The Distributional and Long-Term Effects of Grade Inflation

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

October 2023

Abstract

Teachers have discretion over how they map student achievement into grades, and their leniency in doing so may positively or negatively impact student achievement. In this paper we construct two measures of grading leniency: ”mean grade inflation” which measures how much higher grades are than would be expected, and ”passing grade inflation” which measures leniency in receiving a passing grade. We show that these measures represent related but distinct grading practices of teachers. Grading leniency is not very correlated with other well-established teacher characteristics such as test score and non-cognitive value-added, which suggests that teachers may face tradeoffs in classroom practices. We show that more lenient teachers reduce performance on tests in subsequent years, and that leniency also has persistent effects, decreasing some students’ likelihood of taking the SAT and graduating. However, while mean grade inflation negatively affects outcomes, we find that passing grade inflation is positively associated with grade progression, especially for lower-performing students.

*University of Notre Dame, NBER, IZA, and CESifo, Email: jeffdenning@nd.edu
†University of Maryland, US Census Bureau
‡University of Maryland, Email: npope@umd.edu
§Stanford University

1 Introduction

Teachers are among the most important school-provided inputs in the education production function. Effective teachers benefit students on numerous dimensions in both the short- and long-run, including but not limited to test scores, suspensions, absences, effort, and adult earnings (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).

A small literature showing that teachers' grading practices, biases, and expectations affect student outcomes provides a notable exception to this critique (Carlana, 2019; Figlio and Lucas, 2004; Gershenson et al., 2022; Papageorge et al., 2020). These are malleable factors that pre- and in-service teacher training programs could target. Moreover, teachers, schools, and districts could actively change or adopt specific grading policies to benefit students. Indeed, school districts throughout the country are in the throes of contentious debates about whether and how to change grading standards (Alex, 2022; Randazzo, 2023; Graham, 2017; Las Vegas Review-Journal Editorial, 2023). The root issue of these debates is theoretical ambiguity as to whether high standards boost performance by increasing student effort or hamper performance by discouraging students.

In this paper we first ask whether grading standards affect students' future outcomes including test performance, high school graduation, and SAT test taking. We then ask which students (if any) are harmed and which students (if any) are helped by tougher standards. Existing research shows that students benefit from teachers who have high *average* grading standards (Figlio and Lucas, 2004; Gershenson et al., 2022; Mozenter, 2019). However, teachers' grading practices need not be unidimensional nor do students on different margins necessarily respond uniformly to those grading practices. We address these limitations of prior work by constructing a more nuanced and realistic formulation of teachers' grading standards along both the intensive (course grade) and extensive (pass / fail) margins and estimate the effects of both margins on a variety of short- and long-run educational outcomes. We do so using administrative data from the nation's second-largest school district, Los Angeles Unified.

These analyses are relevant, given current news coverage of rising grade inflation and debates regarding policies and regime changes that may curtail it. As such, a clear policy question is to understand the effects of this persistent increase in grades. We document that grades are increasing in the Los Angeles Unified School District (LAUSD): Figure 6 shows that GPAs increased over the 2004-2014 school cohorts on average and also in math and English separately. Over these 10 years, the mean GPA increased from 2.21 to 2.49, roughly a quarter of a letter grade. We also find that over those same 10 years, for students who score within one tenth of a standard deviations of the average on the state standardized test, the average GPA rose by .41. Students scoring over 3 standard deviations above the mean saw their average GPA increase by .17.[1] Later, we construct more robust measures of grade inflation that confirm a similar pattern. However, the goal of this paper is not to quantify how much rising grades are due to grade inflation. Instead, we use the rise in grades as motivation to understand what the consequences of leniency are for students. We rely on several studies that have shown that grading has become more lenient over time in high school and college, suggesting that these changes in grades represent grade inflation rather than increases in human capital (Zhang and Sanchez, 2013; Gershenson, 2018; Hurwitz and Lee, 2018; Denning et al., 2022; Sanchez and Moore, 2022).[2] Given the large increases in grade inflation over this time, it has become an urgent policy question to understand the effects of this persistent increase in grades over this time period.

Reactions to the presence of grade inflation are split. One camp is dismissive, suggesting that grade inflation does not negatively affect students and that high grading standards may even discourage students (Kohn, 2002). Others suggest that grade inflation is harmful to students, leaving them unprepared for future educational or vocational endeavors and decreasing their overall effort (Wright, 2019). Both of these camps might be right – the effects of grade inflation on student success are theoretically ambiguous. Minimum grades are required to pass classes and ultimately graduate from high school. Hence, grade inflation may help students to pass classes and graduate when they otherwise would not. However, working hard to achieve high grades provides 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: higher grades can reduce failure but

[1] 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.

[2] One example of this is the Equitable Grading movement which has been adopted in approximately 50 school districts since 2013. This approach emphasizes flexibility in deadlines and raising minimum scores, among other practices.

blunt incentives to study and learn material. They also likely differ in importance for different students. Students on the margin of failing a class are more likely to benefit from grade inflation. Students who are easily above the passing margin may be harmed by the reduced incentive to study.

This paper makes several contributions to our understanding of grading practices. First, we construct two measures of grading leniency for each teacher in our sample: one that measures "mean grade inflation," similar in concept to (Figlio and Lucas, 2004; Mozenter, 2019; Gershenson et al., 2022). Mean grade inflation measures how much higher *on average* grades are for a teacher than would be expected, given a student's standardized test score and other characteristics. Additionally, we introduce a novel measure of grading leniency that measures grade inflation in receiving a *passing grade*, which we refer to as "passing grade inflation." These measures allow us to understand the trade offs of greater grading leniency along these two dimensions. We show that these two measures of grading leniency are highly correlated (correlation coefficient of .9) but distinct. Hence we conclude that mean grade inflation and passing grade inflation represent related but sometimes different grading practices of teachers. We show that this distinction is important, with passing grade inflation and mean grade inflation typically having opposite effects on students' outcomes. Additionally, they differ in how they affect students from different parts of the achievement distribution.

Second, we ask how grading leniency 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 test score 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.[3]

Third, we evaluate the effect of both of these measures on longer-term outcomes such as high

[3] 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 not 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.

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). Consistent with this, we show that having higher mean grade inflating teachers reduces performance on future tests.[4] 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 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 contribute to the existing literature by documenting important heterogeneity in the effects of different types of grading leniency. 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 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, has previously been unexplored.

Fourth, we show the measures of grade inflation have heterogeneous effects depending on the characteristics of the students. Having a higher mean grade inflating teacher has similar negative effects on students' future test scores and likelihood of graduation for all students, but is more detrimental on taking the SAT for students in the top of the 8th grade GPA distribution. On the other hand, 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 among students in the bottom of the 8th grade GPA distribution.

[4] Mozenter (2019) finds no effect on longer-term outcomes, but he 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.

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 use administrative student-level panel data from the Los Angeles Unified School District (LAUSD). LAUSD is the second largest school district in the United States. The data contain student-year observations from 2004 to 2013 for high school students. In 2003, the school district was 71.9 percent Hispanic, 12.1 percent Black, and 9.4 percent white. In addition, in 2010 over 80 percent of students received free and reduced price lunch. 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 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.[5] Several characteristics are notable in the LAUSD. First, measures of student success are low. Only 49 percent of ninth graders graduate from the LAUSD within 4 years and 58 percent graduate within 5 years. 12 percent of students are held back at some point, which we define as your administrative grade level being the same as your grade level in the previous year. The average student misses 7 percent of their classes and 20 percent of students are English language learners. In the LAUSD, our best measure of college intention is taking the SAT; 45 percent of students in our data take the SAT. Among those who take the SAT, the average score is a 1335.38 on a 2400 scale, which is approximately the 30th percentile of test takers. Notably, course grades are low, with an average GPA of 2.29. Course specific GPAs are even lower for English and math, with averages of 2.18 and 1.75, 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 non-cognitive teacher value-added measures. Students perform similarly in these non-cognitive dimensions, with an average effort GPA of 2.16 and an average cooperation GPA of 2.39.

Figure 1 plots the distribution of GPAs for different measures including math, English, and the noncognitive effort and cooperation measures. Math and English GPAs range from 0 to 4, whereas noncognitive GPAs range from 1 to 3. Figure 1 shows that, to varying degrees, higher grades are more common than lower for English, total, effort, and cooperative GPAs. The opposite is true for math, with grades being fairly evenly distributed across the scale or even slightly skewed toward low grades.

[5] We require that students have the following information to be included in this sample: school and district code, grade level, math or English grade inflation measures, math or English teacher value-added measures, non-cognitive 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.

3 Estimation

3.1 Constructing our measures of grading leniency

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, the number of students who fail, etc. In assessing the effects of grade inflation, 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 similar to Figlio and Lucas (2004); Gershenson et al. (2022); Mozenter (2019).[6] We model the student's grade as in Equation 1:

$$Grade_{ijst} = GI_{jt}^{mean} + \beta_1 TestScore_{ijst} + \beta_2 Grade_{ist-1} + X_{it}\beta + \varepsilon_{ijst} \tag{1}$$

where i indexes student, j indexes teacher, s indexes school and t indexes year. The object of interest is GI_{jt}^{mean} which is the year-specific teacher fixed effect. This is the teacher's contribution to

[6] 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").

grades after controlling for several important factors. First we account for a student's performance in the subject as measured by their test score. Second, we control for student characteristics. The 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, suspended, and heldback. Importantly, we also control for $Grade_{ijst-1}$ which is the student's grade in the focal subject from the prior year.

For each teacher we calculate $\widehat{GI_{jt}^{mean}}$ which is our measure of grading leniency. We will call this "mean grading leniency" 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 $\widehat{GI_{jt}^{mean}}$ and $\widehat{GI_{jt}^{pass}}$ 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.

$\widehat{GI_{jt}^{mean}}$ is fundamentally defined as a residual. It is worth considering what $\widehat{GI_{jt}^{mean}}$ could be capturing aside from grading leniency. We rule out some potential alternative explanations by controlling for student characteristics. For example, student performance is accounted for in two ways. First, Equation 1 includes the student's contemporaneous 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.[7]

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 contem-

[7] 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.

poraneous 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 such as grades and test scores 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 not passing the class. We expect that these two measures of leniency may have different effects on students' future performance.[8] For our second measure we are still interested in estimating the teacher effect in the modified 1, which we refer to as GI_{jt}^{pass}.

Ultimately, we want our measures to characterize how much grade inflation a student experiences in a given year.[9] 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.[10] We then characterize the grade inflation that a student experiences in a year by averaging over all the student's teachers in a given subject, weighted by the number of classes a student takes with that same teacher.[11] 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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

grade inflation that a student experiences in a given year. Figure 2 shows the distribution of these measures of grade inflation.

3.2 Other teacher measures

In order to understand how related our measure of grade inflation is to a general indicator of teacher quality, we construct both test score and noncognitive value-added measures. We estimate test score value-added using a jack-knife empirical Bayes estimator following Chetty et al. (2014a). We estimate noncognitive valued-added also 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-year level measures of grade inflation. We first construct test score and non-cognitive 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 we combine math and English course measures of days absent, suspension, grade retention, cooperation GPA, total academic GPA, and effort GPA into a single "non cognitive" value-added measure by summing the components and standardizing to be mean zero, standard deviation one. We also create subject-specific "non-cognitive" value-added measures by summing the components from only that subject's teachers and standardizing to be mean zero, standard deviation one.

3.3 The effects of grade inflation on longer-term outcomes

To explore the effects of grade inflation on longer-term outcomes, we estimate specifications similar to Chetty et al. (2014b), Petek and Pope (2023), Gershenson et al. (2022), 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:

$$Y_{it} = \alpha_{mean}\widehat{GI_{it}^{mean}} + \theta_{pass}\widehat{GI_{it}^{pass}} + \delta_{cogVA}\widehat{VA_{it}^{test}} + \psi_{noncogVA}\widehat{VA_{it}^{noncog}} + X_{it}\beta + \eta_{it} \quad (2)$$

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.[12]

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 3.2. 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

[12]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.

test score is the outcome variable.

4 Results

4.1 How Correlated is Grade Inflation with Other Teacher Characteristics?

In Table 2, we correlate our two measures of grade inflation with several measures of teacher value-added such as teacher value-added on test scores, suspensions, absences, and effort as in Petek and Pope (2023). These correlations are inherently not causal, but rather reflect the relationship between grade inflation and other teacher attributes as observed in the data. It is possible that a teacher could start inflating grades and not change their other VA measures. However, it is also possible that these correlations (partly) reflect a causal relationship and so teachers cannot alter their grade inflation practices without affecting other types of value-added.

Since value-added is measured with error, within-teacher correlations could be biased due to random estimation errors or due to cross-outcome within-year correlations between errors. We follow Jackson et al. (2023) to calculate correlations using a split-sample approach. For example, to calculate the correlation between grade inflation and value-added, we first calculate the raw correlation between grade inflation in odd years and value added in even years. We then divide by the square root of the product of the two within-outcome cross-year correlations: the correlation between grade inflation in even years and in odd years, and the correlation between value-added in even years and in odd years. For each pair of outcomes, this produces two correlation estimates. In general, these estimates need not match. In some cases, the magnitudes are different, though the sign always matches.

Mean grade inflation and passing grade inflation are correlated at .88 which is also visualized in Figure 5. We would expect these two measures of grade inflation to be correlated since a teacher who raises the grades of all students would also increase the probability of passing. While these two measures are highly correlated, they are not perfectly correlated – passing grade inflation and mean grade inflation appear to be distinct. Because they are not perfectly correlated, we will be able to use both measures in Equation 2 to tease out the effects of different types of grade inflation.

We also show the correlation between our grade inflation measures and other teacher characteristics such as value-added in Table 2 and Figure 4. Teachers who inflate grades tend to have lower value-added. Grade inflation is negatively correlated with our index of test score value-added at -.40 and -.32 for mean and passing grade inflation respectively. The negative correlation is not surprising given results in Betts and Grogger (2003); Figlio and Lucas (2004); Mozenter (2019); Gershenson et al. (2022) which find that grade inflating teachers reduce test scores. We confirm these findings in our paper. We also explore the relationship between grade inflation and non test score value-added measures. Teacher grade inflation is positively correlated with non cognitive value-added at .31 for mean grade inflation and .30 for passing grade inflation. Our takeaway from these correlations is that grade inflation represents a distinct teacher characteristic from those previously studied. Grade inflation is negatively correlated with test score value-added and positively correlated with non cognitive value added.

A teacher's decision to inflate grades could reflect a tradeoff if the correlations are (partly) causal because grade inflation is both positively and negatively correlated with other desirable teacher characteristics. For instance a grade inflating teacher may induce students to attend class more but reduce future test scores. As a result, the next section of the paper will explore whether grade inflation has positive or negative effects on students in the longer run while holding other teacher characteristics constant.

4.2 Does Grade Inflation Matter for Future Outcomes?

We next explore if grade inflation matters for future outcomes by estimating Equation 2, results of which are shown in Table 3. Throughout we control for mean grade inflation, passing grade inflation, test-score value-added, and noncognitive value-added. Hence, we are interpreting the effects of each of these teacher attributes while holding fixed the other teacher characteristics.

In Table 3 we note several things. First, teacher test-score value-added and noncognitive value-added have their expected signs, improving future test scores, student graduation, and the probability of taking the SAT. Additionally, measures of value-added reduce retention and student absence. These results closely mirror those of Petek and Pope (2023).

We next focus on the effects of mean grade inflation. We find that mean grade inflation reduces future test scores, with a one standard deviation increase in mean grade inflation reducing future math test scores by .034 standard deviations and future English test scores by .024 standard deviations. We find no contemporaneous effect on math test scores but a negative effect on English.

As a comparison, teacher value-added increases future test scores by .108 for math and .049 for English. We interpret our estimates as suggesting that teacher's grading leniency has a meaningful impact on future test score performance – about 30% as large as the effect of traditional value-added measures in math and 48% as large in English. This finding closely mirrors that of (Figlio and Lucas, 2004; Gershenson et al., 2022).

Mean grade inflation negligibly reduces the chance that that a student takes the SAT. This allows us to interpret the effects of grade inflation on a different measure of student achievement, SAT, without worrying about selection into having a reported test score. Our estimates suggest a very small effect overall of the mean grade inflation on our best measure of college intentions. We find that mean grade inflation reduces future SAT scores by 4.6 points (on a 2100 point scale). Mean grade inflation has a similar small negative effect on PSAT scores (see Table A.4). Focusing on the SAT and PSAT is interesting because they are designed for a different purpose and test different skills than end of year standardized tests. That grade inflation has similar negative effects on both types of tests suggests a more general reduction in human capital resulting from grade inflation. In addition, we find that mean grade inflation reduces the probability of graduation within 5 years by .8 percentage points.

We now turn our attention to a teacher's passing grade inflation. In contrast to mean grade inflation, teacher passing grade inflation has a small statistically significant positive effect on future standardized test scores in English and a similar size for math test scores that is significant at the 10 percent level. Passing grade inflation also has a positive effect on PSAT scores.[13] Passing grade inflation reduces the probability of being held back in the current year by 1.1 percentage points, as we would expect given that being held back is partially a function of failing a class. This is a meaningful effect size because being held back is not very common, with 13 percent of students being held back. We estimate a small statistically significant increase in the probability of graduating within 5 years of .006 percentage points. Students are also slightly more likely to take the SAT when they have on average high passing grade inflating teachers.

We explore several alternative specifications and find similar patterns. In Table A.7 we explore the effect of separate math and English grade inflation metrics. We find similar results to our

[13] There is no effect of passing grade inflation on taking the PSAT which makes interpretation of the PSAT score effects easier.

preferred specification generally with larger point estimates for math.

To summarize, we find that passing grade inflation has small positive impacts on student outcomes such as grade progression, graduation, and PSAT score. However, mean grade inflation reduces student standardized test scores and SAT scores. We conclude that not all grade inflation has the same effect – it depends critically on the type of grade inflation a teacher engages in.

4.3 Who is Affected by Grade Inflation?

We next explore which students are affected by grade inflation. In the introduction, we outlined one hypothesis as to how grade inflation might have important heterogeneous effects: students on the margin of failing may be benefited by having a teacher who engages in passing grade inflation. High achieving students may be harmed by mean grade inflation as their incentives to study are reduced. We test this by exploring heterogeneity by student achievement as measured by 8th grade GPA. We split our sample by above and below median 8th grade GPA in order to characterize student achievement using a measure that occur before students enter our sample. We estimate Equation 2 separately among above and below median 8th grade GPA samples in table 5.

We find that students who have below median grades in eighth grade see larger increases in the probability of graduating within 5 years from passing grade inflation at .7 percentage points. We estimate no effect on graduation for above median students. Mean grade inflation has negative effects on statewide test scores across the ability distribution. However, mean grade inflation reduces the probability that higher performing students take the SAT, and those students have lower SAT scores conditional on taking the test.

Our results exploring heterogeneity by student ability suggest that students from the bottom half of the grade distribution benefit more from teachers who engage in passing grade inflation. In contrast, mean grade inflation appears to be harmful for graduation and test score performance for most of the student ability distribution, with some evidence of particular harm among the most high-achieving students.

We also explore heterogeneity by the grade in school that a student experiences grade inflation. Grade inflation experienced in early grades may have a different effect than grade inflation experienced later in high school. In Table 7 we find that negative effects of mean grade inflation on test scores and graduation occur in grades 9-11. The negative effects of mean grade inflation on

SAT scores are strongest in 9th grade.

Passing grade inflation appears to have larger effects on graduation in grades 9 and 10 than in later grades. This may be due to the structuring of when required courses are taken. We have explored heterogeneity by whether a course is required, but essentially all courses that have a standardized end of year test score are required for high school graduation, which makes such an exercise difficult. Mean grade inflation in earlier grades also has a large effect on SAT scores relative to mean grade inflation in later years.

Average grades differ at the school level, so a relevant question is whether the effects of grade inflation are the same in different school environments. To study this we split the sample into above and below median of the school-level average GPA in Table 6. In general, we see similar patterns to our results thus far. Mean grade inflation reduces test scores and graduation; passing grade inflation reduces the chance that a student is held back and increases graduation in five years.

5 Conclusion

We show that teacher's grading practices affect students in important and heterogeneous ways. We find that the type of grading leniency matters meaningfully for student outcomes. A teacher who generally inflates grades has a negative effect on future performance as measured by test scores. They also reduce high school graduation. In contrast, a teacher who inflate grades so that students are less likely to fail improves student graduation rates without reducing future test performance. We further show that the effects of this grade inflation differ depending on student characteristics. In particular, low performing students especially benefit from passing grade inflation and high performing students are particularly harmed by mean grade inflation.

Future work on grading practices should distinguish between different grading practices. Grading policy should also consider the distributional consequences of grade inflation. Our results suggest that reducing the likelihood of a failing grade, while maintaining higher standards for top grades may help the most students and harm the fewest.

References

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

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.

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.

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., and Tyner, A. (2022). Making the grade: The effect of teacher grading standards on student outcomes. *under review*.

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].

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.

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. (2023). Who benefits from attending effective high schools? *Journal of Labor Economics*, 0(ja):null.

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

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.

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.

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

Table 1: Student-level Summary Statistics (Index Sample)

	Mean	Std. Dev.	Min	Max	N
Math CST Score	0.06	1.02	-3.02	7.22	706,922
English CST Score	0.10	1.00	-3.37	5.70	752,431
Math GPA	1.75	1.24	0.00	4.00	910,194
English GPA	2.18	1.20	0.00	4.00	985,020
GPA	2.29	0.98	0.00	4.00	985,020
Effort GPA	2.16	0.53	1.00	3.00	985,020
Learning Skills GPA	2.39	0.45	1.00	3.00	985,020
Fraction of Days Absent	0.07	0.09	0.00	1.00	985,020
Ever Suspended	0.06	0.24	0.00	1.00	985,020
Held Back	0.12	0.32	0.00	1.00	985,020
English Learner	0.20	0.40	0.00	1.00	985,020
Average Teacher Experience	6.48	2.83	2.00	11.00	985,020
Don't Graduate in LAUSD	0.31	0.46	0.00	1.00	733,949
Leave Dataset	0.27	0.45	0.00	1.00	733,949
Graduate on Time	0.49	0.50	0.00	1.00	833,266
Graduate within 5 Years	0.58	0.49	0.00	1.00	733,949
Number of AP Courses	1.11	2.04	0.00	21.00	985,020
Ever Took SAT	0.36	0.48	0.00	1.00	833,266
Ever Took PSAT	0.43	0.49	0.00	1.00	646,242
SAT Score	1335.38	299.80	600.00	2400.00	363,022
PSAT Score	1114.21	258.12	600.00	2370.00	393,366
English CAHSEE Score	0.13	0.99	-3.32	2.46	854,094
Math CAHSEE Score	0.12	1.01	-2.94	2.72	856,998
10th Grade Science CST Score	0.12	1.01	-3.24	5.53	656,977
11th Grade Social CST Score	0.11	1.01	-3.19	5.27	674,673

Table 2: Correlations between GI and VA

	Base GI	Pass GI	Cog VA	Noncog VA
Base GI	1.0000	0.8829	-0.4270	0.2188
Pass GI	0.8737	1.0000	-0.3673	0.1746
Cog VA	-0.4039	-0.3209	1.0000	-0.0729
Noncog VA	0.3072	0.2984	-0.0418	1.0000

Table 3: Student-Year, Index of Math and English GI

	Test Score (math)	Future Test Score (math)	Test Score (ela)	Future Test Score (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
GI Factor	-0.008	-0.034***	-0.021***	-0.024***	0.001	-0.008***	-0.008***	-4.621***
	(0.007)	(0.007)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(1.603)
Passed GI Factor	-0.006	0.011+	0.010***	0.007**	-0.011***	0.006**	0.007***	0.308
	(0.006)	(0.006)	(0.002)	(0.003)	(0.001)	(0.003)	(0.002)	(1.606)
VA Cog. Factor	0.137***	0.108***	0.049***	0.053***	0.000	0.000	0.029***	23.690***
	(0.007)	(0.010)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(3.027)
VA Non-Cog. Factor	-0.014***	-0.012**	0.007***	0.010***	-0.007***	0.013***	0.009***	-0.889
	(0.003)	(0.005)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(1.412)
Outcome Mean	0.06	-0.02	0.10	0.04	0.13	0.58	0.32	1327.79
Observations	706,647	391,782	751,986	432,534	832,002	733,946	680,305	186,350
R^2	0.544	0.521	0.703	0.659	0.167	0.296	0.363	0.737

Table 4: Student-Year, Separate Math and English GI

	Test Score (math)	Future Test Score (math)	Test Score (ela)	Future Test Score (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Grade Inflation (math)	-0.018**	-0.058***	-0.024***	-0.026***	-0.000	-0.011***	-0.012***	-4.583**
	(0.007)	(0.010)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(2.053)
Grade Inflation (ela)	-0.002	-0.011	-0.013***	-0.013***	0.003+	-0.006+	-0.001	-1.414
	(0.005)	(0.008)	(0.003)	(0.004)	(0.002)	(0.003)	(0.004)	(2.169)
Passed Grade Inflation (math)	0.003	0.024***	0.012***	0.012***	-0.010***	0.007**	0.008**	1.582
	(0.006)	(0.009)	(0.003)	(0.005)	(0.002)	(0.003)	(0.003)	(1.536)
Passed Grade Inflation (ela)	0.003	0.006	0.005+	-0.002	-0.011***	0.006	0.003	-3.020
	(0.006)	(0.008)	(0.003)	(0.005)	(0.002)	(0.004)	(0.004)	(2.233)
VA Test (math)	0.188***	0.111***	0.011***	0.014***	0.002	-0.008***	0.010***	17.884***
	(0.005)	(0.012)	(0.002)	(0.003)	(0.001)	(0.002)	(0.003)	(2.666)
VA Test (ela)	0.035***	0.062***	0.070***	0.074***	-0.002	0.015***	0.037***	21.436***
	(0.006)	(0.009)	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)	(3.586)
VA Non-Cog. Factor (math)	-0.007***	-0.010**	0.003***	0.003**	-0.004***	0.008***	0.006***	-2.504**
	(0.003)	(0.004)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(1.044)
VA Non-Cog. Factor (ela)	0.001	-0.002	0.001	0.006***	-0.004***	0.005***	0.002+	-1.669
	(0.003)	(0.005)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(1.514)
Outcome Mean	0.18	0.11	0.30	0.23	0.11	0.64	0.40	1367.84
Observations	368,480	201,188	380,468	217,860	415,677	330,016	326,271	174,478
R^2	0.590	0.563	0.719	0.679	0.161	0.307	0.378	0.745

Table 5: Student-Year, Split by Quartiles of Student 8th Grade GPA

	Test Score (math)	Test Score (ela)	Future Test Score (math)	Future Test Score (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Below Median								
GI Factor	-0.004	-0.020***	-0.031***	-0.025***	-0.004	-0.007***	-0.002	-2.360+
	(0.005)	(0.003)	(0.006)	(0.005)	(0.003)	(0.003)	(0.002)	(1.335)
Passed GI Factor	-0.004	0.008***	0.006	0.006	-0.012***	0.007***	0.003+	-2.570+
	(0.004)	(0.003)	(0.005)	(0.004)	(0.002)	(0.002)	(0.002)	(1.459)
Outcome Mean	-0.25	-0.26	-0.33	-0.31	0.23	0.47	0.14	1213.85
Observations	283,306	302,081	161,249	181,382	343,437	263,059	272,601	55,051
R^2	0.367	0.583	0.323	0.522	0.164	0.289	0.223	0.651
Above Median								
GI Factor	-0.008	-0.020***	-0.037***	-0.022***	0.002+	-0.005**	-0.008***	-4.165**
	(0.009)	(0.003)	(0.007)	(0.003)	(0.001)	(0.002)	(0.003)	(1.678)
Passed GI Factor	-0.011	0.011***	0.013+	0.009***	-0.009***	0.001	0.007**	0.640
	(0.007)	(0.003)	(0.007)	(0.003)	(0.001)	(0.002)	(0.003)	(1.653)
Outcome Mean	0.37	0.47	0.27	0.40	0.05	0.81	0.51	1371.51
Observations	311,350	324,274	189,020	202,392	337,968	283,415	268,546	178,682
R^2	0.587	0.724	0.568	0.683	0.083	0.182	0.287	0.751

Table 6: Student-Year, Split by Quartiles of School Average GPA

	Test Score (math)	Test Score (ela)	Future Test Score (math)	Future Test Score (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
Below Median								
GI Factor	-0.008	-0.024***	-0.037***	-0.030***	0.001	-0.009**	-0.013***	-4.779***
	(0.007)	(0.004)	(0.010)	(0.004)	(0.002)	(0.003)	(0.004)	(1.477)
Passed GI Factor	-0.005	0.013***	0.013	0.012***	-0.011***	0.007**	0.010***	1.350
	(0.006)	(0.003)	(0.008)	(0.004)	(0.001)	(0.003)	(0.003)	(1.277)
Outcome Mean	-0.16	-0.12	-0.22	-0.17	0.16	0.52	0.25	1242.77
Observations	339,950	362,136	186,543	205,858	410,334	354,675	336,215	113,256
R^2	0.464	0.659	0.435	0.612	0.170	0.292	0.322	0.702
Above Median								
GI Factor	-0.006	-0.020***	-0.029***	-0.020***	0.001	-0.007	-0.004	-3.190
	(0.010)	(0.003)	(0.008)	(0.004)	(0.002)	(0.005)	(0.003)	(2.120)
Passed GI Factor	-0.010	0.009***	0.006	0.006	-0.011***	0.006	0.004	-0.440
	(0.008)	(0.002)	(0.009)	(0.004)	(0.003)	(0.004)	(0.003)	(2.169)
Outcome Mean	0.25	0.30	0.18	0.23	0.10	0.64	0.38	1392.60
Observations	366,697	389,850	205,239	226,676	421,668	379,270	344,089	165,865
R^2	0.567	0.713	0.544	0.669	0.155	0.280	0.377	0.738

Table 7: Student-Year, Split by Grade

	Test Score (math)	Test Score (ela)	Future Test Score (math)	Future Test Score (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
9th								
GI Factor	-0.002	-0.020***	-0.035***	-0.026***	-0.004	-0.008***	-0.008***	-7.292***
	(0.008)	(0.003)	(0.007)	(0.004)	(0.003)	(0.003)	(0.003)	(1.760)
Passed GI Factor	-0.007	0.010***	0.011	0.008**	-0.012***	0.009***	0.008***	3.134^{+}
	(0.007)	(0.003)	(0.007)	(0.004)	(0.002)	(0.002)	(0.002)	(1.759)
Outcome Mean	0.05	0.07	-0.04	0.01	0.20	0.49	0.24	1322.40
Observations	313,354	317,617	234,734	250,456	344,601	225,071	258,079	87,499
R^2	0.533	0.710	0.534	0.663	0.220	0.326	0.352	0.740
10th								
GI Factor	-0.019**	-0.027***	-0.035***	-0.024***	0.003	-0.014***	-0.015***	-3.681**
	(0.008)	(0.003)	(0.007)	(0.004)	(0.002)	(0.002)	(0.003)	(1.805)
Passed GI Factor	-0.011	0.011***	0.004	0.006	-0.009***	0.012***	0.011***	-1.935
	(0.007)	(0.003)	(0.006)	(0.004)	(0.002)	(0.003)	(0.002)	(1.572)
Outcome Mean	0.05	0.09	0.03	0.09	0.11	0.60	0.32	1332.58
Observations	236,607	255,234	157,035	182,066	276,673	201,635	229,606	98,842
R^2	0.579	0.715	0.575	0.669	0.119	0.330	0.380	0.759
11th								
GI Factor	-0.014^{+}	-0.017***			-0.000	-0.008***	-0.001	-2.209
	(0.008)	(0.004)			(0.001)	(0.003)	(0.006)	(1.833)
Passed GI Factor	-0.007	0.003			-0.003^{+}	0.006**	0.002	-1.064
	(0.007)	(0.004)			(0.002)	(0.003)	(0.005)	(1.912)
Outcome Mean	0.07	0.15			0.06	0.62	0.41	1339.90
Observations	156,674	179,127			210,725	171,137	192,614	92,763
R^2	0.621	0.703			0.091	0.400	0.394	0.766
12th								
GI Factor						-0.000	-0.013**	-3.017
						(0.003)	(0.005)	(2.193)
Passed GI Factor						-0.000	0.010**	0.312
						(0.003)	(0.005)	(1.869)
Outcome Mean						0.67	0.55	1347.29
Observations						136,076	152,948	83,887
R^2						0.552	0.329	0.771

Figure 1: Distribution of High School GPAs

Figure 2: Distribution of High School Grade Inflation

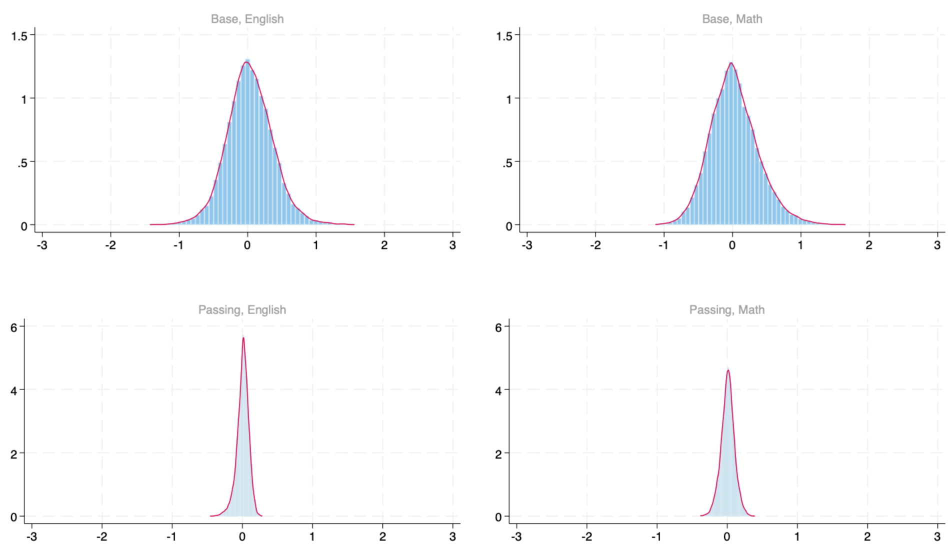


Figure 3: Distribution of High School Value-Added

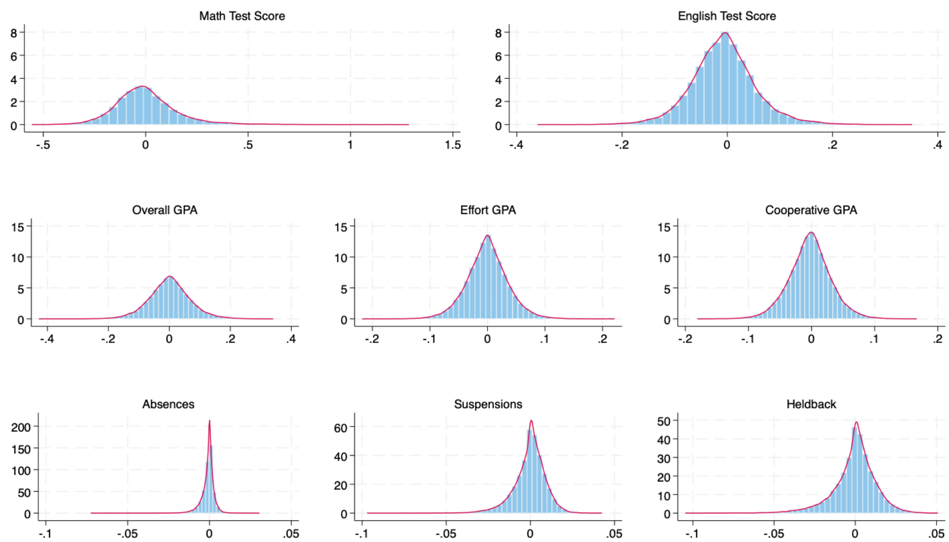


Figure 4: Correlation of GI and VA

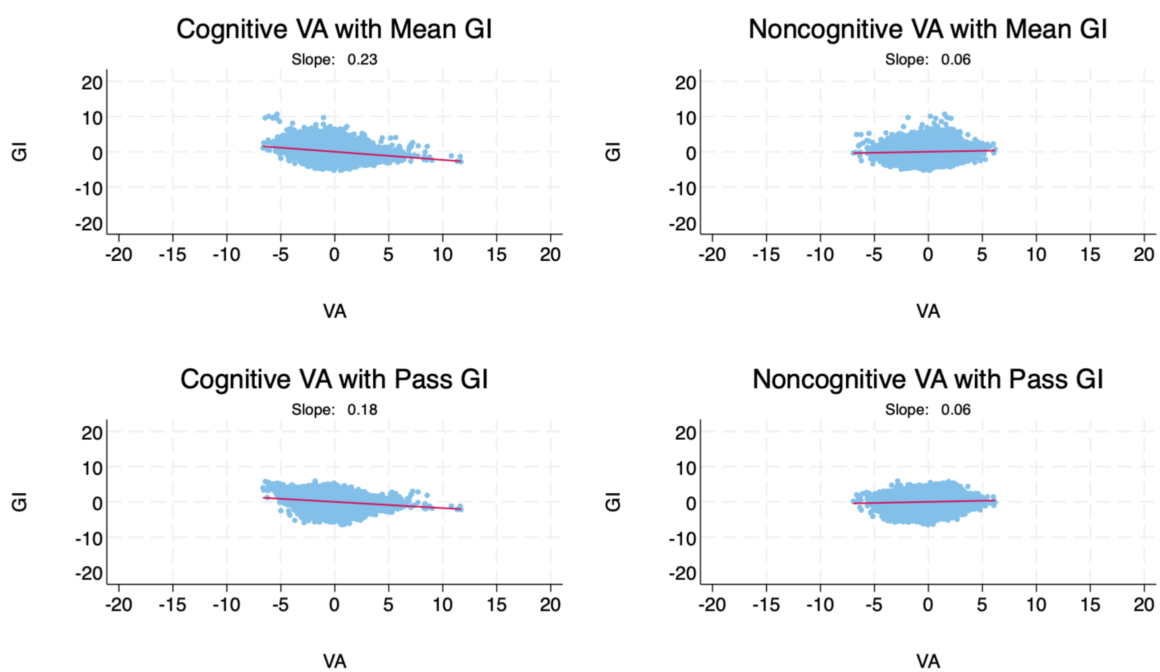


Figure 5: Correlation of GI Measures

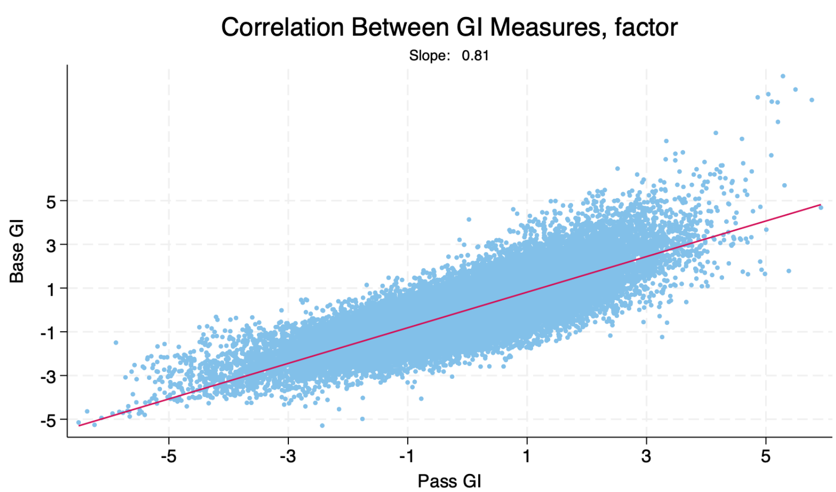


Figure 6: School-level Average Grades Over Time

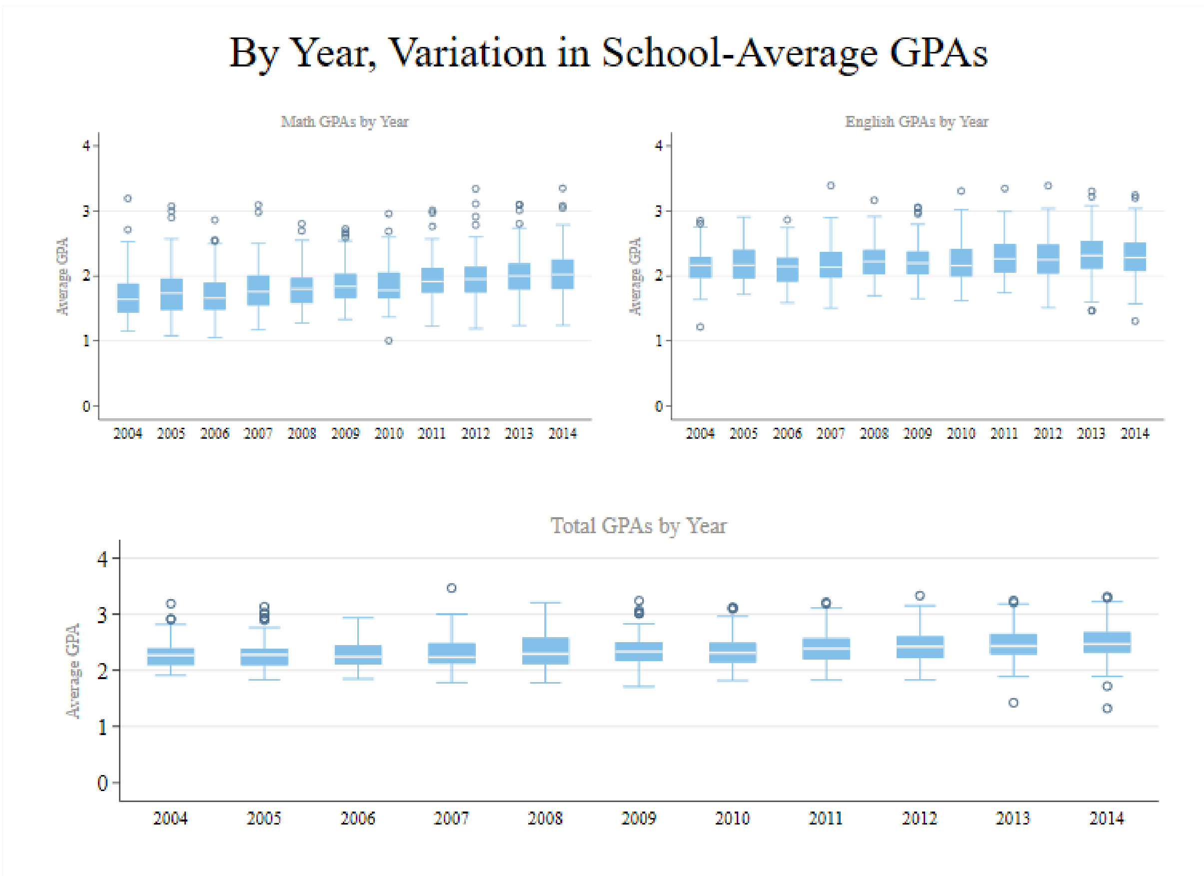
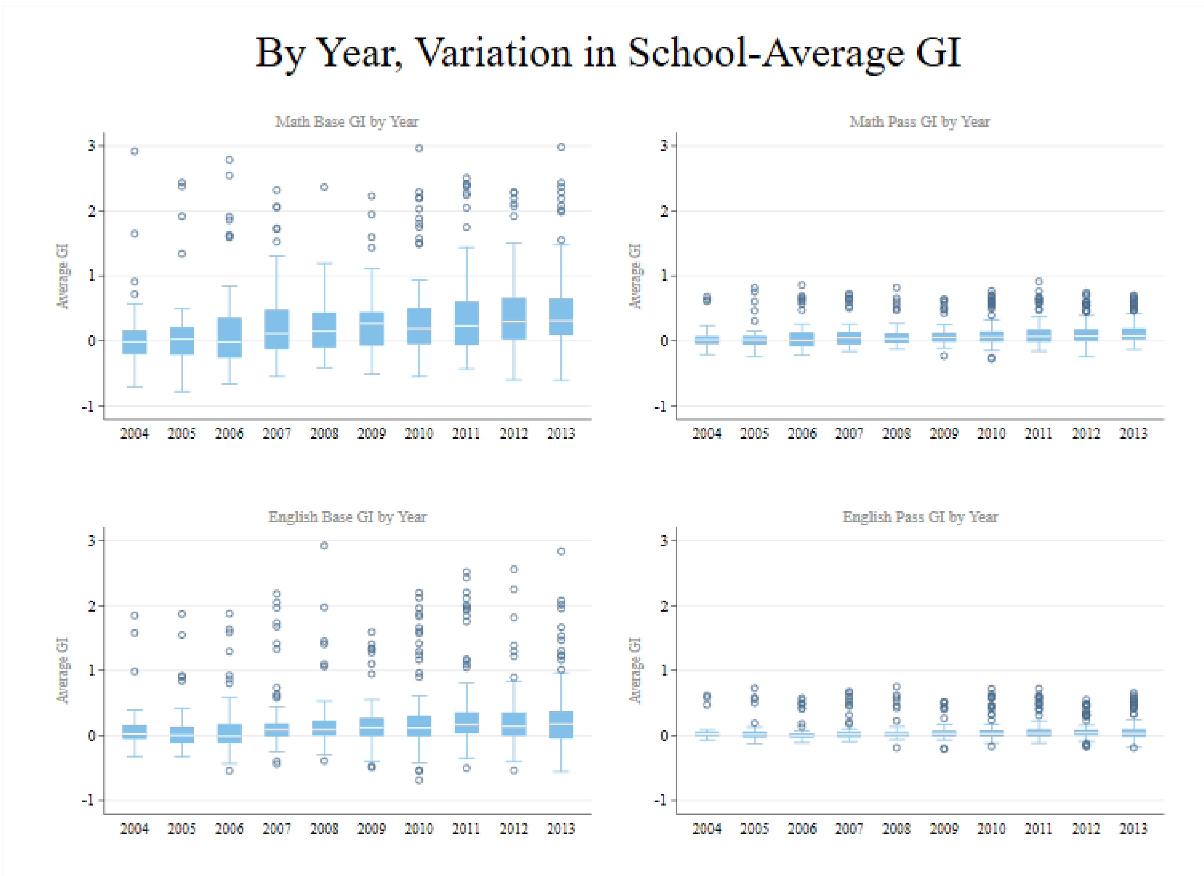


Figure 7: School-level Average Grade Inflation Over Time



Appendix Figures and Tables

Table A.1: Student-level Summary Statistics (Separate Sample)

	Mean	Std. Dev.	Min	Max	N
Math CST Score	0.18	1.08	-3.02	7.22	368,561
English CST Score	0.30	1.00	-3.37	5.70	380,576
Math GPA	1.93	1.27	0.00	4.00	473,509
English GPA	2.33	1.22	0.00	4.00	473,509
GPA	2.43	0.98	0.00	4.00	473,509
Effort GPA	2.24	0.53	1.00	3.00	473,509
Learning Skills GPA	2.45	0.44	1.00	3.00	473,509
Fraction of Days Absent	0.06	0.09	0.00	1.00	473,509
Ever Suspended	0.05	0.22	0.00	1.00	473,509
Held Back	0.10	0.29	0.00	1.00	473,509
English Learner	0.15	0.35	0.00	1.00	473,509
Average Teacher Experience	6.80	2.81	2.00	11.00	473,509
Don't Graduate in LAUSD	0.27	0.44	0.00	1.00	330,022
Leave Dataset	0.24	0.43	0.00	1.00	330,022
Graduate on Time	0.54	0.50	0.00	1.00	384,085
Graduate within 5 Years	0.64	0.48	0.00	1.00	330,022
Number of AP Courses	1.53	2.38	0.00	21.00	473,509
Ever Took SAT	0.45	0.50	0.00	1.00	384,085
Ever Took PSAT	0.45	0.50	0.00	1.00	331,618
SAT Score	1371.72	300.85	600.00	2400.00	213,832
PSAT Score	1160.54	263.47	600.00	2350.00	221,369
English CAHSEE Score	0.32	0.97	-3.07	2.46	418,499
Math CAHSEE Score	0.32	1.02	-2.94	2.72	419,710
10th Grade Science CST Score	0.30	1.03	-3.24	5.53	337,241
11th Grade Social CST Score	0.28	1.02	-3.07	5.27	331,570

Table A.2: Correlations between GI and VA, Math

	Base GI	Pass GI	Test VA	GPA VA	Eff GPA VA	Coop GPA VA	Absences VA	Suspended VA	Heldback VA
Base GI	1.0000	0.9114	-0.3638	0.3299	0.2022	0.0135	0.1631	-0.1929	0.1681
Pass GI	0.8879	1.0000	-0.2799	0.3332	0.1759	0.0293	0.0885	-0.2475	0.1342
Test VA	-0.3189	-0.2511	1.0000	-0.0316	0.0890	0.1261	-0.0699	0.0634	-0.1748
GPA VA	0.3689	0.3471	-0.0689	1.0000	0.9914	0.7167	-0.0541	-0.0278	-0.0629
Eff GPA VA	0.1678	0.1706	0.0502	0.7555	1.0000	0.8078	-0.1249	-0.0873	-0.1807
Coop GPA VA	-0.0899	-0.0753	0.1877	0.5186	0.8492	1.0000	-0.3356	-0.1774	-0.3765
Absences VA	0.2746	0.3685	0.0008	0.4665	0.5210	0.2926	1.0000	0.8684	0.0023
Suspended VA	-0.0848	-0.1633	0.0068	-0.3163	-0.3700	-0.2640	0.4354	1.0000	-0.0750
Heldback VA	0.2610	0.2380	-0.1568	0.0218	-0.1433	-0.3265	0.3625	0.6545	1.0000

Table A.3: Correlations between GI and VA, ELA

	Base GI	Pass GI	Test VA	GPA VA	Eff GPA VA	Coop GPA VA	Absences VA	Suspended VA	Heldback VA
Base GI	1.0000	0.8400	-0.3908	0.3111	0.1565	-0.0436	0.0560	-0.0829	0.1655
Pass GI	0.8528	1.0000	-0.3620	0.3452	0.1615	-0.0531	-0.0377	-0.0478	0.1940
Test VA	-0.3719	-0.2778	1.0000	-0.2450	-0.0147	0.3168	-0.1899	-0.1561	-0.1872
GPA VA	0.3194	0.3360	-0.2469	1.0000	0.8189	0.5948	0.0528	0.0954	-0.1403
Eff GPA VA	0.2065	0.1954	-0.0578	0.8485	1.0000	0.8605	-0.0110	0.0097	-0.2741
Coop GPA VA	-0.0274	-0.0295	0.2579	0.5707	0.8083	1.0000	-0.2204	0.0217	-0.4292
Absences VA	0.1898	0.1271	-0.2126	0.1371	0.0881	-0.0057	1.0000	0.5481	0.0082
Suspended VA	0.0649	-0.0033	-0.1177	-0.2020	-0.2631	-0.2316	0.1598	1.0000	0.0940
Heldback VA	0.2974	0.2415	-0.2604	-0.0129	-0.1131	-0.3400	0.0899	0.2464	1.0000

Table A.4: Student-Year, Index of Math and English GI

	Future Held Back	Future Frac. Days Absent	Future Suspension	Graduate On Time	Don't Graduate	Leave Dataset Next Year	PSAT Score	Took PSAT
GI Factor	0.003	0.001***	0.001+	-0.007**	0.009***	0.002+	-7.001***	-0.004
	(0.002)	(0.000)	(0.001)	(0.003)	(0.002)	(0.001)	(1.364)	(0.003)
Passed GI Factor	-0.004***	-0.001**	0.000	0.004	-0.008***	-0.005***	5.335***	0.004
	(0.001)	(0.000)	(0.001)	(0.003)	(0.002)	(0.001)	(1.463)	(0.004)
VA Cog. Factor	0.001	0.000	-0.001+	-0.001	0.000	-0.001	21.389***	0.013***
	(0.001)	(0.000)	(0.000)	(0.003)	(0.002)	(0.001)	(3.047)	(0.003)
VA Non-Cog. Factor	-0.010***	-0.002***	-0.004***	0.014***	-0.014***	-0.000	0.678	-0.001
	(0.001)	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)	(1.251)	(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	0.285	0.306	0.085	0.726	0.338

Table A.5: Student-Year, Index of Math and English GI

	Missing Math Test	Missing Future Math Test	Missing English Test	Missing Future English Test
GI Factor	0.003	0.003	0.003+	-0.002
	(0.002)	(0.002)	(0.002)	(0.001)
Passed GI Factor	-0.007***	0.001	-0.004***	0.001
	(0.002)	(0.002)	(0.002)	(0.001)
VA Cog. Factor	-0.003+	-0.002	-0.001	0.001
	(0.001)	(0.002)	(0.001)	(0.001)
VA Non-Cog. Factor	0.000	-0.002	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Outcome Mean	0.15	0.24	0.10	0.19
Observations	832,002	477,270	832,002	485,680
R^2	0.160	0.404	0.133	0.520

Table A.6: Student-Year, Separate Math and English GI

	Future Held Back	Future Frac. Days Absent	Future Suspension	Graduate On Time	Don't Graduate	Leave Dataset Next Year	PSAT Score	Took PSAT
Grade Inflation (math)	0.003	0.001	0.001	-0.008**	0.007***	-0.001	-6.389***	-0.013**
	(0.002)	(0.000)	(0.001)	(0.004)	(0.002)	(0.001)	(2.269)	(0.006)
Grade Inflation (ela)	0.001	0.001**	0.001	-0.006	0.006***	0.002	-2.185	0.002
	(0.003)	(0.001)	(0.001)	(0.004)	(0.002)	(0.002)	(1.374)	(0.007)
Passed Grade Inflation (math)	-0.003+	-0.000	0.000	0.003	-0.002	-0.002	6.522***	0.009
	(0.002)	(0.000)	(0.001)	(0.004)	(0.002)	(0.001)	(1.646)	(0.006)
Passed Grade Inflation (ela)	-0.003	-0.001	0.001	0.005	-0.010***	-0.007***	0.255	0.002
	(0.003)	(0.001)	(0.001)	(0.004)	(0.003)	(0.002)	(1.858)	(0.008)
VA Test (math)	0.001	0.000	0.000	-0.008***	0.008***	0.000	14.488***	0.002
	(0.001)	(0.000)	(0.000)	(0.003)	(0.002)	(0.001)	(2.751)	(0.004)
VA Test (ela)	-0.001	-0.001**	-0.002***	0.012***	-0.011***	-0.001	20.494***	0.021***
	(0.001)	(0.000)	(0.001)	(0.003)	(0.002)	(0.001)	(3.362)	(0.005)
VA Non-Cog. Factor (math)	-0.006***	-0.002***	-0.002***	0.007***	-0.009***	-0.000	-1.133	-0.004+
	(0.001)	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)	(1.033)	(0.002)
VA Non-Cog. Factor (ela)	-0.011***	-0.002***	-0.003***	0.007***	-0.008***	0.000	-1.253	-0.004
	(0.001)	(0.000)	(0.000)	(0.002)	(0.001)	(0.001)	(1.515)	(0.004)
Outcome Mean	0.11	0.07	0.04	0.54	0.31	0.07	1153.33	0.42
Observations	330,426	371,018	384,801	384,082	279,449	415,677	198,918	298,397
R^2	0.152	0.312	0.057	0.303	0.324	0.087	0.742	0.383

Table A.7: Student-Year, Separate Math and English GI

	Missing Math Test	Missing Future Math Test	Missing English Test	Missing Future English Test
Grade Inflation (math)	-0.004**	0.003	-0.001	-0.002
	(0.002)	(0.003)	(0.001)	(0.002)
Grade Inflation (ela)	0.004	0.002	0.003	-0.000
	(0.002)	(0.003)	(0.002)	(0.002)
Passed Grade Inflation (math)	-0.004**	0.006**	-0.002	0.001
	(0.002)	(0.003)	(0.001)	(0.002)
Passed Grade Inflation (ela)	-0.007***	-0.004	-0.006***	-0.001
	(0.002)	(0.003)	(0.002)	(0.002)
VA Test (math)	0.000	0.000	0.002+	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
VA Test (ela)	-0.007***	-0.007***	-0.003***	-0.002
	(0.001)	(0.002)	(0.001)	(0.001)
VA Non-Cog. Factor (math)	-0.001	-0.002+	-0.001	-0.002***
	(0.001)	(0.001)	(0.001)	(0.001)
VA Non-Cog. Factor (ela)	0.002**	0.002	0.000	0.005***
	(0.001)	(0.002)	(0.001)	(0.001)
Outcome Mean	0.11	0.23	0.08	0.19
Observations	415,677	246,271	415,677	249,574
R^2	0.157	0.491	0.144	0.618

Table A.8: Student-Year, Split by Quartiles of Student 8th Grade GPA

	Future Held Back	Future Frac Days Absent	Future Suspension	Graduate On-Time	Don't Graduate	Leave Dataset Next Year	PSAT Score	Took PSAT
Below Median								
GI Factor	-0.000	0.001^{+}	0.001	-0.005^{+}	0.007***	0.001	-5.515***	-0.007**
	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)	(0.001)	(0.878)	(0.003)
Passed GI Factor	-0.003	-0.001	0.001	0.005**	-0.008***	-0.004**	3.770***	0.004
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.935)	(0.004)
Outcome Mean	0.20	0.10	0.07	0.37	0.54	0.10	1003.09	0.35
Observations	243,576	293,344	310,532	305,868	234,915	343,437	139,752	257,004
R^2	0.083	0.267	0.060	0.296	0.248	0.079	0.584	0.295
Above Median								
GI Factor	0.004**	0.001**	0.001	-0.003	0.006***	0.001	-5.270***	0.001
	(0.002)	(0.000)	(0.001)	(0.003)	(0.002)	(0.001)	(1.410)	(0.004)
Passed GI Factor	-0.004**	-0.001**	-0.000	-0.002	-0.002	-0.002^{+}	3.579**	-0.001
	(0.002)	(0.000)	(0.001)	(0.003)	(0.002)	(0.001)	(1.512)	(0.004)
Outcome Mean	0.07	0.05	0.02	0.69	0.20	0.04	1195.02	0.45
Observations	299,469	316,510	324,011	330,028	231,294	337,968	174,880	255,362
R^2	0.163	0.260	0.031	0.310	0.174	0.048	0.749	0.388

Table A.9: Student-Year, Split by Quartiles of School Average GPA

	Future Held Back	Future Frac Days Absent	Future Suspension	Graduate On-Time	Don't Graduate	Leave Dataset Next Year	PSAT Score	Took PSAT
Below Median								
GI Factor	0.005**	0.001^{+}	0.001	-0.008**	0.007***	0.005***	-8.748***	-0.004
	(0.002)	(0.001)	(0.001)	(0.004)	(0.002)	(0.002)	(1.322)	(0.004)
Passed GI Factor	-0.006***	-0.001	0.001	0.006	-0.008***	-0.007***	6.045***	0.001
	(0.002)	(0.001)	(0.001)	(0.004)	(0.002)	(0.002)	(0.969)	(0.005)
Outcome Mean	0.16	0.09	0.06	0.43	0.44	0.10	1045.10	0.38
Observations	295,415	351,253	369,550	403,336	295,487	410,334	159,861	274,330
R^2	0.114	0.289	0.063	0.277	0.296	0.093	0.671	0.312
Above Median								
GI Factor	0.000	0.001***	0.002***	-0.005	0.009***	0.000	-3.965**	-0.006
	(0.002)	(0.000)	(0.001)	(0.005)	(0.002)	(0.001)	(1.682)	(0.006)
Passed GI Factor	-0.001	-0.001***	-0.001	0.002	-0.007***	-0.002	3.031	0.008^{+}
	(0.002)	(0.000)	(0.001)	(0.005)	(0.002)	(0.002)	(1.977)	(0.005)
Outcome Mean	0.11	0.07	0.04	0.55	0.30	0.07	1161.13	0.41
Observations	334,769	380,385	393,800	429,929	302,364	421,668	181,855	288,447
R^2	0.140	0.298	0.050	0.275	0.292	0.070	0.744	0.364

Table A.10: Student-Course-Year, Only Algebra 1

	Test Score (math)	Test Score (ela)	Future Test Score (math)	Future Test Score (ela)	Held Back	Graduate in 5 Years	Took SAT	SAT Score
9th Graders								
Grade Inflation	-0.006	-0.028***	-0.056***	-0.039***	-0.012***	-0.015***	-0.012***	-6.563**
	(0.013)	(0.006)	(0.011)	(0.007)	(0.005)	(0.005)	(0.004)	(2.615)
Passed Grade Inflation	-0.012	0.011+	0.012	0.017**	-0.015**	0.007+	0.008**	2.716
	(0.011)	(0.006)	(0.010)	(0.007)	(0.006)	(0.004)	(0.003)	(2.714)
Outcome Mean	-0.03	-0.11	-0.26	-0.17	0.21	0.52	0.22	1187.57
Observations	110,336	111,243	87,069	91,234	114,411	71,678	84,997	26,042
R^2	0.437	0.606	0.357	0.548	0.207	0.259	0.216	0.600
Higher Grades								
Grade Inflation	-0.029**	-0.028***	-0.039***	-0.016	0.005	-0.016+	-0.011+	-7.656
	(0.014)	(0.010)	(0.014)	(0.010)	(0.009)	(0.008)	(0.006)	(5.678)
Passed Grade Inflation	-0.007	0.003	-0.005	0.000	-0.013+	0.008	0.011**	1.983
	(0.014)	(0.010)	(0.014)	(0.011)	(0.007)	(0.008)	(0.005)	(5.189)
Outcome Mean	-0.11	-0.44	-0.36	-0.45	0.19	0.44	0.11	1101.44
Observations	19,478	20,403	10,261	11,849	21,791	17,062	19,079	2,859
R^2	0.311	0.508	0.258	0.454	0.110	0.224	0.143	0.568