time	0.88	0.32	652,155
Graduate within 5 Years	0.93	0.25	588,912
Enrolled in Higher Ed 1 year Post HS	0.66	0.47	588,912
Enrolled in Higher Ed 2 years Post HS	0.58	0.49	520,522
Graduated from Higher Ed 4 years Post HS	0.27	0.44	359,479
Graduated from Higher Ed 6 years Post HS	0.41	0.49	189,756
Employed 6 years Post HS	0.69	0.46	588,912
Employed 6 years after Expected HS Graduation	0.63	0.48	189,756
Earnings 1 year Post HS	7945.15	7708.00	407,761
Earnings 6 years Post HS	28132.63	21325.57	120,397

Notes: This table reports summary statistics (mean, standard deviation, and number of observations) for all high school students enrolled in Maryland between 2013-2023. The sample is restricted to students with non-missing test scores, grades, and behavioral outcomes and requires students to have taken both math and English in a given year. Math tests are from Algebra I, Algebra II, and Geometry courses; English tests are from grades 9 and 10. Test scores are standardized before making sample restrictions. "Held back" is measured as whether we see the student with the same administrative grade code in the following year. The timing of the college and career outcomes are measured relative to expected high school graduation as of 9th grade, so "1 year post HS" refers to 5 years after we see the student in 9th grade. Earnings are conditional on having observed any earnings and are winsorized at the 99th percentile.

Table A.3: Correlations between Grade Inflation and Value-Added Measures in LAUSD (Math Teachers)

	Mean GI	Passing GI	Cog. VA	GPA Total VA	Effort GPA VA	Coop. GPA VA	Frac. Absent VA	Held Back VA	Suspended VA
Mean GI	1.0000	.	.	.	.	.	.	.	.
Passing GI	0.8863	1.0000	.	.	.	.	.	.	.
Cog. VA	-0.4289	-0.3492	1.0000	.	.	.	.	.	.
GPA Total VA	0.2610	0.2902	0.0166	1.0000	.	.	.	.	.
Effort GPA VA	0.0698	0.1290	0.1881	0.8296	1.0000	.	.	.	.
Coop. GPA VA	-0.1749	-0.0772	0.2790	0.5862	0.8293	1.0000	.	.	.
Frac. Absent VA	0.0882	0.0780	-0.0525	0.0974	-0.1770	-0.2266	1.0000	.	.
Held Back VA	0.0597	0.0194	-0.0177	-0.0485	-0.0689	-0.1040	0.0463	1.0000	.
Suspended VA	-0.0046	0.0037	-0.0732	-0.3544	-0.3381	-0.2945	0.0754	-0.0775	1.0000

Notes: Within-teacher-year correlations are estimated using an error correction procedure based on randomly splitting the data within classroom. Reported correlations are the average across 200 bootstrapped estimates. The grade inflation and value-added measures used in estimating these correlations are constructed without shrinkage. This sample includes only math teachers.

Table A.4: Correlations between Grade Inflation and Value-Added Measures in LAUSD (English Teachers)

	Mean GI	Passing GI	Cog. VA	GPA Total VA	Effort GPA VA	Coop. GPA VA	Frac. Absent VA	Held Back VA	Suspended VA
Mean GI	1.0000	.	.	.	.	.	.	.	.
Passing GI	0.8379	1.0000	.	.	.	.	.	.	.
Cog. VA	-0.3747	-0.2431	1.0000	.	.	.	.	.	.
GPA Total VA	0.2526	0.2518	0.0333	1.0000	.	.	.	.	.
Effort GPA VA	0.1819	0.2008	0.1641	0.8424	1.0000	.	.	.	.
Coop. GPA VA	0.0213	0.1020	0.3720	0.5920	0.8228	1.0000	.	.	.
Frac. Absent VA	0.0879	0.0407	-0.1814	0.0710	-0.1826	-0.2101	1.0000	.	.
Held Back VA	0.0498	0.0063	-0.1324	-0.0521	-0.0782	-0.1023	0.0806	1.0000	.
Suspended VA	0.0171	0.0168	-0.1531	-0.3366	-0.3128	-0.3074	0.0636	-0.0538	1.0000

Notes: Within-teacher-year correlations are estimated using an error correction procedure based on randomly splitting the data within classroom. Reported correlations are the average across 200 bootstrapped estimates. The grade inflation and value-added measures used in estimating these correlations are constructed without shrinkage. This sample includes only English teachers.

Table A.5: Correlations between Grade Inflation and Value-Added Measures in Maryland (Math Teachers)

	Mean GI	Passing GI	Cog. VA	GPA VA	Frac. Absent VA	Held Back VA	Suspended VA
Mean GI	1.0000	.	.	.	.	.	.
Passing GI	0.3850	1.0000	.	.	.	.	.
Cog. VA	-0.3316	-0.1282	1.0000	.	.	.	.
GPA VA	0.0168	0.2013	0.1377	1.0000	.	.	.
Frac Absent VA	0.1961	-0.1912	-0.2528	-0.3812	1.0000	.	.
Held Back VA	0.0906	-0.0457	-0.1426	-0.2121	0.3704	1.0000	.
Suspended VA	0.0667	0.1637	-0.0952	-0.0586	-0.0896	-0.1802	1.0000

Notes: Within-teacher-year correlations are estimated using an error correction procedure based on randomly splitting the data within classroom. Reported correlations are the average across 200 bootstrapped estimates. The grade inflation and value-added measures used in estimating these correlations are constructed without shrinkage. This sample includes only math teachers.

Table A.6: Correlations between Grade Inflation and Value-Added Measures in Maryland (English Teachers)

	Mean GI	Passing GI	Cog. VA	GPA VA	Frac. Absent VA	Held Back VA	Suspended VA
Mean GI	1.0000	.	.	.	.	.	.
Passing GI	0.3309	1.0000	.	.	.	.	.
Cog. VA	-0.2961	-0.0104	1.0000	.	.	.	.
GPA VA	0.0500	0.1934	0.0810	1.0000	.	.	.
Frac Absent VA	0.0983	-0.2060	-0.2042	-0.3634	1.0000	.	.
Held Back VA	0.0220	-0.0988	-0.0702	-0.2486	0.3711	1.0000	.
Suspended VA	0.0615	0.1802	-0.1169	-0.0216	-0.1231	-0.1791	1.0000

Notes: Within-teacher-year correlations are estimated using an error correction procedure based on randomly splitting the data within classroom. Reported correlations are the average across 200 bootstrapped estimates. The grade inflation and value-added measures used in estimating these correlations are constructed without shrinkage. This sample includes only English teachers.

Table A.7: Regressions of Single Teacher Measures on College and Career Outcomes in Maryland

	Enrolled in Postsecondary		Graduated from Postsecondary		Employed		Log Conditional Earnings		Unconditional Winz. Earnings	
	1 year	2 years	4 years	6 years	1 year	6 years	1 year	6 years	1 year	6 years
Mean GI	-0.005***	-0.005***	0.000	-0.003***	-0.003***	-0.002**	-0.003	-0.006**	-39.118***	-163.664***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(12.443)	(44.686)
Passing GI	-0.000	-0.001	-0.001+	-0.003***	-0.001	-0.000	0.003	-0.002	-4.630	-71.329**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(11.847)	(35.540)
Cog. VA	0.002***	0.002**	-0.003***	-0.000	-0.000	-0.001	-0.006***	-0.002	-25.438**	-18.649
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(12.737)	(42.155)
Noncog. VA	0.007***	0.006***	-0.001	0.001	0.002***	0.002**	0.006***	0.006+	31.558***	39.826
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(9.536)	(44.529)
Outcome Mean	0.64	0.56	0.27	0.40	0.68	0.62	8.39	9.78	5444.70	17801.59
Observations	979,425	838,616	552,566	274,354	979,425	274,354	664,075	170,840	979,425	274,354

Notes: This table presents estimated coefficients from regressions of the relevant outcome (column) on the relevant teacher measure (row) and our main vector of controls. Thus, each *cell* reports an estimate from a different regression. The outcomes in this table are only available in the Maryland data, not in LAUSD. The timing of enrollment is measured relative to expected high school graduation as of 9th grade; for example, "1 year post HS" refers to 5 years after we see the student in 9th grade. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Effects of Subject-Specific Teacher Measures on High School Outcomes in Los Angeles

	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI (math)	-0.046***	-0.022***	0.000	-0.011***	-0.009**	-1.129
	(0.010)	(0.005)	(0.003)	(0.003)	(0.004)	(2.369)
Mean GI (ela)	-0.005	-0.011***	0.003	-0.006	0.001	0.679
	(0.008)	(0.004)	(0.002)	(0.004)	(0.004)	(2.131)
Passing GI (math)	0.011	0.007	-0.011***	0.006+	0.003	-2.459
	(0.009)	(0.005)	(0.003)	(0.004)	(0.003)	(1.960)
Passing GI (ela)	-0.001	-0.006	-0.012***	0.005	-0.001	-6.945**
	(0.009)	(0.005)	(0.002)	(0.004)	(0.004)	(2.644)
Cog. VA (math)	0.126***	0.017***	0.002+	-0.010***	0.011***	21.299***
	(0.013)	(0.003)	(0.001)	(0.003)	(0.004)	(2.862)
Cog. VA (ela)	0.058***	0.066***	-0.002	0.013***	0.033***	19.492***
	(0.008)	(0.003)	(0.001)	(0.002)	(0.002)	(3.287)
Noncog. VA (math)	-0.005	0.005***	-0.003***	0.007***	0.007***	-0.306
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.605)
Noncog. VA (ela)	0.004	0.008***	-0.004***	0.005***	0.004***	0.492
	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(1.166)
Outcome Mean	0.11	0.23	0.11	0.65	0.40	1368.58
Observations	199,038	215,559	411,275	327,528	323,635	173,195
R^2	0.564	0.679	0.160	0.307	0.379	0.746

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Effects of Subject-Specific Teacher Measures on High School Outcomes in Maryland

	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI (math)	0.001	-0.007	-0.002**	-0.001	-0.003	-0.502
	(0.008)	(0.005)	(0.001)	(0.001)	(0.003)	(2.372)
Mean GI (ela)	-0.000	-0.001	-0.001**	-0.001	-0.004+	-0.138
	(0.007)	(0.007)	(0.001)	(0.001)	(0.003)	(2.308)
Passing GI (math)	-0.010	-0.003	-0.007***	0.004**	0.004	-8.151**
	(0.012)	(0.007)	(0.002)	(0.002)	(0.004)	(4.003)
Passing GI (ela)	-0.013	0.009	-0.010**	0.002	0.008	-2.334
	(0.015)	(0.013)	(0.005)	(0.003)	(0.009)	(4.603)
Cog. VA (math)	0.004	0.036***	-0.001	0.002+	0.008**	14.186***
	(0.011)	(0.007)	(0.001)	(0.001)	(0.003)	(3.911)
Cog. VA (ela)	0.034***	0.043***	0.002***	-0.001	0.006+	-11.621***
	(0.006)	(0.006)	(0.001)	(0.001)	(0.003)	(1.779)
Noncog. VA (math)	0.002	0.009**	-0.002***	0.004***	0.000	-3.382**
	(0.005)	(0.004)	(0.000)	(0.000)	(0.002)	(1.379)
Noncog. VA (ela)	0.010+	0.002	-0.002***	0.004***	0.001	-3.843***
	(0.005)	(0.004)	(0.001)	(0.001)	(0.002)	(1.373)
Outcome Mean	-0.19	0.05	0.03	0.92	0.68	1150.20
Observations	65,988	130,701	491,914	444,650	492,054	111,200
R^2	0.383	0.578	0.145	0.209	0.239	0.760

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Effects of Subject-Specific Teacher Measures on College and Career Outcomes in Maryland

	Enrolled in Postsecondary		Graduated from Postsecondary		Employed		Log Conditional Earnings		Unconditional Winz. Earnings	
	1 year	2 years	4 years	6 years	1 year	6 years	1 year	6 years	1 year	6 years
Mean GI (math)	-0.006***	-0.004**	0.001	0.000	-0.005***	-0.005**	-0.010**	0.008	-74.643***	-110.426
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.005)	(0.007)	(24.071)	(108.179)
Mean GI (ela)	-0.006***	-0.005***	0.002	-0.002	-0.003**	-0.001	-0.005	-0.010+	-44.788**	-181.627**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.004)	(0.006)	(21.835)	(81.448)
Passing GI (math)	0.002	-0.002	-0.005**	-0.007**	0.002	0.003	0.010+	-0.005	45.384	-65.554
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.006)	(0.010)	(34.094)	(125.684)
Passing GI (ela)	0.005	0.005	-0.001	-0.003	0.000	-0.003	0.013	-0.019	19.245	-136.213
	(0.004)	(0.004)	(0.004)	(0.005)	(0.002)	(0.004)	(0.008)	(0.014)	(43.935)	(247.180)
Cog. VA (math)	-0.002	-0.003	-0.003	-0.003	-0.003	-0.007***	-0.004	-0.005	-10.622	-227.701**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.007)	(0.008)	(38.391)	(110.234)
Cog. VA (ela)	0.002+	0.002+	-0.004***	0.000	-0.001	0.000	-0.015***	-0.001	-79.798***	46.044
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.006)	(16.232)	(80.076)
Noncog. VA (math)	0.004***	0.004***	0.000	0.000	0.001	0.001	0.006***	0.003	22.707**	-14.261
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(11.518)	(53.920)
Noncog. VA (ela)	0.004***	0.004***	-0.001	0.000	0.000	0.001	-0.001	0.003	-3.207	57.061
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(13.620)	(57.590)
Outcome Mean	0.65	0.57	0.26	0.40	0.69	0.63	8.40	9.79	5546.57	18175.75
Observations	444,650	388,229	260,877	127,076	444,650	127,076	307,329	80,360	444,650	127,076
R^2	0.269	0.268	0.230	0.329	0.032	0.048	0.089	0.049	0.078	0.041

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. The outcomes in this table are only available in the Maryland data, not in LAUSD. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Effects of Teacher Measures on Additional High School Outcomes in Los Angeles

	Future Held Back	Future Frac. Days Absent	Future Suspension	Graduate On Time	Don't Graduate	Leave Dataset Next Year	PSAT Score	Took PSAT
Mean GI	0.003^{+}	0.001***	0.001**	-0.007**	0.009***	0.002	-3.956**	-0.004
	(0.002)	(0.000)	(0.001)	(0.003)	(0.002)	(0.001)	(1.629)	(0.003)
Passing GI	-0.004***	-0.001^{+}	0.000	0.004	-0.008***	-0.004***	2.235	0.002
	(0.002)	(0.000)	(0.001)	(0.003)	(0.002)	(0.001)	(1.784)	(0.004)
Cog. VA	0.002**	0.000	-0.001	-0.001	0.002	-0.001	22.855***	0.012***
	(0.001)	(0.000)	(0.000)	(0.003)	(0.002)	(0.001)	(3.083)	(0.003)
Noncog. VA	-0.011***	-0.002***	-0.004***	0.012***	-0.012***	0.001	2.441***	-0.003
	(0.001)	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)	(0.806)	(0.003)
Outcome Mean	0.13	0.08	0.05	0.49	0.37	0.08	1097.94	0.40
Observations	630,184	731,638	763,350	833,265	597,851	832,002	261,265	562,777
R^2	0.129	0.299	0.057	0.284	0.306	0.085	0.727	0.338

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Effects of Teacher Measures on Additional High School Outcomes in Maryland

	Future Held Back	Future Frac. Days Absent	Future Suspension	Graduate On Time	PSAT Score	Ever Took PSAT	Ever Took SAT or ACT
Mean GI	0.001***	0.001***	0.002***	-0.001**	-0.015	-0.003+	-0.004***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.128)	(0.001)	(0.001)
Passing GI	-0.001	0.000	-0.000	0.002**	-0.194	0.001	0.002
	(0.001)	(0.000)	(0.000)	(0.001)	(0.130)	(0.001)	(0.002)
Cog. VA	-0.000	-0.000	0.000	0.000	0.163+	-0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.098)	(0.001)	(0.001)
Noncog. VA	-0.004***	-0.001***	-0.004***	0.010***	-0.742***	-0.001	0.005***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.143)	(0.001)	(0.002)
Outcome Mean	0.04	0.08	0.04	0.84	128.39	0.25	0.65
Observations	1,059,643	1,059,391	1,059,643	1,117,945	155,708	1,126,690	1,178,611
R^2	0.108	0.366	0.065	0.342	0.738	0.645	0.327

Notes: This table presents estimates from regressions of the relevant outcome (column) on mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added, as well as our main vector of controls. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: Forecast Bias Tests of Value-Added in the LAUSD

Panel A: Individual Test, Math	Test Score	GPA	Effort GPA	Coop GPA	Frac Days Absent	Suspended	Heldback
Corresponding VA Measure	1.11	0.68	0.85	1.12	-0.63	-1.79	-1.99
	(0.02)	(0.05)	(0.06)	(0.05)	(0.09)	(0.15)	(0.13)
N	479,237	701,518	699,042	699,043	691,505	701,698	619,341
Panel B: Individual Test, ELA	Test Score	GPA	Effort GPA	Coop GPA	Frac Days Absent	Suspended	Heldback
Corresponding VA Measure	1.16	0.59	0.68	0.92	-0.29	-1.77	-1.68
	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.20)	(0.16)
N	548,508	778,299	775,450	775,448	766,664	779,165	688,482
Panel C: Prediction Test, Math	Test Score	GPA	Effort GPA	Coop GPA	Frac Days Absent	Suspended	Heldback
Corresponding VA Measure	0.10	0.00	0.01	0.01	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	538,602	541,501	539,501	539,501	541,102	541,564	540,674
Panel D: Prediction Test, ELA	Test Score	GPA	Effort GPA	Coop GPA	Frac Days Absent	Suspended	Heldback
Corresponding VA Measure	0.26	0.01	0.01	0.01	0.00	-0.00	-0.00
	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	603,609	607,046	604,596	604,596	605,768	607,479	603,901

Notes: This table presents results from two different forecast bias tests (the second test is only practical in LAUSD where there are more frequent standardized tests). For the "individual" test, an estimate close to 1 shows that these measures, though subject to estimation error, are actually predictive of student outcomes out of sample. For the 'prediction' test, an estimate close to 0 shows that these measures are not biased by sorting on omitted variables. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14: Forecast Bias Tests of Value-Added in Maryland

Panel A: Individual Test, Math	Test Score	GPA	Frac Days Absent	Suspended	Heldback
Corresponding VA Measure	1.07	0.85	-1.13	-1.32	-0.68
	(0.09)	(0.06)	(0.09)	(0.09)	(0.06)
N	226,260	1,143,672	1,144,560	1,144,560	1,069,973
Panel B: Individual Test, ELA	Test Score	GPA	Frac Days Absent	Suspended	Heldback
Corresponding VA Measure	0.88	0.88	-1.19	-1.40	-0.43
	(0.11)	(0.06)	(0.10)	(0.10)	(0.06)
N	254,944	1,163,448	1,167,090	1,167,186	1,101,485

Notes: This table presents results from the forecast bias test that is relevant in Maryland. For the "individual" test, an estimate close to 1 shows that these measures, though subject to estimation error, are actually predictive of student outcomes out of sample. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15: Effects of Measures of Teacher Grade Inflation and Value-Added with Additional Controls on High School Outcomes in Maryland

	Future Test (math)	Future Test (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Mean GI	0.001	-0.007+	-0.001***	-0.001***	-0.004***	-0.407
	(0.004)	(0.003)	(0.000)	(0.000)	(0.001)	(1.254)
Passing GI	-0.003	0.002	-0.004***	0.002**	0.001	-2.793**
	(0.004)	(0.003)	(0.001)	(0.001)	(0.002)	(1.400)
Cog. VA	0.016***	0.031***	0.001+	0.000	0.004***	-0.705
	(0.004)	(0.003)	(0.000)	(0.000)	(0.001)	(1.268)
Noncog. VA	0.001	0.011***	-0.004***	0.007***	0.002	-6.083***
	(0.004)	(0.003)	(0.001)	(0.001)	(0.002)	(1.440)
Outcome Mean	-0.20	0.03	0.04	0.90	0.65	1158.08
Observations	123,205	279,802	1,117,227	979,425	1,117,945	252,950
R^2	0.356	0.591	0.153	0.238	0.241	0.772

Notes: This table presents estimates from regressions of the relevant outcome (column) on alternative measures of mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added. The alternative measures are created using additional demographic controls available only in Maryland: race, sex, and use of Free and Reduced Price Lunch. The regressions also include those additional controls along with our standard controls. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

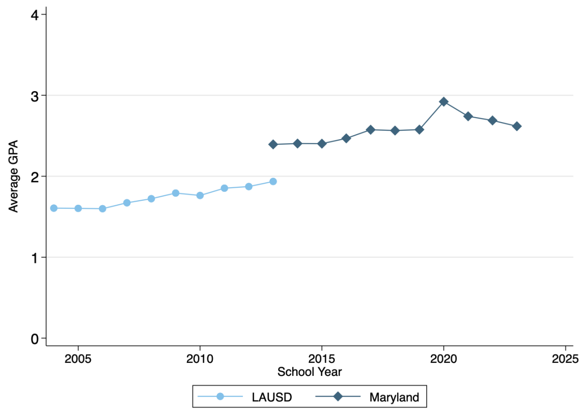
Table A.16: Effects of Teacher Grade Inflation and Value-Added with Additional Controls on College and Career Outcomes in Maryland

	Enrolled in Postsecondary		Graduated from Postsecondary		Employed		Unconditional Winz. Earnings	
	1 year	2 years	4 years	6 years	1 year	6 years	1 year	6 years
Mean GI	-0.006***	-0.005***	0.000	-0.002+	-0.004***	-0.002+	-48.557***	-168.062***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(12.284)	(51.641)
Passing GI	0.002**	0.001	-0.002+	-0.002**	0.000	0.000	11.538	-6.060
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(11.629)	(41.400)
Cog. VA	0.001	0.001	-0.004***	-0.001	-0.000	-0.001	-29.773**	-53.385
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(12.433)	(41.588)
Noncog. VA	0.006***	0.005***	-0.001	0.000	0.002***	0.001	33.330***	-24.619
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(9.711)	(43.546)
Outcome Mean	0.64	0.56	0.27	0.40	0.68	0.62	5444.70	17801.59
Observations	979,425	838,616	552,566	274,354	979,425	274,354	979,425	274,354
R^2	0.285	0.278	0.237	0.338	0.034	0.049	0.080	0.040

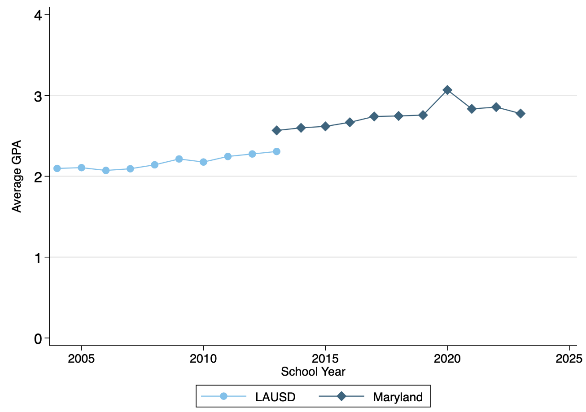
Notes: This table presents estimates from regressions of the relevant outcome (column) on alternative measures of mean grade inflation, passing grade inflation, cognitive value-added, and noncognitive value-added. The alternative measures are created using additional demographic controls available only in Maryland: race, sex, and use of Free and Reduced Price Lunch. The regressions also include those additional controls along with our standard controls. The outcomes in this table are only available in the Maryland data, not in LAUSD. For this student-year level analysis, value-added and grade inflation are measured as averages across the student's math and English instructors in each year. The controls include school, grade, and year fixed effects, ELL, and previous year values of math test score, English test score, total GPA, fraction of days absent, suspension, and held back. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.1: Average GPA over Time by Subject

(a) Math



(b) ELA



Notes: This figure plots average math and English GPA over time for students in our Los Angeles and Maryland samples. The data from the LAUSD span 2004 to 2013 and the data from Maryland span 2013 to 2023. GPA is measured on a 0 to 4 scale (e.g., A = 4.0, B = 3.0).