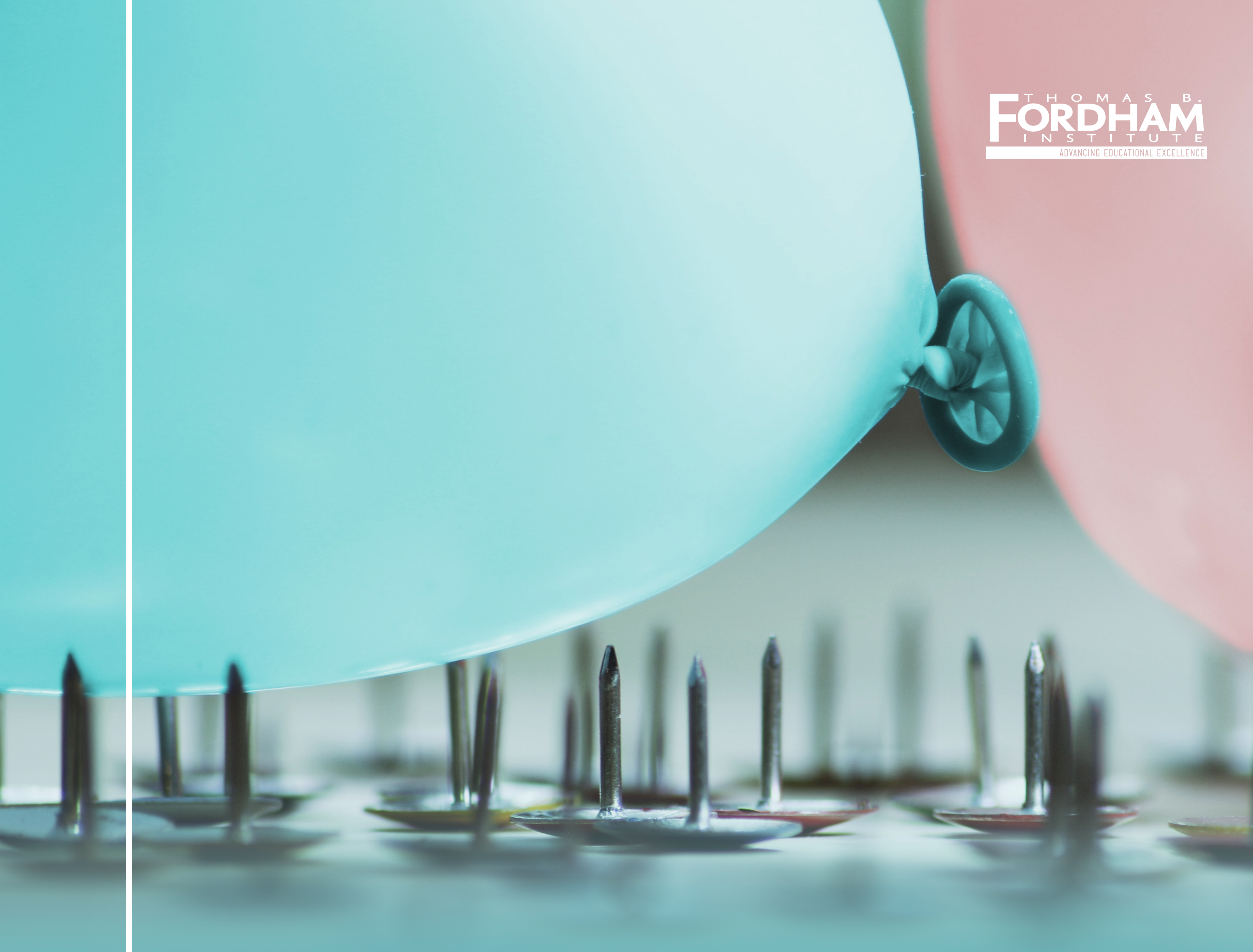


THOMAS B. FORDHAM INSTITUTE
ADVANCING EDUCATIONAL EXCELLENCE

Grade Inflation in High Schools (2005–2016)

By Seth Gershenson

Foreword and Executive Summary by Amber M. Northern and Michael J. Petrilli

September 2018

The Thomas B. Fordham Institute promotes educational excellence for every child in America via quality research, analysis, and commentary, as well as advocacy and exemplary charter school authorizing in Ohio. It is affiliated with the Thomas B. Fordham Foundation, and this publication is a joint project of the Foundation and the Institute. For further information, please visit our website at www.edexcellence.net. The Institute is neither connected with nor sponsored by Fordham University.

Contents

Foreword & Executive Summary

By Amber M. Northern and Michael J. Petrilli

Many of us, if we're lucky, can fondly recall a time in elementary school when our parents proudly posted one of our A papers on the refrigerator door. Maybe it was a spelling test or set of multiplication problems—no matter. What mattered, though, was the outstanding achievement that mom, dad, and kid believed was embodied in that A, and the pride and satisfaction that we felt in seeing it every time we opened the fridge for a sandwich.

Back then, we didn't question whether that A was actually earned. We assumed that we had mastered whatever was being graded and our hard work had paid off.

Unfortunately, it's getting harder and harder to assume that an A still represents excellence. And that's a real problem.

Here at Fordham, we've had a longstanding interest in helping to ensure that parents know the truth about how their kids are doing in school. Notably, more than a decade ago, we published *The Proficiency Illusion*—a groundbreaking study that found that levels of reading and math "proficiency" varied wildly from state to state because of where states had set their "cut-off scores." What it took to pass the state test ranged from reading or doing math at the 6th percentile all the way up to the 77th.

That's why, when the Common Core Standards and related consortia tests came onto the scene, we saw them as invitations to increased rigor, transparency, and truth-telling. Finally, parents would receive accurate, useful information about their children's academic challenges and whether they were on track for college and career. The news might not always be positive. But knowledge is power, right?

Except the message has yet to hit home. The tests are indeed tougher than ever, with *Education Next* and others finding that most states now set the proficiency bar at much higher levels than before.[1] Yet a 2018 survey published by Learning Heroes, a parent information group, found that 90 percent of parents believe their child is performing at or above grade level, and two out of three believe their child is "above average" in school. Eighty-five percent say their kid is on track for academic success—and just 8 percent believe that their child is performing below average.[2]

That's a lot of misinformed parents, given that one-third of U.S. teenagers, at most, leave high school ready for credit-bearing courses.[3]

One of us recently mused that the reason dismal state test scores don't resonate with parents is because they conflict with what parents see coming home from school:

> *Conscientious parents are constantly getting feedback about the academic performance of their children, almost all of it from teachers. We see worksheets and papers marked up on a daily or weekly basis; we receive report cards every quarter; and of course there's the annual (or, if we're lucky, semiannual) parent-teacher conference. If the message from most of these data points is "your kid is doing fine!" then it's going to be tough for a single "score report" from a distant state test administered months earlier to convince us otherwise. After all, who knows my kid better: his or her teacher, or a faceless test provider?*[4]

The teacher, of course. But what if the grades that teacher assigns don't reflect the state's high standards? What if practically everyone in that class is getting As and Bs from the teacher—but parents don't know that?

That was the impetus for this study. We wanted to know how easy or hard it is today to get a good grade in high school and whether that has changed over time. We can't develop solutions until we've accurately identified the problem. And in this case, we suspect that one reason for stalled student achievement across the land is that historically trusted grades are telling a vastly different story than other academic measures.

So we teamed up with American University economist Seth Gershenson, who is keenly interested in this topic, and whose prior research on the role of teacher expectations in student outcomes made him a perfect fit to conduct the research.

The study asks three questions:

1. How frequent and how large are discrepancies between student grades and test scores? Do they vary by school demographics?
2. To what extent do high school test scores, course grades, attendance, and cumulative GPAs align with student performance on college entrance exams?
3. How, if at all, have the nature of such discrepancies and the difficulty of receiving an A changed in recent years?

Although other studies have addressed similar questions, this is the first to use official transcript data and standardized test scores for the entire population of eligible students in a state. By including such a broad set of students, rather than a subset, Dr. Gershenson's analysis breaks new ground.

He focused on student-level data for all public school students taking Algebra 1 in North Carolina from the 2004–05 school year to 2015–16. He had access to course transcripts, end-of-course (EOC) exam scores, and ACT scores. His primary analysis compared students' course grades with their scores on EOC exams to evaluate the extent of grade inflation. The study yielded three key findings, all of which should cause concern about present-day grading practices.

Finding 1: While many students are awarded good grades, few earn top marks on the statewide end-of-course exams for those classes.

On average, students who score higher on the EOC exams also earn higher grades. But a significant number of students who receive high marks also do poorly on the EOC. In fact, more than one-third of the students who received Bs from their teachers in Algebra 1 failed to score "proficient" on the EOC exam.

Finding 2: Algebra 1 end-of-course exam scores predict math ACT scores much better than do course grades.

Since grades and EOC scores sometimes provide conflicting information, it's important to understand the extent to which these measures predict later math achievement and college readiness, as gauged by ACT scores. Results show that algebra EOCs are strongly predictive of math ACT results, much more so than course grades.

Finding 3: From 2005 to 2016, more grade inflation occurred in schools attended by more affluent youngsters than in those attended by the less affluent.

While the median GPA rose in all schools, it rose by 0.27 points in affluent schools but just 0.17 points in less affluent schools. In other words, it's gotten easier to get a good grade in more affluent schools, but not in less affluent ones. Thus, the "GPA Gap" has widened.

An analysis of ACT scores also shows that grade inflation accelerated from about 2011 onward, mostly in schools serving advantaged students. This is consistent with a similar finding from a recent College Board study that examined GPAs versus SAT scores.[5]

To us, these findings raise several implications.

First, course grades and test scores each have their place. Just because EOC scores better predict math ACT scores than do course grades, the point isn't that tests are reliable and grades aren't. Much research shows that students' cumulative high school GPAs—which are typically an average of grades in twenty-five or more courses—are highly correlated with later academic outcomes.[6]

Grades and test scores each provide valuable information because they measure different aspects of student performance and potential. Grades reflect not only students' mastery of content and skills, but also their classroom behavior, participation, and effort. And test scores tend to be informative measures of general cognitive ability. We need both.

Yet parents don't appear to value both equally. When there's a big difference between what the two measures communicate, parents are apt to take the test scores less seriously—especially if the scores are low. "My child doesn't test well," goes the refrain. In our view, this is a form of confirmation bias that's leading to greater complacency not only on the part of students, but parents too. Why should youngsters invest more time to learn something if an A or B grade says they already know it? Why should mothers work to help their children catch up if grades don't signal that they're behind? That's particularly true when parents are blissfully unaware of how widespread those good grades are. The sad fact is that some will only become aware that their child is marching off a cliff with regard to college readiness —along with many others—after it's too late.

While external exams are valuable sources of information, educating teachers about high expectations is key. Dr. Gershenson suggests in the pages that follow that one way to combat grade inflation is through content-based external tests like EOC exams. Having an external measure that is not developed or graded by the classroom teacher can be an effective way to preserve high standards, and it also serves as an "audit" of course grades and progress. That's how EOCs were used in the current analysis, and that's the role that Advanced Placement exams play for many high school students.

But what if teachers don't truly know what high standards look like? That's the issue that Success Academy's Eva Moskowitz recently raised:

> *Educators can't hold students to a high bar if they don't have a clear vision of what excellence looks like. Often teachers—and principals—have a definition of excellence that defaults to the best work produced in their classroom or school; if the "best" work is not great, expectations for all their students inevitably shift downwards....*
>
> *The battle for excellence is tough and ongoing, and merely commanding teachers to raise their standards is grossly insufficient. Ultimately, holding students to a high bar requires a zealous and persistent commitment by everyone—from superintendents, principals, and parents, to assistant teachers and office staff. Everyone must share a clear understanding of what excellence is, and give students the realistic feedback and dedicated support they need to meet the ambitious expectations of which we know they are capable.*[7]

Hear, hear! We couldn't have said it better. The question is: Are we ready to take this charge seriously?

Acknowledgements

This report was made possible through the generous support of the Bill & Melinda Gates Foundation and our sister organization, the Thomas B. Fordham Foundation. A grant from Carnegie Corporation of New York also helped to support an earlier draft of this paper presented at an invitation-only convening on the "Information Gap" in Washington, D.C., in June 2018. We are especially grateful to Seth Gershenson, who conducted the research and authored this report. His efficiency, timeliness, and collegiality made working together a pleasure. Thanks also to external reviewers Cory Koedel (University of Missouri) and Michael Hurwitz (College Board), who provided valuable input on report drafts.

At Fordham, we extend a hearty thank you to associate director of research Adam Tyner, who managed the project and asked insightful questions that improved the final product, Chester E. Finn, Jr. for reviewing drafts, Victoria McDougald and Nicholas Munyan-Penny for handling media and funder relations (respectively), and Jonathan Lutton (along with Munyan-Penney) for creating the report's layout and design. Fordham intern Emily Howell also provided assistance at various stages in the process. Kudos also to Shannon Last, who copyedited the report, as well as Grahamhaynes of GettyImages for our cover's photo.

Finally, the statements made and views expressed both in the main report and the foreword are solely the responsibility of the respective authors.

I. Introduction

In 2017, a high-profile investigation into the Washington, D.C. public school system revealed that one-third of that year's graduating class should not have received a diploma. The *Washington Post* reported that "the school system has turned virtually overnight into an embarrassment for the city and its elected leaders, who are publicly re-examining their assumptions about the system's progress."[8] At one high school, nearly two-thirds of the 177 graduates who received diplomas lacked the proper credits to graduate, had excessive absences, or completed "credit recovery" programs for courses they hadn't yet attempted.[9]

The scandal unfortunately highlights yet another instance of the now-familiar education phenomenon known as "grade inflation." It occurs when the course grades subjectively assigned by teachers do not comport with objective measures of student performance. Just as price inflation can harm individual consumers and firms and slow economic growth,[10] grade inflation can harm students, schools, and the larger education system. Accordingly, grade inflation merits serious attention from policymakers and practitioners. To date, however, it has received relatively short shrift as a K–12 education policy challenge in need of attention and remedy.[11]

This is disappointing, not least because researchers have been documenting the mismatch between school grades and external measures of student learning for decades. In fact, rising high school grade point averages (GPAs) have been accompanied by stagnant SAT, ACT, and NAEP scores, strongly suggesting lowered classroom standards.[12] And in higher education, As are now the most common grade awarded, despite constituting just 15 percent of grades in the early 1960s.[13]

Why should we bother addressing this problem? Mostly because grade inflation clouds measures of students' true knowledge and skills. This means grades may mislead students, parents, and subsequent educators—not to mention potential employers and policymakers—about how children and schools are performing and how well students are prepared for what follows, be it work or postsecondary education. A middle school student who receives high grades despite being two grade levels behind in reading might be excluded from a tutoring opportunity that would help him catch up. Later, that same student—now of college age—may find himself woefully unprepared to comprehend college-level textbooks, decreasing the likelihood that he will persist and graduate.

Grade inflation can also exacerbate socioeconomic inequities in educational outcomes when it varies systematically by student or school background. On the one hand, teachers may be more apt to inflate the grades of higher-performing, higher-income students in an effort to appease their pushy parents. On the other, teachers may be more apt to inflate the grades of low-performing, low-income students to avoid both the difficulty and associated costs of remediating them. Both scenarios widen socioeconomic gaps in postsecondary education and labor market success: they enhance the college opportunities of advantaged students and depress those of the disadvantaged students, since the latter may unknowingly graduate high school without the necessary skills to succeed later.

The extent to which course grades and course-specific test scores provide conflicting information to families, and the extent to which such conflicts vary by school background, are unclear to many in education. That's because the results of few prior studies are generalizable to the population—for example, they might look only at students who took the SAT—and are intentionally limited to specific schools and institutions.

This study, however, uses statewide data on all public school students in North Carolina who took Algebra 1 between 2005 and 2016. (See *Appendix A. How This Study Improves Upon Earlier Efforts*.) We compare these students' grades to scores on the Tarheel State's end-of-course (EOC) standardized exam. We also compare ACT college entrance exams to cumulative GPAs (using data from three cohorts of students who were required to take the ACT exam in eleventh grade).[14,15]

Specifically, the study addresses three research questions:

1. How frequent and how large are discrepancies between grades and test scores? Do these discrepancies vary by schools' demographics?

2. To what extent do test scores, course grades, attendance, and cumulative GPAs align with student performance on college entrance exams?

3. How have the nature of such discrepancies and the "difficulty" of receiving an A changed in recent years?

The report proceeds as follows: *Section II* provides additional background for the study, *Section III* briefly describes the data sources and methods, *Section IV* presents the results, and *Section V* suggests implications of the report's findings.

II. Background

The Harm in Grade Inflation

Grade inflation may occur for a variety of reasons. Teachers may have an incentive to provide overly rosy evaluations of student performance to appease students and parents or to enhance the reputation of their schools or classrooms. No matter the reason—whether it's part of a broader upward or downward drift of standards, or the result of missing guidance as to what constitutes an "A" or a "C"—grade inflation carries many potential negative consequences.

First, grade inflation may breed an unwarranted sense of complacency that leads to underinvestment in education and prevents students from reaching their full potential. For example, when a student thinks he has already mastered material that he has not, the student will not invest the extra effort and time necessary to truly learn it. Likewise, parents will not recognize the need to help their child catch up. In this case, the complacency is not due to lack of desire or ability, but rather faulty information about the student's academic standing. The consequences are real: Research on college students suggests that they study less—by as much as 50 percent—when they expect the average class grade to be an A rather than a C.[16,17]

Second, grade inflation results in promoting students to subsequent grades and later accepting them to postsecondary institutions for which they are academically ill-prepared.[18] Consequently, they struggle and risk dropping out.

Third, grade inflation misleads (and potentially harms) schools and employers who use grades as signals of ability and content mastery. It will lead them to favor—for jobs, scholarships, and admissions slots—candidates who display higher grades but who may actually possess weak skills.

Fourth, grade inflation can perpetuate and worsen sociodemographic and racial gaps in educational success when it occurs inconsistently across school types and/or pupil populations. Studies have found that recent GPA gains have been concentrated in wealthier and majority-white schools.[19] Although GPAs rose rapidly between 1990 and 2000 and continued to rise at a lower rate through 2009, the more recent growth in average GPAs occurred only for white students.

Lastly, grade inflation may have the political consequence of encouraging families to believe everything is going well at school, even when a school is troubled and needs reform. It is easy for parents to ignore systemic mediocrity when their children's grades seem strong. In this way, grade inflation may contribute to a more general sense of complacency about schools and help explain why parents tend to report satisfaction with their own schools at the same time that they hold deep concerns about the state of the country's education system.

Aside from these harms, there exists a larger debate about whether grades or test scores, such as college entrance exams, are better predictors of postsecondary success. In short, the literature shows that grades are better predictors of later success than test scores, but that the best predictions are those based on both grades and test scores because test scores provide valuable information beyond what grades can show (see *The Predictive Power of Grades Versus Test Scores*).

The Predictive Power of Grades Versus Test Scores

Numerous studies have examined the predictive power of grades and college entrance exam scores for student postsecondary success.[20] Generally, they have found that grades and college entrance exam scores are correlated with each other and that both are predictive of postsecondary success measures (such as first-year college GPA and college graduation). Although the analysis in the 2009 book *Crossing the Finish Line* is perhaps the best known, all studies in recent years have shown high school GPA to be a better predictor of college completion than entrance exam scores, and these findings have led to the commonplace belief that "grades are more important than test scores." Similar results have been found when researchers used GPA and entrance exam score data to predict first-year college GPA.

These results have contributed to anti-testing sentiment embodied in testing "opt-out" movements in K–12 and "testing optional" policies in higher education. Yet the studies on which these reactions are based have almost exclusively compared the predictive power of cumulative GPA, which reflects student performance over multiple years and many courses, with that of one-shot college entrance exams. The claim that "grades are more important than test scores" is supported by this type of comparison, but many other comparisons of grades and test scores are possible and have not been explored. Research has not established whether, for example, a single course grade is more predictive of later success than a corresponding EOC score.

Note also that grade inflation leads to what has been called "grade compression," since average grades can rise but the "ceiling" does not, resulting in a narrower distribution of grades. Researchers worry that grades may become less useful predictors of success if this trend continues.[21] The first study to examine grade compression found that high school cumulative GPA rose 0.25 points from 1996 to 2006, but the researchers found no weakening of the relationship between GPA and postsecondary success.[22] When researchers at the College Board updated the analysis from *Crossing the Finish Line* using more recent data, they found the correlation between grades and college graduation weakening somewhat,[23] although the authors of *Crossing the Finish Line* attribute some of this change to differing data sources and methodologies.[24]

Perhaps more important than the question of the predictive power of grades versus test scores is that the studies cited here have consistently demonstrated that the best predictions come from models including both of these measures. For example, a 2017 study found that grades were more predictive of adult workforce outcomes (such as income) when those test scores were incorporated into the grades, as they often are in North Carolina.[25] In sum, the research literature is clear that test scores provide valuable information beyond what grades can show.

III. Methods

Data Sources

The study utilizes administrative data from North Carolina's Department of Public Instruction via the North Carolina Education Research Data Center.[26] The data include course grades and end-of-course test scores for nearly all students in North Carolina between 2005 and 2016 who were enrolled in Algebra 1, a required course for general education high school students.

The primary analyses focus on three types of data: course transcripts, EOC exams, and ACT scores:

1. **Course transcripts.** These are files for high school classes that contain either letter (A–F) or numeric (0–100) course grades for each student. So that all students can be compared on a common metric, numeric grades are converted to letter grades of A, B, C, D, and F as follows: A = 90–100 course points; B = 80–89; C = 70–79; and D or F <70. In the analysis below, we evaluate both Algebra 1 grades and cumulative GPA, which is the average of all high school course grades according to the standard four-point scale, with no weighting for course difficulty. Note that EOC outcomes are included in course grades in some cases, and this causes the two measures to converge to some extent. (See *North Carolina High School EOC Policies*.)

2. **End-of-course exams.** Standardized, statewide EOC exam data target one common high school course, Algebra 1, primarily because this subject was consistently coded and identified throughout the study's time period. In fact, EOC tests in North Carolina have included Math I (i.e., Algebra 1) for the twelve years covered in this study.[27]

   Because the state has changed scales over time, we standardize the EOC scores by year to have a mean of zero and a standard deviation of one, which makes them comparable across years. When analyzing the ordinal levels of achievement on EOCs, we restrict the sample to 2014 through 2016, when these levels were based on a consistent scale. In these years, scores above 249 were considered "proficient at grade level" or "sufficient"; scores above 252 were "solid"; and scores above 263 deemed "superior."

3. **ACT scores.** These scores measure college preparedness. The ACT provides a composite score based on a scale of 1–36, as well as subject-specific scores in English, reading, math, and science. The ACT was mandatory for all eleventh graders in North Carolina from 2013 to 2015.

North Carolina High School EOC Policies

Subject-specific end-of-course tests in high school were implemented in the 1984 North Carolina Elementary and Secondary Reform Act and have been a part of the state's education system ever since. EOC tests have direct implications on students' grades and eligibility for graduation, the school's achievement scores, and the incentives for teachers tied to school-level growth (such as performance-based pay). Starting in 2007, North Carolina policy required students to take EOC tests in five core subjects: Biology, Algebra 1, English 1, English 2, and U.S. History. As of 2012, however, the English 1 and U.S. History EOC tests were eliminated by a state mandate that removed tests not required by federal law or federal grant receipt.[28]

State board of education policy mandates a proficient score on these EOC tests required for high school graduation.[29] Students who pass the course but fail the EOC have two chances to retake it, each time within three weeks of receiving the test result. If the student has not passed the EOC after two retakes, a committee of teachers reviews the case and the principal ultimately decides whether the student has mastered the course content or needs to retake the course.

That said, districts vary in how they handle students who retake a course. For example, Winston-Salem/Forsyth County Schools allow students who fail and retake a class to use the old EOC score, as long as it was above the proficient threshold. However, they also give students the option to retake the EOC and then use the higher of the two EOC scores when computing final grades. Students are notified of their EOC score by the end of July, via a mailed report from the school that benchmarks performance against those of same-grade peers in the same school and in the state.

Performance on the EOC factors directly into the course grade for the corresponding subject, though how this is implemented also varies widely across districts. The scores, which districts receive in time to incorporate into student grades, generally contribute a minimum of 20 percent to the final course grade, though some districts only include a score if it improves the course grade. Districts may also use different weights for students on different tracks. For example, Buncombe County Schools count EOC tests as 10 percent of the course grade for students enrolled in a career and technical education program and 25 percent for those enrolled in an academic core program.

In interpreting the analysis that follows, the reader should keep in mind that grades may be influenced by EOC scores as described here. Since the point is generally to describe cases of contrast, note that the inclusion of EOC scores in course grades makes the two measures converge somewhat. In other words, without the inclusion of EOC scores in course grades, the differences between the two measures would be even greater.

Methods

Each of the three research questions (see page 9) is addressed using the methods described in this section.

First, the report examines correlations and identifies discrepancies between course grades and EOC exam scores. The discrepancies are calculated by determining the most likely grade for a student based on a particular Algebra 1 EOC score in a given grade and year,[30] and then comparing that "expected" grade to the one actually received by the student.[31] The relationship between grades and achievement levels during the period from 2014 to 2016 is also examined (denoted as "limited," "partial," "sufficient," "solid," and "superior").

Second, the analysis investigates whether grade inflation varies by school socioeconomic status and whether that gap has changed over time. These are within-school comparisons that eliminate potentially confounding differences between schools in teacher quality, resources, average parental resources, and so on. Schools are split into more and less affluent groups based on whether or not a majority or minority of students is eligible for free or reduced-price lunch. This enables us to see whether the likelihood of receiving an A for students in the same school with the same test scores has changed over time, and whether these changes look different depending on the school's affluence.

Finally, regression techniques capture the share of variation in the ACT score that is explained by various predictors. This analysis enables us to evaluate which variables, including EOC scores, student absences, and course grades, best predict ACT scores.

IV. Results

Finding 1: While many students are awarded good grades, few earn top marks on the EOC.

As Figure 1 shows, on average, students who score higher on EOC exams also earn higher grades.[32] The correlation between EOC scores and grades is strong enough that EOC scores accurately predict about half of awarded course grades. (See *Appendix B* for additional information on the correlation between grades and test scores.)

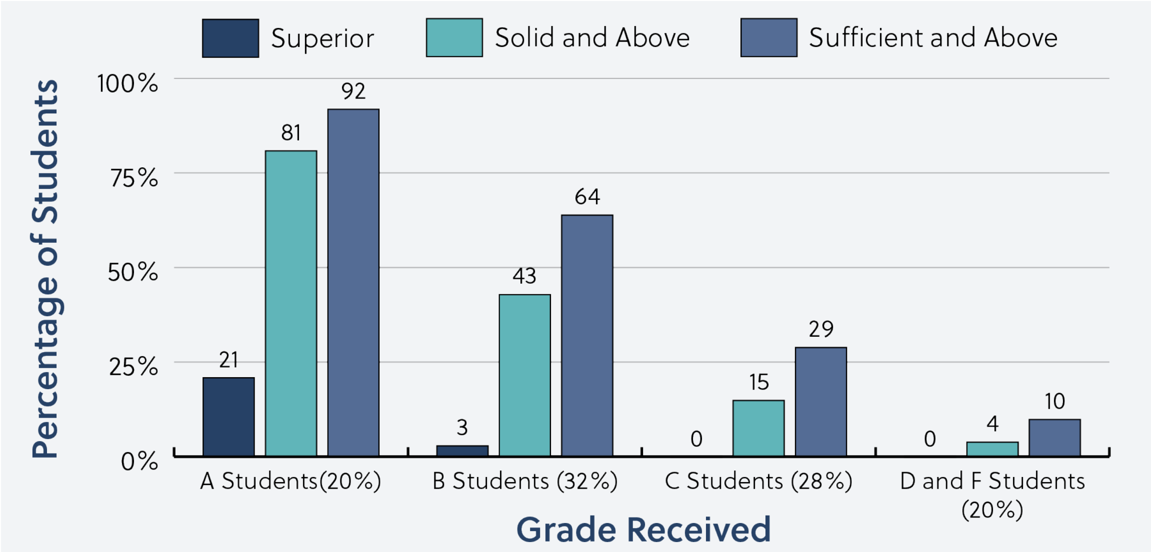
Figure 1. Students who get better grades are likely to perform better on EOCs, and vice versa.

Note: The North Carolina Department of Public Instruction administrative data includes approximately 750,000 students who took Algebra 1 for the first time in grades 8, 9, or 10 from 2005 to 2016. Letter grades are defined as follows: A = 90–100 course points; B = 80–89; C = 70–79; and D or F <70. The D and F grades are combined because they are relatively rare, which may indicate an aversion among teachers to award failing grades. The y-axis represents standardized EOC scores, with a mean of zero and a standard deviation of one.

But then there's the other half. Despite the correlation between EOC scale scores and grades, many students receive good grades while failing to demonstrate high levels of mastery on the EOC, and some students with good grades fail even to achieve simple proficiency on the EOC.

Figure 2 plots the percentage of students earning each course grade who achieve top proficiency levels on the EOCs. (See *Appendix B* for the complete distribution of grades and EOC achievement levels.) As with Figure 1, we see an intuitive pattern, whereby students who receive higher course grades are more likely to attain EOC proficiency and those receiving lower grades are less likely to do so.

Figure 2. Just 21 percent of students who receive As achieve in the highest score range on the EOC.

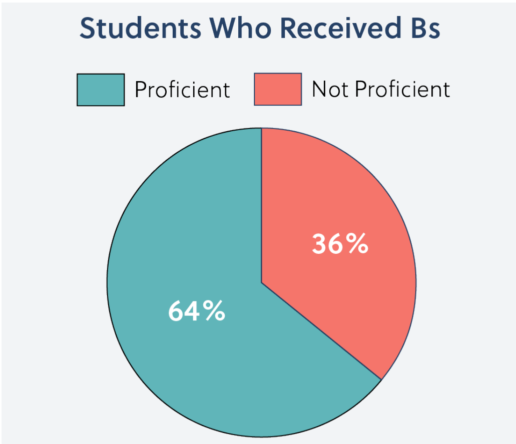


Note: Author's calculations of NC administrative data. Based on approximately 250,000 unique students who took Algebra 1 for the first time in grades 8, 9, or 10 in NC public schools from 2014 to 2016. Letter grades are defined as follows: A = 90–100 course points; B = 80–89; C = 70–79; D = 60–69; F = 0–59. Achievement levels are defined by state-determined EOC scale score thresholds of 243, 249, 252, and 263, with scores above 249 constituting grade-level proficiency.

Still, few students overall achieve mastery of the material to the point of earning the top designations on the Algebra 1 EOC. Just one student in twenty (5 percent) earns the "superior" designation on the EOC, while only one in three (34 percent) achieves the "solid" level or above, indicating college and career readiness. (See *Appendix B* for the complete distribution of grades and EOC achievement levels.) Among students with top grades, just 3 percent of students earning a B and 21 percent of students earning an A reach the highest level of achievement on the EOC.

For students receiving Bs, the larger picture is concerning. Figure 3 shows that more than one-third (36 percent) of students who received Bs failed to score "proficient" on the EOC exam. While some of these students might have simply had a bad testing day, that can't explain why so many B students are failing the corresponding EOC. Even fewer B students are meeting the "solid" level on the EOC, which indicates college and career readiness: more than half (57 percent) receiving Bs fail to demonstrate college and career readiness on the EOC. Considering that a B is generally considered to be a good grade, these findings do indeed suggest inflated grades.

Figure 3. More than one-third of B students do not reach proficiency on the EOC.



Note: Author's calculations of NC administrative data. Based on approximately 80,000 unique students who took Algebra 1 for the first time in grades 8, 9, or 10 and received a grade of B in NC public schools from 2014 to 2016. Proficiency is indicated as "sufficient or above" in the EOC.

Unsurprisingly, of students receiving a C, D, or F in the course, few earn a "superior" score (the highest proficiency level) on the EOC. For the "solid" and "sufficient" proficiency levels, a nontrivial number of students receive conflicting signals from their course grades and EOC proficiency status. Specifically, small numbers of D and F students score at the "solid" (4 percent) and "sufficient" (10 percent) levels, which indicate at least grade-level proficiency (Figure 2). At the other end of the spectrum, some students received good grades while failing to demonstrate even the lowest level of proficiency on the Algebra 1 EOC: While 92 percent of students who received As reached at least the "sufficient" level on the EOC, 8 percent of these A students did not attain proficiency on the EOC at all (Figure 2).

Finding 2: Algebra 1 EOC scores are much more predictive of math ACT scores than are course grades.

Since grades and EOC scores sometimes provide conflicting information (see Figure 3), it is valuable to examine the extent to which these academic measures are predictive of later math achievement and college readiness as gauged by math ACT scores.

Figure 4 shows that EOC scores are indeed strongly predictive of math ACT results, explaining 60 percent and 55 percent of the variation in math ACT scores in more and less affluent schools, respectively.

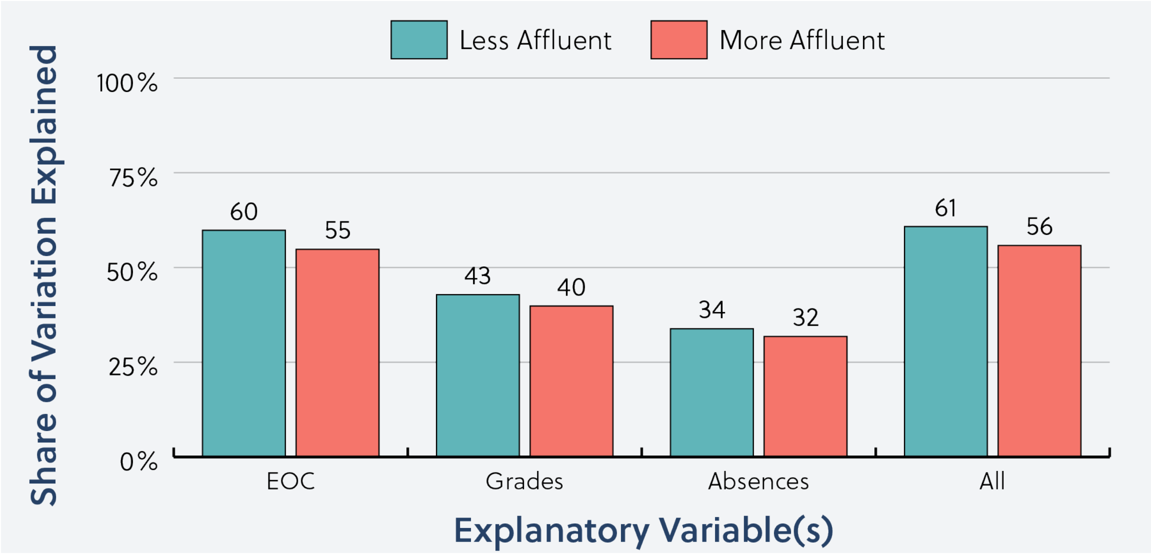
Course grades and absence rates, however, explain much less of this variation.[33] In fact, combining course grades, absences, and EOC scores explains about 61 percent of the variation in less affluent schools and 56 percent in more affluent schools, just slightly more than EOC scores alone can explain.[34] In other words, EOC scores are such strong predictors of ACT scores that course grades and absences provide virtually no additional explanatory power.

Finding 3: During the years examined here, grade inflation occurred in schools attended by more affluent students but not in schools attended by less affluent ones.

Next, we turn to the extent of grade inflation between 2005 and 2016.

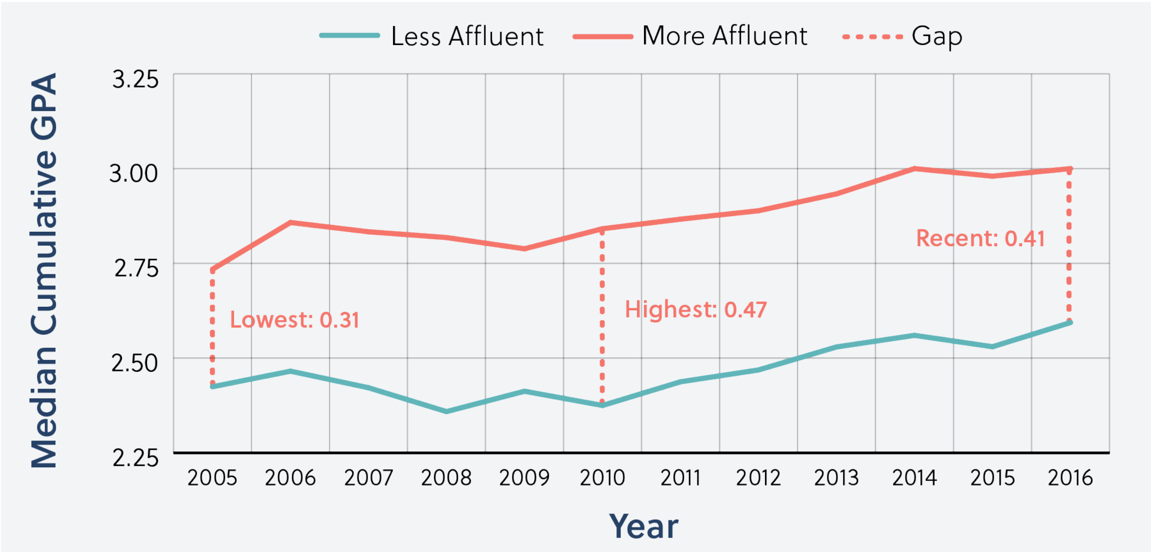
Figure 5 plots trends in the median cumulative high-school GPA (on a four-point scale) separately for more and less affluent schools. The median GPA increased in both school types between 2005 and 2016, but it increased more in the more affluent schools. The GPA growth is substantial, with the median GPA in more affluent schools increasing by 0.27 points, from 2.73 in 2005 to 3.00 in 2016, and median GPA in less affluent schools increasing 0.17 points (from 2.42 to 2.59).

Figure 4. Knowing a student's grades and absences provides little additional information about her likely ACT scores beyond what her EOC scores already tell you.



Note: Variation is measured by the R^2 from linear regressions of math ACT scores on students' EOC exam score in Algebra 1, letter grade earned in Algebra 1, and annual absences in the year students took Algebra 1. All models also control for the year and grade in which the student took Algebra 1. Less affluent schools are defined as those with more than 50 percent of students eligible for free or reduced-priced lunch; more affluent schools have less than 50 percent.

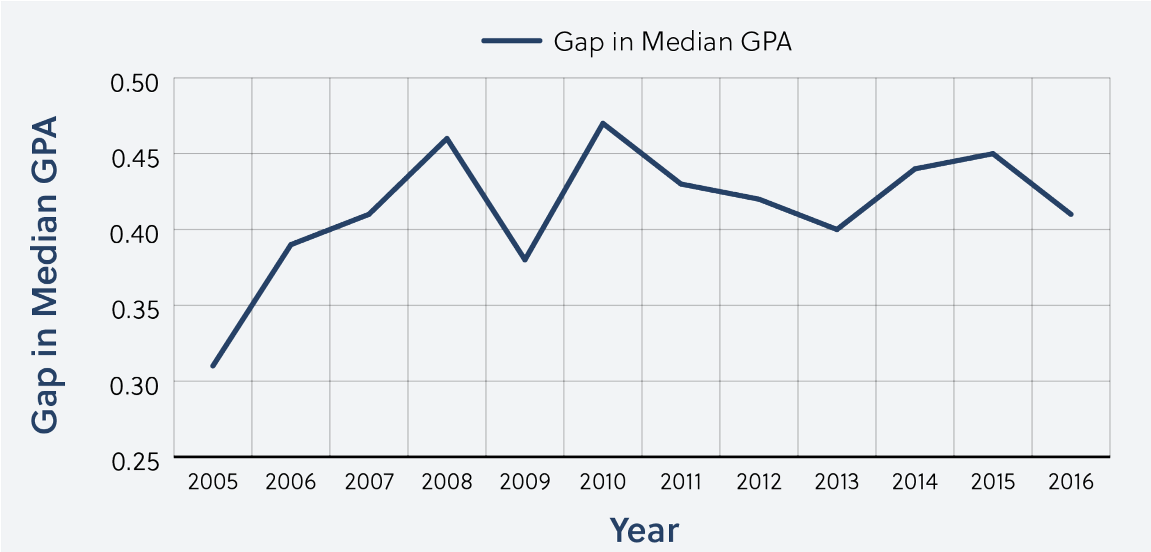
Figure 5. GPAs rose in all schools, but rose faster in more affluent schools.



Note: Less affluent schools are defined as those with more than 50 percent of students eligible for free or reduced-priced lunch; more affluent schools have less than 50 percent.

Figure 6 shows that the gap in median GPAs for more and less affluent students increased to 0.39 points in 2006 and then bounced between 0.38 and 0.47 points through 2016. Since the average GPA in less affluent schools rose at a slower pace than at more affluent schools during this period, the gap in median GPA widened as well: The 0.10 point increase from 2005 to 2016 represents a 30 percent increase relative to the 2005 baseline gap.

Figure 6. The GPA gap between more and less affluent schools has expanded.



Note: Less affluent schools are defined as those with more than 50 percent of students eligible for free or reduced-priced lunch; more affluent schools have less than 50 percent.

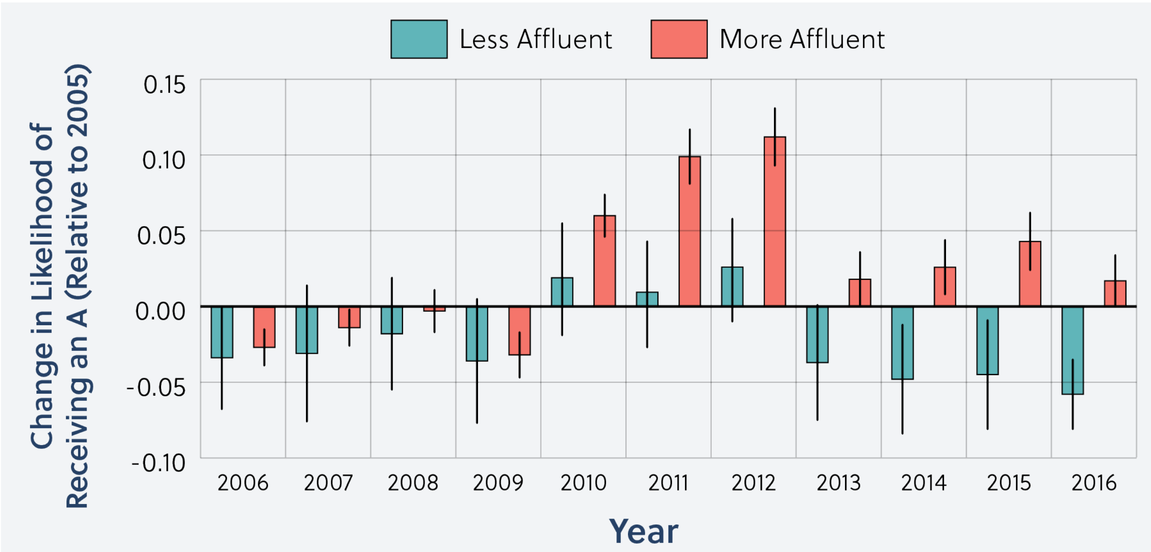
Subfinding 3.A: Algebra 1 grade increases outpace EOC increases in affluent schools.

Although GPAs rose during this period, test scores did, too, so it is not clear from the rise of GPA alone whether grade inflation is occurring. If grades rise while students actually learn more, those higher grades may not indicate grade inflation. To investigate, we first take a look at the likelihood of receiving an A in Algebra 1.

Overall, among students attending the same school and who scored similarly on the EOCs, the likelihood of receiving an A remained about constant overall between 2005 and 2016.[35] But more and less affluent schools experienced very different trends in that likelihood during this period: Starting around 2010, the probability of receiving an A in more affluent schools increased significantly, while starting in 2013 the probability of receiving an A in less affluent schools decreased significantly relative to 2005.[36,37]

Figure 7 plots changes in the probability of receiving an A separately for more and less affluent schools, relative to the baseline 2005 school year. A positive (negative) value in a given year indicates that among students in the same school and who earned the same EOC scores, the probability of receiving an A increased (decreased) relative to 2005. Because these estimates control for EOC scores, the divergence between more and less affluent schools in the likelihood of receiving an A indicates differential rates of grade inflation by school type. Since EOC score growth did not keep pace with grade increases in more affluent schools, it appears that the grade increases for students in these schools reflect actual grade inflation.

Figure 7. It has become easier for students in more affluent schools to get As while getting harder for students in less affluent schools, controlling for EOC scores.



Note: Each set of bars represents the regression-adjusted change, relative to 2005, in the likelihood of receiving an A in Algebra 1, for students in the same school who earned the same EOC score. Error bars are 95 percent confidence intervals. See Figure B.1 in Appendix B for the difference in likelihood of receiving an A over time between the less affluent and more affluent schools.

Let's see how this trend plays out in the hypothetical experience of Jack, who attended relatively affluent City High, and his friend Jill, who attended less affluent County High. As shown in Figure 7, there was essentially no difference in grade inflation through 2009 (relative to baseline 2005) between the two schools. That is, in those years, there was no change in the likelihood of receiving an A by a student in the same school with the same EOC score from 2005 for students in both City High and County High.

In 2010, however, the likelihood of receiving an A in City High started to evolve differently than it did in County High. Specifically, relative to past students in the same school with the same EOC scores, As became less likely for students in County High (the less affluent school) and more likely for students in City High (the more affluent school). In fact, by 2011, the difference in the likelihood of receiving an A between students in more versus less affluent schools became both substantively and statistically significant—at nearly 10 percentage points through 2016.[38,39]

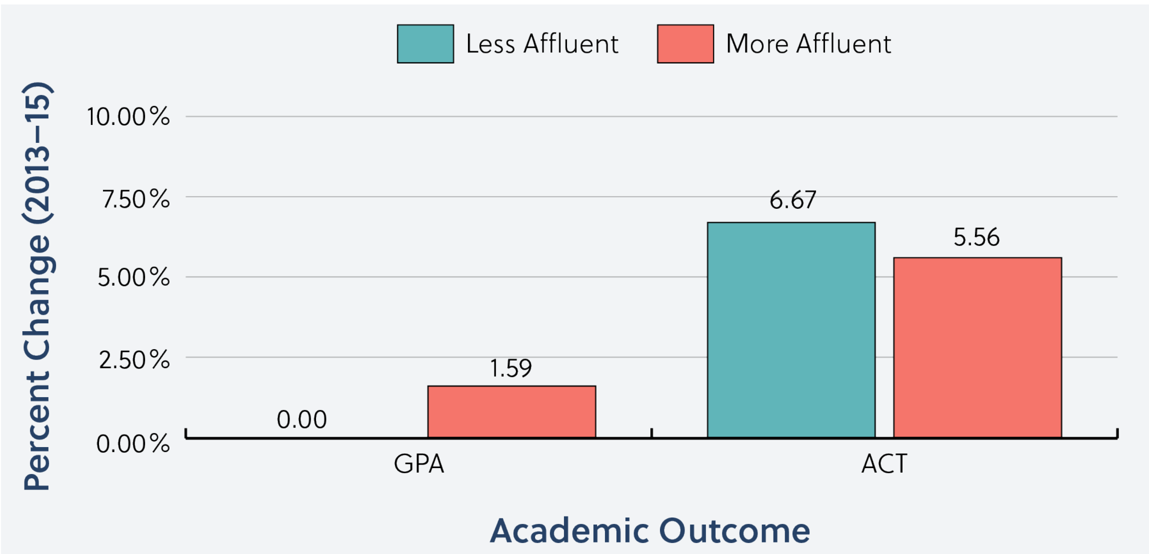
Subfinding 3.B: Cumulative GPA increases outpace ACT increases in affluent schools.

Cumulative GPAs are rising and As are becoming more common for students in affluent schools, even after controlling for EOC scores. This again suggests differential grade inflation by school type, but we take a deeper look with an analysis of ACT scores, which were available for all students for a short time period (2013 to 2015). Importantly, the ACT was taken by nearly all eleventh graders in North Carolina from 2013 to 2015, which eliminates concerns about selection bias. The idea is that if both measures—ACT scores and cumulative GPAs—move together in more and less affluent schools, we can't rule out the possibility that the cumulative GPA growth (shown in Figure 5) is due to true improvements in achievement among students.

But consistent with the patterns in our analysis of Algebra 1 EOC scores, we find that composite ACT scores increased in all schools while cumulative GPAs increased only in more affluent schools. Figure 8 shows the percent change from 2013 to 2015 for each measure, separately by school type; the GPA results on the left-hand side of the figure are calculated from the same data as Figure 5.

Median GPA increases by 2 percent in more affluent schools and remains unchanged in less affluent schools. (Over the longer eleven-year time period shown in Figure 5, GPAs rise for both groups, but ACT data are unavailable for analysis over the longer time period.) The right-hand side of the figure shows growth in median composite ACT scores. Here, we see that they grow by about the same amount in both types of schools, and actually grow a bit more in the less affluent schools.

Figure 8. ACT scores increased in all schools while GPAs only increased in more affluent schools.



Note: Less affluent schools are defined as those with more than 50 percent of students eligible for free or reduced-priced lunch; more affluent schools have less than 50 percent.

The fact that the composite ACT score, an objective measure of academic preparation for college, increased by about the same amount in all schools—during a time period in which cumulative GPAs increased only in relatively affluent schools—bolsters the claim that grade inflation occurred disproportionately in the latter. These results broadly confirm the key finding from the Algebra 1 course analyses: Grade inflation increased from about 2011 onward, but mostly in schools serving relatively advantaged students.

V. Implications

This report highlights five aspects of grade inflation (and grades in general) that teachers, parents, administrators, and policymakers ought to consider when using student grades to inform practical decisions, as well as changes in state and local policy.

First, earning a good grade in a course is no guarantee that a student has learned what the state expects her to have learned in that course.

Results show that even students who earn the best grades often fail to demonstrate mastery of key skills and knowledge when measured on the state test. Recall that just 21 percent of A students and 3 percent of B students attain the "superior" designation on the EOC, and more than one-third of B students don't reach proficiency at all. That's clearly a problem since receiving an A or B in a course signals academic success to most students and their families.

When students earn passing grades despite not mastering the academic material, a vicious cycle can follow, whereby they're set up for failure via unmerited promotion to the next course or grade level. This pattern occurs in part because we associate passing a course with accumulating hours of seat time, rather than mastering the material. Promotion based on demonstrated mastery, a model sometimes called "mastery learning," is particularly well suited to math, where foundational skills sequentially build upon one another. Adopting such approaches would go a long way toward stemming the tide of grade inflation. That said, stricter retention policies are no panacea and introduce costs of their own in terms of increased spending, behavioral problems, and higher dropout rates.[40] States and districts should keep these tradeoffs in mind when considering strengthening grade-retention policies.

Second, course grades and test scores should be viewed as complements to one another—not substitutes.

In recent years, education researchers have engaged in a wonky debate about whether course grades or test scores are more predictive of student success. The results of this report suggest that debate is a red herring.

Although they are strongly correlated, course grades and EOC scores don't always agree. This is not an indictment of either metric that, by design, each measures different aspects of student performance and potential. Grades are assigned not just based on the level of skill and content mastery, but may also include classroom behavior, teamwork, performance on intermediate assignments, and other factors. Given that Algebra 1 EOC scores strongly predict eleventh-grade math ACT scores two years later, they should be rightly viewed as highly informative measures of students' general cognitive ability and mathematics content mastery.[41] In fact, they are such good forecasters of ACT scores that course grades and absences provide virtually no additional predictive power. Yet course grades and other measures contain broader information about other skills, including "noncognitive" skills—those that help students persist and complete college—complementing the information that EOC scores provide.

Given the healthy amount of divergence between course grades and test scores, we should view them as complements, not substitutes. Taken together, they provide a fuller picture of a child's current achievement status, future trajectory, and areas of potential and improvement.

Third, grade inflation can perpetuate existing socioeconomic gaps in educational success.

Students can only be admitted to schools to which they apply, and here too, grade inflation can have unintended consequences. Grades send a signal to students, parents, and school counselors, which in turn affects students' educational aspirations, external guidance, and ultimately the perceived "choice set" of schools to which they apply. When students in more affluent schools systematically receive more optimistic evaluations of their current and future performance than their more disadvantaged peers, they will act on this misleading information. That means, among other things, that they will apply to and attend more selective postsecondary institutions. In this way, inflated grades trigger a self-fulfilling prophecy that perpetuates—even exacerbates—existing socioeconomic gaps in educational access and success.[42]

Fourth, college admissions officers should take into account how grade inflation could impact their decisions.

When making admissions decisions, postsecondary institutions should take into account that earning an A has become easier, particularly in schools serving more affluent students. Anecdotally, we know that some postsecondary institutions already do this. At minimum, these results should raise greater awareness about the need to right-size perceptions of grades at more and less advantaged schools.

A more aggressive response would be development of a formal approach to account for systematic variation across schools in grading practices—one that would ensure that all students receive fair consideration by admissions committees. That might mean, for instance, constructing something akin to the GDP deflator commonly used in macroeconomics.[43] A "grade deflator" could be constructed at the school level (as a function of school characteristics), or even at the school-subgroup level, to ensure a common relationship between EOC scores, course grades, and other measures (like attendance) for all students in all schools. For schools that are prone to award easy As (or are especially tough graders), the deflator would adjust student GPA accordingly. Obviously, explaining how and why such adjustments are made and carrying them out transparently would be key to the deflator's success.

Lastly, states should embrace EOC tests.

Scholars have long advocated for content-based external assessments such as EOCs, arguing that they are an effective mechanism for student accountability, and backing up this assertion with data showing that they can increase student effort and learning.[44] Although adopted for a range of reasons, over the past two decades, more than twenty-five states began using EOCs as tools of school and student accountability. Not only can these assessments increase student stakes in their learning, they can also serve as powerful diagnostic tools for developing strategies to identify sociodemographic disparities in learning, combat grade inflation, and preserve high standards. Finally, they can also function as an "audit" of course grades and student progress.

Yet over the past several years, some states—such as Oklahoma, Ohio, and Texas—have moved to eliminate EOCs or reduce their importance. Although state policymakers must weigh the costs and benefits of any set of policies, this study demonstrates that EOCs have tremendous value as diagnostic tools besides the benefits scholars have attributed to them in promoting student accountability. As described above, the EOCs provide complementary information to students, parents, and teachers that can help them better understand student achievement. And maintaining external assessments of student achievement enables policymakers and researchers to identify potentially inflated grades and independently gauge the extent of student learning. Given all these benefits, states should think twice before abandoning these valuable tools.

Conclusion

There's no debate that grade inflation exists. It's unequally distributed across schools. It has pernicious effects on everyone—not just on students, but also those who examine those grades for purposes of making decisions ranging from further education to employment. It is especially perilous for disadvantaged students, as it worsens sociodemographic gaps in educational and professional opportunity and success. We should beware of its presence and resist the urge to eliminate EOCs and related standardized gauges, which provide valuable, complementary information to course grades. Finally, we should work hard to ensure that the "signal" course grades send to parents and students is clearly red, yellow, or green—not inoperably blinking.

Appendix A. How This Study Improves Upon Earlier Efforts

Although previous work has documented grade inflation, several concerns have been raised about the varying methodologies and data sources. First, because GPA is a broad measure that captures both cognitive and noncognitive skills (like effort and persistence),[45] increases in GPA may be due to legitimate improvement in student effort, class participation, class behavior, and so on—as opposed to true grade inflation. Yet this alternative explanation has not been tested, likely due to a lack of reliable data on noncognitive skills. The current study attempts to address this issue by adjusting for student absences, which are an objectively observable behavior associated with noncognitive skills, effort, and school engagement.[46]

Second, the early literature on grade inflation relies on students' self-reported GPA and/or SAT/ACT data for students who chose to sit for these exams. Accordingly, extant estimates of grade inflation may be biased by both measurement error in reported GPAs and non-random selection. Thus, results from prior studies may not apply to the most disadvantaged schools in which relatively few students sit for college entrance exams. The current study mitigates these concerns by utilizing ACT scores, GPAs, EOCs, and course grades from administrative data for cohorts of students who were required by state law to take the ACT.

Appendix B. Descriptive Statistics and Model Fit

Descriptive Statistics

Table B.1 summarizes the data analyzed in this report. The main analyses focus on over one million student records of student performance in Algebra 1 courses from 2005 to 2016. Column 1 of Table B.1 shows the overall breakdown in course grades. The EOC average grade is 0.04. It is not precisely 0.00 because the scores were standardized using all available test scores, but means are only reported for students for whom course grades and some other basic data are observed. The average student was absent about six times per year in high school, and finished school with a cumulative GPA of 2.77 on an unweighted four-point scale. About half of students were male and about 60 percent were white.

About one-quarter of students in the main analytic sample attended a school in which the majority of students were eligible for free or reduced-price lunch (the measure for school affluence). Columns 2 and 3 summarize the characteristics of students separately in each type of school. Consistent with past research, in less advantaged schools, academic performance is significantly lower, and black and Hispanic students are overrepresented.

Table B.1. Descriptive Statistics at the Student Level

	All Schools	More Affluent Schools	Less Affluent Schools
	1	2	3
Less Affluent School	0.24	0.00	1.00
A (Algebra 1)	0.27	0.28	0.23
B (Algebra 1)	0.36	0.36	0.35
C (Algebra 1)	0.24	0.23	0.26
D or F (Algebra 1)	0.14	0.13	0.16
EOC Score Normalized (Algebra 1)	0.04	0.10	-0.17
	(-0.95)	(-0.94)	(-0.97)
Absences (HS yearly average)	6.00	5.80	6.50
	(-5.10)	(-4.90)	(-5.70)
Math ACT	19.80	20.1	18.60
	(-4.40)	(-4.50)	(-4.00)
Composite ACT	19.10	19.40	17.70
	(-4.80)	(-4.80)	(-4.50)
N	1,293,736	986,449	307,287

Table B.1. Descriptive Statistics at the Student Level, Continued

	All Schools	More Affluent Schools	Less Affluent Schools
Cumulative HS GPA	2.77	2.82	2.50
	(-0.71)	(-0.70)	(-0.72)
Male	0.48	0.48	0.47
White	0.61	0.67	0.40
Black	0.24	0.19	0.40
Asian	0.02	0.02	0.02
Hispanic	0.09	0.08	0.12
N	1,293,736	986,449	307,287

Notes: N reports the number of students for whom both course grades and EOC scores were observed in Algebra 1 in 2005 to 2016. The sample size for other variables is smaller, due to the fact that in certain years some students are missing data (e.g., ACT data are available in only three years). "Less affluent" is an indicator equal to 1 if the majority of a school's students are eligible for free or reduced-price lunch.

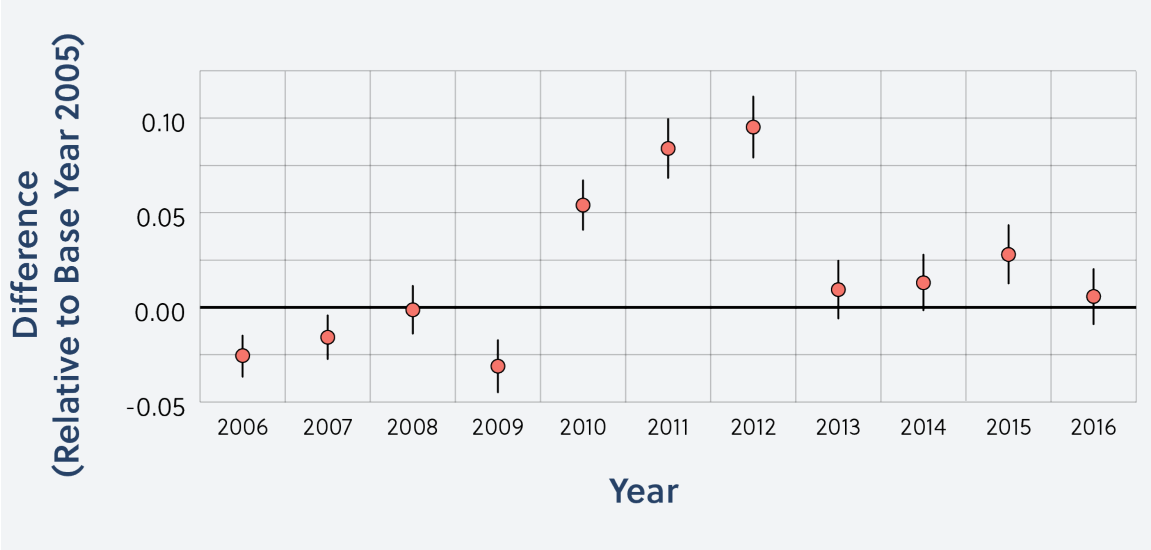
Table B.2 describes the full distribution of grades and EOC achievement levels for students who took Algebra 1 for the first time in eighth, ninth, or tenth grade in the years 2014 to 2016. The rows represent the EOC achievement levels and include information about whether the level indicates proficiency or college and career readiness according to North Carolina's standards. The columns represent course grades.

Table B.2. Algebra 1 Grades and Achievement Levels (2014 to 2016)

Proficient	College and Career Ready	EOC Achievement Level	Grade Received					
			A	B	C	D	F	Total
Yes	Yes	Superior	4%	1%	0%	0%	0%	**5%**
Yes	Yes	Solid	11%	13%	4%	1%	0%	**29%**
Yes	No	Sufficient	2%	7%	4%	1%	0%	**14%**
No	No	Partial	1%	7%	8%	3%	1%	**21%**
No	No	Limited	0%	4%	12%	9%	6%	**31%**
Total			**19%**	**32%**	**28%**	**13%**	**7%**	**100%**

While Figure 7 in the body of the report describes the changes in the share of As in Algebra 1 for each type of school separately, Figure B.1 shows the difference between these two types of schools over time. After 2010, it becomes easier—controlling for EOC scores—for students in more affluent schools to get As.

Figure B.1. After 2010, it becomes easier for students in more affluent schools to get As in Algebra 1 (controlling for EOC scores).



Note: Each point represents the change, relative to 2005, in the likelihood of receiving an A in Algebra 1, for students in the same school, who took Algebra 1 in the same grade level, and had the same EOC score. Error bars are 95 percent confidence intervals.

Model Fit

The predictive models using EOC scores to predict course grades accurately predict 51 percent of course grades. Even when EOC scores fail to accurately predict the course grade, the discrepancies are generally small. In other words, when the prediction is incorrect, it is usually just one letter grade off in either direction. Table B.3 shows the alignment between EOC scores and student grades. The columns in the table represent grades predicted by a model including EOC scores, the grade in which Algebra 1 was taken, and year fixed effects; the rows represent the students' actual course grades. Elements along the diagonal of the matrix (in bold) are cases in which the student received the predicted grade—these elements amount to 51 percent, indicating that EOC scores correctly predict slightly more than half of the students' actual course grades.

Entries below and to the left (in salmon) of the Table B.3 diagonal are cases where the grade is higher than what we would expect from the EOC score and the other information in the model. In short, EOCs predict a higher than expected course grade 22 percent of the time, meaning that, for example, the model predicted an A based on the EOC score, but the student earned a B. Entries to the upper right of the diagonal (in turquoise) are cases where the grade is lower than what we would expect based on the EOC score; for example, a student earned a C when the model based on the EOC score predicted a B.

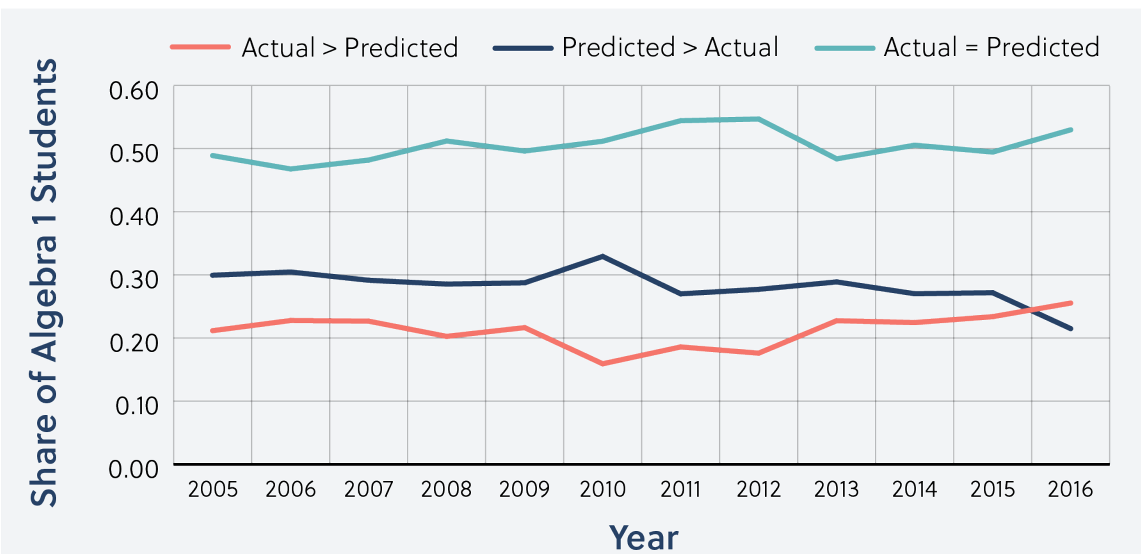
Table B.3. Most student grades can be predicted from their EOC scores.

Prediction	Actual Grade				
	A	B	C	D or F	Total
A	**14%**	6%	1%	1%	21%
B	11%	**23%**	11%	4%	50%
C	1%	6%	**8%**	5%	19%
D or F	0%	1%	3%	**6%**	10%
Total	25%	35%	24%	15%	100%

Notes: Author's calculations of NC administrative data. Based on approximately 750,000 unique students who took Algebra 1 for the first time in grades 8, 9, or 10 in NC public schools between 2005 and 2016. Letter grades are defined as follows: A = 90–100 course points; B = 80–89; C = 70–79; and D or F <70. Predictions are made by an ordered logistic regression that adjusts for EOC test scores, grade in which Algebra 1 was taken, and year fixed effects.

As can be seen in Figure B.2, over time, cases where the EOC predicted a higher grade than was awarded are more common than the opposite case in every year except for the most recent year of 2016. The gap between over- and under-predictions is fairly constant, at about 10 percentage points, from 2005 to 2009. It then widens, maxing out at about 15 percentage points in 2011 and 2012, before beginning to narrow in 2013 and eventually reversing in 2016. Because the percent correctly predicted is approximately constant, the narrowing and eventual reversal is due to increases in under-predictions that mirror decreases in over-predictions.

Figure B.2. Trends in EOC Scores' Ability to Predict Course Grades



Notes: Predictions were made using an ordered logit model, unique to each year, which adjusts EOC scores and the grade in which the course was taken.

Endnotes

1. P. Peterson and M. Ackerman, "States Raise Proficiency Standards in Math and Reading," *Education Next* 15, no. 3 (Summer 2015): 16–21, https://www.educationnext.org/states-raise-proficiency-standards-math-reading/.

2. Hart Research Associates, "Parents 2017: Unleashing Their Power and Potential," Survey, Learning Heroes, August 2017, https://r50gh2ss1ic2mww8s3uvjvq1-wpengine.netdna-ssl.com/wp-content/uploads/2017/07/LH_ParentsReport2017_20170719_MasterScreen_SinglePages_AN.pdf.

3. "The Condition of College & Career Readiness 2016," ACT, Inc. (July 2016), https://www.act.org/content/dam/act/unsecured/documents/CCCR_National_2016.pdf.

4. M. J. Petrilli, "Common Confusion: Most Kids in America Aren't on Track for Success. Why Don't They and Their Parents Know It?" *Education Next* 17, no. 1 (Winter 2017): 84–85, https://www.educationnext.org/common-confusion-state-standards-tests-proficiency/.

5. M. Hurwitz and J. Lee, "Grade Inflation and the Role of Standardized Testing." In *Measuring Success: Testing, Grades, and the Future of College Admissions*, ed. Jack Buckley, Lynn Letukas, and Ben Wildavsky (Baltimore: Johns Hopkins University Press, 2017), 64-93.

6. W. G. Bowen et al., *Crossing the Finish Line: Completing College at America's Public Universities* (Princeton: Princeton University Press, 2009); W. J. Camara and G. Echternacht, "The SAT I and High School Grades: Utility in Predicting Success," Research Notes (July 2000), 10; M. Hurwitz and M. Welch, "Comment," in J. Buckley et al., Eds., *Measuring Success: Testing, Grades, and the Future of College Admissions* (Baltimore: Johns Hopkins University Press, 2018), 226-234; J. P. Noble and R. L. Sawyer, "Is High School GPA Better than Admission Test Scores for Predicting Academic Success in College?" *College and University* 79, no. 4 (2004): 17.

7. E. Moskowitz, "'Holding Students to High Expectations is Harder than It Sounds," *Flypaper* (blog), September 7, 2018, https://edexcellence.net/articles/holding-students-to-high-expectations-is-harder-than-it-sounds.

8. P. Jamison and F. Nirappil, "Once a National Model, Now D.C. Public Schools Target of FBI Investigation," *Washington Post*, February 2, 2018, https://www.washingtonpost.com/local/dc-politics/dc-public-schools-were-once-a-success-story-are-they-now-an-embarrassment/2018/02/01/fb15dd4c-069d-11e8-b48c-b07fea957bd5_story.html?utm_term=.163ebeda99de.

9. K. McGee, "Ballou Allowed Chronically Absent Students To Graduate, Report Confirms," WAMU.org, https://wamu.org/story/18/01/16/report-confirms-policy-violations-ballou-allowed-chronically-absent-students-graduate/.

10. J. Andres and I. Hernando, "Does Inflation Harm Economic Growth? Evidence for the OECD" (Cambridge, MA: National Bureau of Economic Research, June 1997), https://doi.org/10.3386/w6062.

11. See, e.g., N. Mankiw, *Principles of Macroeconomics*, 7th Edition (Boston, MA: Cengage Learning, 2014).

12. From 1964 to 1978, GPA increased while SAT stayed constant. See I. Bejar and E. Blew, "Grade Inflation and the Validity of the Scholastic Aptitude Test," *American Educational Research Journal* 18, no. 2 (1981), 143, https://doi.org/10.2307/1162378. Similar evidence of grade inflation has been identified in the 1990s: See D. Woodruff and R. Ziomek, "High School Grade Inflation from 1991 to 2003," ACT Inc. (March 2004), https://files.eric.ed.gov/fulltext/ED484784.pdf. Researchers at ACT found less evidence of grade inflation relative to the ACT from 2004 to 2011 using data from public high schools. See also Q. Zhang and E. Sanchez, "High School Grade Inflation from 2004 to 2011," ACT Inc. (June 2013), https://eric.ed.gov/?q=grade+inflation&id=ED555577. More recently, researchers used SAT scores to document grade inflation through 2016. M. Hurwitz and J. Lee, "Grade Inflation and the Role of Standardized Testing," in in J. Buckley et al., Eds., *Measuring Success: Testing, Grades, and the Future of College Admissions* (Baltimore: Johns Hopkins University Press, 2018), 64.

13. Independent researcher Stuart Rojstaczer has compiled grading trends on his website, www.gradeinflation.com.

14. Years refer to the spring semester of the academic year (e.g., 2016 refers to 2015–16).

15. Some other subjects, including Algebra 2, Geometry, and English 1, are only observed through 2013. The reason is that the course codes used in the transcript files changed after 2013 and complicate identifying these specific courses in the more recent school years.

16. P. Babcock, "Real Costs of Nominal Grade Inflation? New Evidence from Student Course Evaluations," *Economic Inquiry* 48, no. 4 (2010), 983–996.

17. Similarly, an average teacher whose students routinely earn As might be incorrectly viewed as a highly effective teacher, and be excluded from limited spots in professional development programs, preventing the teacher from reaching his or her full potential. This type of market failure can even operate at the school level, causing parents to send their children to inferior schools.

18. This is a variant of the overmatch hypothesis. See, e.g., E. Dillon and J. Smith, "Determinants of the Match between Student Ability and College Quality," *Journal of Labor Economics* 35, no. 1 (2017), 45–66. See also Hurwitz and Lee, "Grade Inflation and the Role of Standardized Testing."

19. C. Nord et al., "America's High School Graduates: Results of the 2009 NAEP High School Transcript Study" (Washington, D.C., National Center for Education Statistics, 2011), https://nces.ed.gov/nationsreportcard/pdf/studies/2011462.pdf.

20. W. Bowen, M. Chingos, and M. McPherson, *Crossing the Finish Line: Completing College at America's Public Universities* (Princeton, NJ: Princeton University Press, 2009), https://books.google.com/books?id=LIdcVVxZb6gC; W. Camara and G. Echternacht, "The SAT I and High School Grades: Utility in Predicting Success," *Research Notes*, 10 (July 2000); M. Hurwitz and M. Welch, "Comment," in J. Buckley et al., Eds., *Measuring Success: Testing, Grades, and the Future of College Admissions* (Baltimore, MD: Johns Hopkins University Press, 2018); J. Noble and R. Sawyer, "Is High School GPA Better than Admission Test Scores for Predicting Academic Success in College?" *College and University* 79, no. 4 (2004), 17; J. Radunzel and J. Noble, "Predicting Long-Term College Success through Degree Completion Using ACT [R] Composite Score, ACT Benchmarks, and High School Grade Point Average," ACT,

Inc., 2012; P. Westrick et al., "College Performance and Retention: A Meta-Analysis of the Predictive Validities of ACT Scores, High School Grades, and SES," *Educational Assessment* 20, no. 1 (2015), 23–45.

21. W. Chan and W. Suen, "A Signaling Theory of Grade Inflation," *International Economic Review* 48, no. 3 (2007), 1065–1090.

22. E. Pattison et al., "Is the Sky Falling? Grade Inflation and the Signaling Power of Grades," *Educational Researcher* 42, no. 5 (2013), 259–265.

23. Hurwitz and Welch, 2018.

24. W. Bowen, M. Chingos, and M. McPherson, 2018.

25. G. Schwerdt and L. Woessmann, "The Information Value of Central School Exams," *Economics of Education Review* 56 (2017), 65–79.

26. Founded in 2000, the North Carolina Education Research Data Center provides researchers and the broader policy community with access to data for policy-oriented research. It cleans, codes, and standardizes longitudinal data files and makes these files available to approved researchers. The center is housed in the Social Science Research Institute at Duke University and continues to process new data released by the DPI annually. For more on the NCERDC, see C. Muschkin et al., *North Carolina Education Research Data Center Grant #200300138: Final Report to the Spencer Foundation* (June 15, 2011), https://childandfamilypolicy.duke.edu/pdfs/projects/NCERDC_SpencerFoundationReport.pdf.

27. The North Carolina Elementary and Secondary Reform Act of 1984 established EOC exams to measure student mastery in particular subjects.

28. North Carolina State Board of Education website, http://www.ncpublicschools.org/accountability/reasonselimtests.

29. Beginning with the ninth-grade class of 2009–10, North Carolina increased student requirements for high school graduation by launching its Future Ready Core program. The main change was the additional requirement of a fourth math class; prior cohorts could graduate with only three math credits. For ninth graders entering in the 2012–13 school year, the academic requirements for graduation were again ratcheted up by requiring a fourth social studies credit. The impact of the reform on graduation rates is unclear, since the raised standards coincided with a large increase in North Carolina's graduation rate, from 68 percent for the 2006 class to 86 percent of the 2016 class.

30. The Algebra 1 EOC test has a total of 280 possible points. Thresholds set by the state transform the scale score into one of five categories, the top three of which are deemed "proficient at grade level," though only the top two are considered "college and career ready."

31. The modal "on time" student takes Algebra 1 in ninth grade, but substantial shares of the student body take the test either one grade earlier or later, which can influence performance. The analysis is limited to students who took Algebra 1 in grades 8, 9, or 10. See E. Parsons et al., "Incorporating End-of-Course Exam Timing into Educational Performance Evaluations," *Journal of Research on Educational Effectiveness* 8, no. 1 (2015), 130–147.

32. This pattern is consistent over time and by school type. That is, reproducing Figure 1 separately by school year or by school type yields similarly downward-sloped gradients.

33. Recall that EOCs are sometimes incorporated into course grades. Since EOCs are a much better predictor of math ACT scores than course grades, some of the predictive power of the course grades is coming from the inclusion of EOC scores. This means that if EOC scores were never included in course grades, the course grades would be even worse predictors of math ACT scores.

34. These basic results do not vary by students' race or sex. Students' self-reported college intent is equally well predicted by EOC scores and course grades.

35. This result is displayed in *Appendix B*, Figure B.2.

36. It is also interesting that this gap emerges in 2011, and a brief increase in grade inflation emerges between 2010 and 2012 in *Appendix B* Figure B.2, as 2010 is the first year of North Carolina's Future Ready Core change to the high school graduation requirements. Students who entered ninth grade in 2010 were the first cohort required to take Algebra 1 (or Integrated Math 1) to graduate. In 2013, the graduation requirements changed again, though the math requirement did not change. See http://www.ncpublicschools.org/curriculum/graduation/.

37. The modal "on time" student takes Algebra 1 in ninth grade, but substantial shares of students take the test either one grade earlier or later, which can influence performance.

38. Identifying the cause of this divergence is outside the scope of this report, though is a worthy area for future research. That said, it is interesting that these divergent patterns coincide with reforms in North Carolina's graduation requirements (see *North Carolina High School EOC Policies*). However, it is unlikely that the change in graduation rates fully explains the changes in grade inflation, particularly in more affluent schools. Still, regardless of whether differential grade inflation occurred as a result of changes in graduation requirements, the bottom line is that As became more plentiful in more affluent schools, even after accounting for objective performance differences on the EOC.

39. This pattern holds even after adjusting for student absences. This means that the widening of the inflation gap between more and less advantaged schools was not driven by differential changes in attendance patterns. More importantly, absences are a proxy for more general noncognitive skills, or effort, meaning it is less likely that differences in noncognitive skills are driving divergent outcomes for less and more affluent schools.

40. B. Jacob, "The Wisdom of Mandatory Grade Retention," *Education Digest* 82, no. 7, 29.

41. That said, the correlation is at least partly driven by students' general test-taking skills.

42. For an overview of related evidence on the role of biased expectations in creating self-fulfilling prophecies, see S. Gershenson and N. Papageorge, "The Power of Teacher Expectations: How Racial Bias Hinders Student Attainment," *Education Next* 18, no. 1 (2018), 65–70, http://educationnext.org/power-of-teacher-expectations-racial-bias-hinders-student-attainment/.

43. See, e.g., "Frequently Asked Questions" (Washington, D.C.: Bureau of Economic Analysis), https://www.bea.gov/help/faq.

44. J. Bishop, "Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn—Externalities, Information Problems and Peer Pressure," *Handbook of the Economics of Education*, 2 (2006), 909–44; J. Bishop, "Do Curriculum-Based External Exit Exam Systems Enhance Student Achievement?" CAHRS Working Paper Series (Ithaca, NY: Cornell University, November 1, 1997); H. Jürges et al., "The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany," *Journal of the European Economic Association* 3, no. 5 (2005), 1134–1155.

45. In some cases, GPA also includes non-academic course grades such as gym, though the increase in GPA observed over time is present even in math-only GPA.

46. See, for example, S. Gershenson, "Linking Teacher Quality, Student Attendance, and Student Achievement," *Education Finance and Policy* 11, no. 2 (2016), 125-149.