Brandon Lehr*

Information and Inflation: An Analysis of Grading Behavior

DOI 10.1515/bejeap-2015-0138

Published online February 19, 2016

Abstract: I study the impact on grades assigned at Occidental College, a selective private liberal arts college, following the introduction of a policy that provides information about average grades across campus to instructors each semester. Using transcript level data from 2009 to 2014, I find that after the information provision, previously below average grading courses increased grades by 0.08 grade points more than the previously above average grading courses. This finding of grade compression holds across all course levels and divisions, except for in the sciences. With respect to students, the relative increase in grades in the previously low grading courses disproportionately benefited Black and Hispanic students relative to White and Asian students. In addition, the grade distribution shifted with previously below average grading courses increasing the share of A's and decreasing the share of B's and C's following the grade information provision.

Keywords: grade compression, grade inflation, information provision, higher education

JEL Classification: I23, D83, D03

1 Introduction

There has been a dramatic increase in average grades at colleges and universities in the United States over the past 50 years. Summarizing the evidence, Babcock (2010) notes that, save for a period of nonincreasing grades in the 1970s, from 1960 to 2004 grade point averages (GPAs) increased by approximately 0.15 points per decade. This trend is consistent with the GPA across all private and public schools increasing from 2.4 (just above a C+) in 1960 to 3.0 (a B) by 2006 (Rojstaczer and Healy 2010). Focusing on the change from the 1980s to the 1990s, Kuh and Hu (1999) find the most dramatic grade inflation occurring at research universities and selective liberal arts colleges.

*Corresponding author: **Brandon Lehr**, Department of Economics, Occidental College, 1600 Campus Road, Los Angeles, CA 90041, USA, E-mail: lehr@oxy.edu

The distribution of grades is also notable, with A's constituting 43 % of all letter grades in 2008, a 28 percentage point increase since 1960, and D's and F's totaling less than 10 % of all letter grades (Rojstaczer and Healy 2012).

Regardless of the reasons for this widespread grade inflation, if grades motivate the behavior of students and the perceptions of students by employers and graduate schools, then such changes can have widespread consequences. There is evidence that grades play a role in students' choices of courses (Sabot and Wakeman-Linn 1991) and effort (Babcock 2010). And although college major choice is a complex decision that depends on idiosyncratic and heterogeneous tastes (Wiswall and Zafar 2015), grade differences across departments likely play a role in this decision. From the perspective of those who will evaluate and rank students, with a maximum grade of an A, grade inflation induces compression at the top of the distribution, potentially making it more difficult for students to signal their ability relative to others.¹ Although even with the majority of grades allocated across A's and B's, the continuous nature of cumulative GPAs may still permit a sufficiently fine ranking of students (Millman, Slovacek, Kulick and Mitchell 1983).

Given the potential distortionary impacts of grade inflation, a number of institutions have adopted varying policies to address this issue either directly or indirectly. Princeton adopted a grade deflation policy in 2004 that recommended that no department give more than 35 % A's overall, only to have it reversed 10 years later. Wellesley College similarly in 2004 implemented a policy by which average grades in courses at the introductory and intermediate levels with at least 10 students should not exceed a B+ average. The effects of this policy, which is still in effect, are examined by Butcher, McEwan and Weerapana (2014). In a less direct way, Cornell University sought to better contextualize course grades by adopting a policy starting in 1998 by which course median grades would be published online, the consequences of which are analyzed by Bar, Kadiyali and Zussman (2009).

¹ Empirically, Summary and Weber (2011) find that increases in grades are associated with a decline in the information content of those grades, as measured with Shannon's entropy index. This idea has also been explored theoretically. Chan, Hao and Suen (2007) construct a signaling model where grade inflation obtains in equilibrium and moreover can induce other schools to inflate grades. Ostrovsky and Schwarz (2010) characterize equilibria where it is optimal to obscure or conceal information about job candidates, one means by which is to reduce the variation in grades assigned. And Boleslavsky and Cotton (2015) consider the strategic choices of schools to choose grading standards that conceal information about graduates and invest in their own school quality.

Turning to the analysis in this paper, I study the impacts of a related policy at Occidental College, a small private liberal arts college in Los Angeles, adopted after grade assignments in the Fall semester of 2012. In particular, Occidental began providing to each instructor at the end of the semester Grade Comparison Sheets, one for each course. These Grade Comparison Sheets contain, among other summary statistics, the college GPA and the GPAs in the course, the course's division, and the course's department, both at all levels and at the same course level (e.g., 100-level introductory courses).² Such information provision was intended to help faculty understand how their grading behavior compared with that of their colleagues in similar courses and disciplines, without formally imposing a quota on grades, as pursued by Princeton and Wellesley.

Using transcript-level data of every grade assigned between the Fall of 2009 and Fall of 2014 at Occidental, I find that the provision of this grade information to instructors generated some unintended consequences. The introduction of the Grade Comparison Sheets evolved out of concerns related to grade inflation at Occidental and in particular the large variation in grading practices across departments, which could be incentivizing students toward majors with higher grades. By providing grade information, one hypothesis was that the information would anchor expectations about grading norms, and those assigning the highest grades would consequently move to lower their average grades. In fact, I find that with the introduction of the grade information, the below average grading courses³ increased grades by 0.08 grade points more than the above average grading courses, which in one preferred specification show no evidence of average grade changes.⁴ This qualitative finding holds across all course levels and divisions, except for in the sciences. With respect to students, the relative increase in grades in the previously low grading courses disproportionately benefited Black and Hispanic students relative to White and Asian students. In addition, the grade distribution shifted with previously below average grading

² All reported GPAs, except for own course GPA, are calculated over observations from the previous 5 years.

³ By “below average grading courses” I mean the courses with average grades below the average grades in the corresponding division-level pair in the pre-information period.

⁴ Although these estimated effects of the Grade Comparison Sheets are statistically significant at conventional levels, the effects are sufficiently small in magnitude that I am unable to detect shifts in student behavior, such as changing enrollments. This is in contrast to the analysis in Butcher et al. (2014), which is closely related, but studies a hard policy cap on grades at Wellesley that shifts grades more dramatically and subsequently generates changes in enrollments and major choices.

courses increasing the share of A's and decreasing the share of B's and C's following the grade information provision. A consequence of the main result is that although overall grades are not lower in the post-information period, the information policy did help to reduce grade dispersion.

The finding that grading behavior is responsive to information about average grades at the college reveals that the grading practices of others are a relevant part of the grade allocation decision-making process, but does not identify the particular mechanism generating this behavior. Within a course, particular student grades likely depend on relative ability, attitude, or maturity, as well as student characteristics like race, ethnicity, and gender (van Ewijk 2011; Ouazad and Page 2013). However, an instructor must also determine a course-level average grade, which I find depends on knowing the grades assigned elsewhere in the college. Previously low grading instructors, upon learning their position in the distribution, may increase this average grade for a number of reasons: to improve student evaluations that are relevant for the instructor's promotion, to attract more students to a department, or to increase the relative competitiveness of her students when they apply for scholarships, employment, or graduate school.⁵ Alternatively, if the college-wide grade distribution is a public good that provides information about student quality to students, faculty, and outside evaluators, then making grade information public to faculty reveals that high graders are free riding on the production of this information content, inducing low graders to raise their grades to free ride as well.⁶

The analysis of information provision in the context of grades in higher education is related to a broader literature on the role of information as a policy lever. In the setting of education more generally, providing low-income families with information about school quality impacts school choice (Hastings and Weinstein 2008) and providing low-income high school students with information about college costs and the application processes increases their applications and admittance to college (Bettinger et al. 2009; Hoxby and Turner 2013). In addition, there is evidence that information on the returns to education reduces dropout rates (Jensen 2010) and increases test scores (Nguyen 2008). The importance of information for health-related decisions has been investigated in a range of settings. For example, the direct provision of already free and

⁵ Johnson (2003) surveys the literature on grading practices and finds robust evidence for the impact of higher grades on better course evaluations and for students choosing courses with higher grades to improve their personal GPA.

⁶ This thoughtful observation was suggested by an anonymous referee.

widely advertised information about Medicare Part D prescription plans can generate substantial welfare gains (Abaluck and Gruber 2011; Kling et al. 2012), providing information about the risk of human immunodeficiency virus (HIV) infection can dramatically decrease teen pregnancy rates (Dupas 2011), and nutritional labeling can improve the quality of consumer choices (Abaluck 2011; Bollinger et al. 2011). The public finance literature has explored the impact of information on enrollment in retirement plans (Duflo and Saez 2003), as well as on labor supply decisions with respect to learning about the Earned Income Tax Credit (Chetty and Saez 2013) and Social Security (Liebman and Luttmer 2015). Notable about these last two studies is that like this paper, they consider the impact of information on the behavior of suppliers instead of consumers, where the suppliers in my setting are teachers producing an education for the consumer students.

An important feature of this literature on information is carefully distinguishing between an information provision effect and a salience effect. For example, by posting tax-inclusive prices on grocery items, the subsequent decrease in demand may be due to both a salience effect and an information effect if shoppers mistakenly believe there is no tax (Chetty et al. 2009; Zheng et al. 2013). In the present study, it is likely that the mechanism driving the responsiveness to the Grade Comparison Sheets is entirely due to an information effect, as grade averages across the college, let alone disaggregated by department, division, or level, were not publicly available.⁷

The rest of this paper is organized as follows. Section 2 outlines the data and institutional context for the introduction of the Grade Comparison Sheet policy. Section 3 outlines the empirical strategy for the analysis and Section 4 presents the results. Section 5 concludes.

2 Background

2.1 Data

The data used in this analysis were supplied by the Office of the Registrar at Occidental College. Each observation is at the transcript level, spanning the grades assigned in all courses at the college between Fall 2009 and Fall 2014, for a total of 89,502 observations (from a student body that averages about 2,100

⁷ Even now, average grades by department are not publicly available at Occidental.

full-time equivalent students). Associated with each grade is the corresponding course information, which includes the department, division, level, units, and number of enrolled students.⁸ In addition to the grades earned by each student, the data set includes the following student demographic variables: ethnicity/race, standardized test scores. Scholastic Aptitude Test (SAT) and/or American College Testing (ACT), and whether or not the student is a legacy or first-generation student.

I make three restrictions to the original data to arrive at my core sample. I first restrict observations to grades distributed in non-summer courses.⁹ Next, I restrict attention to 100-, 200-, and 300-level courses with more than one student. The 0-level courses indicate those that belong to the cultural studies component of the Core Program, which is not in fact a proper major with dedicated faculty, but primarily an assortment of first-year writing seminars taught by a changing sample of faculty from across the college. Since my analysis concerns grading behavior by faculty within departments over time, I drop classes offered in this cultural studies program. I also drop 400- and 500-level courses, the former of which designate senior seminars and the latter courses intended for graduate students.¹⁰ The restriction to include only those classes for which more than one student is enrolled is to omit independent studies or study abroad classes, which are coded with zero enrolled students, since grading behavior is likely to be more idiosyncratic in such nonstandard course settings. Finally, I drop grades in courses that did not exist any time before the Grade Comparison Sheets were introduced in Fall 2012, as well as grades in departments that did not exist after Fall 2012.¹¹

This trimmed data set consists of 5,046 unique students and 69,635 grades, the summary statistics for which are displayed in Tables 1 and 2. The student body is 54.1% White, 16.3% Asian, 12.7% Hispanic or Latino, and 6.3% Black or African-American. Relative to the national average at 4-year colleges in 2012, Occidental has fewer White (61.5%), Hispanic (16.0%), and Black (13.8%) students, but more

8 The course itself is not identified for privacy reasons, but a random course identifier allows me to track courses across instructors and over time.

9 Only a limited number of courses are offered during the summer. There are an average of 70.4 grades each summer term, and over 80 % of the summer grades in the sample are from just two departments: American studies and education.

10 More than 80 % of the 500-level grades in the sample were issued by the education department, which ended its Master's program in Summer 2012.

11 The last psychobiology and Asian studies courses were offered in the 2009–2010 academic year and since Spring 2013 courses in two newly formed departments (Latino/a and Latin American studies and writing and rhetoric) have been introduced.

Table 1: Student-level summary statistics.

Variable	Obs.	Mean	Std. dev.	Min.	Max.
Ethnicity/race					
White	5,046	0.541	0.50	0	1
Asian	5,046	0.163	0.37	0	1
Hispanic/Latino	5,046	0.127	0.33	0	1
Black/African-American	5,046	0.063	0.24	0	1
American Indian/Alaskan Native	5,046	0.017	0.13	0	1
Native Hawaiian/Other Pacific Islander	5,046	0.012	0.11	0	1
Nonresident alien	5,046	0.052	0.22	0	1
Unknown	5,046	0.026	0.16	0	1
Legacy	5,046	0.087	0.28	0	1
First-generation	5,046	0.144	0.35	0	1
SAT verbal	4,160	640.43	72.16	320	800
SAT math	4,160	640.72	71.03	300	800
SAT writing	4,096	644.99	70.12	350	800
ACT	2,173	28.40	3.09	10	36

Table 2: Course and grade-level summary statistics.

Variable	Obs.	Mean	Std. dev.	Min.	Max.
Enrolled	3,808*	17.82	10.158	2	112
Units	3,808*	3.669	0.905	1	8
Grade points	69,635	3.313	0.674	0	4
Grades					
A	69,635	0.282	0.450	0	1
A–	69,635	0.204	0.403	0	1
B+	69,635	0.166	0.372	0	1
B	69,635	0.160	0.367	0	1
B–	69,635	0.080	0.271	0	1
Any C	69,635	0.093	0.290	0	1
Below C–	69,635	0.015	0.123	0	1
Division					
Arts and humanities	69,635	0.355	0.479	0	1
Sciences	69,635	0.353	0.478	0	1
Social sciences	69,635	0.292	0.455	0	1
Level					
100 level	69,635	0.426	0.494	0	1
200 level	69,635	0.322	0.467	0	1
300 level	69,635	0.252	0.434	0	1

*The summary statistics for the enrollment and units of each class are computed across the 3,808 courses identified in the data set.

Asian (6.5%) students.¹² In addition, 8.7% of the student body are legacy students and 14.4% are first-generation students. Most (97.5%) students have a recorded SAT or ACT score. The average SAT math and verbal composite score is 1,281 (1,926 with writing component) and the average ACT score is 28.4, in which using a linear interpolation of the concordance table provided by the ACT is equivalent to a SAT math–verbal composite score of 1,276. These average scores are near the 90th percentile for college-bound seniors.

The grade-level data show that the GPA¹³ in the college across all semesters in the data set is 3.31,¹⁴ corresponding to just above a B+. This average is characterized by a distribution in which 48.6% of the grades are either an A or A– and only 10.8% of the grades are C+ or below. The arts and humanities and sciences divisions distribute the largest number of grades, with the social sciences division distributing the least. The number of grades assigned is also decreasing in the course level, as would be expected with larger introductory and smaller advanced courses.

There exists substantial variation in average grades across departments, divisions, and levels in the pre-information period (Fall 2012 and earlier). Figure 1 illustrates the deviation of departmental GPAs from the pre-information college GPA of 3.312.¹⁵ The range spans a GPA of 2.95–3.64. There is a strong divisional correlation with departmental GPAs for the sciences and the arts and humanities. All but one of the science departments assigns grades below the college average and all but one of the arts and humanities departments assigns grades above the college average. The social sciences are more evenly distributed, with three of the eight departments assigning below average grades. To anticipate the forthcoming econometric analysis, Figure 2 presents the same breakdown of departmental GPAs, but in the period after the grade information has been provided to instructors (where the college GPA in this post-period increased slightly to 3.316). Visual inspection reveals that in general the

12 U.S. Census Bureau, ‘Type of College and Year Enrolled for College Students 15 Years Old and Over, by Age, Sex, Race, Attendance Status, Control of School, and Enrollment Status: October 2012,’ Table 5, <https://www.census.gov/hhes/school/data/cps/2012/tables.html> (accessed January 30, 2016).

13 All reported GPAs throughout the paper are weighted averages across all grades (within the college, department, division, level, or student demographic group as specified), with weights equal to class units.

14 GPAs at Occidental are calculated on a 4-point scale, with the highest grade of A equal to 4 points and the lowest grade of F equal to 0 points. Plus and minus modifications to a grade adjust the numerical score up or down by 0.3 points, respectively. The only exceptions to this rule are that there are no D– or F+ grades permitted.

15 Department labels are suppressed as per the request of the college.

Figure 1: Pre-information grade differences.

Figure 2: Post-information grade differences.

Table 3: Pre-information average grades by division and level.

	Arts and humanities	Sciences	Social sciences	All divisions
100 level	3.45	3.11	3.21	3.26
200 level	3.38	3.16	3.36	3.32
300 level	3.50	3.33	3.37	3.40
All levels	3.43	3.19	3.31	3.31

magnitude of the departmental deviations from the college mean decreased, although the highest grading department deviated even further from the college mean in the post-information period. Table 3 tabulates average grades in the pre-information period across divisions and levels. The sciences assign the lowest grades, with an average of 3.19, followed by the social sciences at 3.31, and finally the arts and humanities at 3.43. The differences are also substantial across levels, with grades increasing in the level: 100 level at 3.26, 200 level at 3.32, and 300 level at 3.40. To better visualize the variation in grades across departments and levels, Figure 3 plots the deviation of average grades in a department-level pair from the average grades in its corresponding division-level pair. While most departments have little variation in the sign of their grade deviations by level, there are some notable exceptions; for example, the lowest grading department on campus, a science department, offers higher than average grades at the 200-level relative to 200-level science courses.

2.2 Grade Comparison Sheets

A campus-wide discussion on grades beginning in May 2012 eventually culminated in the provision of Grade Comparison Sheets for each course at the end of the Fall 2012 semester. In May 2012, a half-day Teaching Faculty Retreat on Academic Culture was held on campus. As articulated in an email announcing the retreat to the faculty, the purpose was “to engage all of us (the Teaching Faculty), in a discussion about ways to improve the academic culture at our College.” The first item on the agenda was a presentation by the Director of Institutional Research, Assessment & Planning entitled “Academic Environment & Achievement at Oxy.” The report included information about student and faculty behaviors and attitudes. As part of this report, GPAs and the share of A/A– grades by department broken into upper and lower division courses were presented, with the report subsequently emailed to the faculty. A discussion among the faculty present at the retreat covered issues related to grade inflation,

Figure 3: Pre-information grade differences across departments by level.

the large and perhaps distortionary impacts of differential grading practices across campus, as well as the difficulty of mandating a cap on grades, as was done by Wellesley. Many faculty admitted that they had never received any guidance within their department about norms for grades and would be reluctant to share their grading behavior with their colleagues.

To follow up on the issues raised in the retreat, a faculty committee on Academic Culture and Intellectual Life was constituted in Fall 2012. This committee, in conjunction with the Center for Teaching Excellence, hosted a lunch in September for faculty to again talk about grading philosophies and practices. After discussions such as these with faculty from across campus, the committee presented a review of their findings at the November Faculty Meeting. They discussed the lack of information about grading practices, particularly in light of the fact that grading practices play some role in the promotion and tenure of faculty. To address this, the committee proposed and the faculty accepted that beginning with the current semester, when student evaluations are returned to the instructor, a Grade Comparison Sheet would be included. This report provides the average grade in the course and the average grades over the past 5 years in the college and the instructor’s own division and department, both

overall and within the same course level, i. e., average grades in the 200-level courses if the own course is a 200 level.¹⁶ These Grade Comparison Sheets have been in place through the present.

3 Empirical Strategy

I seek to estimate how the provision of the grade information impacted grades differentially for courses that assigned below average and above average grades in the pre-information period. I allow for such a heterogeneous response since learning about average grades may serve as a signal of the grading norm, and instructors of courses on both sides of the average may adjust their grades to more closely align with this average. For estimation purposes, I must take a stand on the grade average with which to partition the grade observations into below and above average grade groups. In the main analysis I use the finest level at which I observe a grade, which is in a course (e. g., Economics 101). I then partition courses into two groups, depending on whether the average grade across all instances of the course (i. e., across instructors and multiple sections, if applicable) is above or below the average grade in the corresponding division-level pair (e. g., all 100-level social science courses) in the pre-information period.¹⁷

There exist other means by which I can define grades as belonging to below or above average grading groups. For example, I can compare course averages to department-level averages (instead of division-level averages) in the pre-information period. In spite of the concern that the observed effects of the information provision on grade assignment might be less apparent given that many department-level pairs consist of only one or two courses, the estimation results are similar to the main analysis. Alternatively I can conduct coarser analyses by comparing average grades in department-level pairs to the average of the corresponding division-level pair or simply average grades in departments to the college-wide average grade in the pre-information period. Each of these

16 The Grade Comparison Sheets also include the standard deviations of the grades in each category, as well as the percentage of grades higher than the average grade in the course for the comparison group.

17 Recall that the Grade Comparison Sheets provide to each faculty member the corresponding 5-year lagged division-level grade average for each course they teach. Thus, even in the last semester for which I have data, over half of the data used to compute this lagged average is from the pre-information period.

cases are considered in the Appendix (see Tables 9, 10, and 11, respectively), with findings that are qualitatively similar to the main analysis, but with smaller (in magnitude) and less precise effects for the coarser analyses.

I therefore test for differential effects on grades in courses with below and above average pre-information grades as follows:

$$G_{ict} = \beta_0 + \beta_1 \cdot \text{Below}_c + \beta_2 \cdot T_t \cdot \text{Below}_c + \beta_3 \cdot T_t \cdot \text{Above}_c + X_{ict}\Gamma + \mu_{ict} \quad [1]$$

where G is the numeric value of the course grade; i indexes the student, c the course, and t the semester. *Below* (*Above*) is a dummy variable equal to 1 if the grade is assigned in a course with average grade below (above) the corresponding division-level average grade in the pre-information period. T is a dummy variable equal to 1 for the treated (post-provision of grade information) observations and 0 for the untreated (pre-provision of grade information) observations. X is a vector of fixed and time-varying student demographic and course (department, division, level, units, and enrollment) characteristics. The main parameters of interest are β_2 and β_3 , which measure the average effects of information treatment for below and above average grading courses.

Note that since instructors had no access to average grades across campus or of other departments (and few even had access to average grades within their own department¹⁸) before the Grade Comparison Sheets, and once the Grade Comparison Sheets were implemented they were included with student evaluations, which are carefully read by instructors, compliance with information treatment is unlikely to be an issue in this setting. Thus, estimates of the “intention to treat” will be equivalent to the “treatment on the treated” effect in this setting. If, however, not all instructors viewed or understood the Grade Comparison Sheets, in spite of numerous announcements and explanations of their content, then the estimates of the intention to treat reported in this paper underestimate the treatment on the treated.

To obtain an informal graphical sense of eq. [1], Figure 4 plots average grades over time for two sets of courses: those with average grades below the average grade in their corresponding division-level pair in the pre-information period and those with above average grades. It looks as though after the provision of the Grade Comparison Sheets, as indicated by the vertical line, grades in the previously low grading courses began to increase, while those in

¹⁸ Department chairs had access to own department grades, but I find no evidence that this data was shared directly with the members of the department, except in the case of the economics department. Anecdotally, chairs have traditionally provided guidance to new faculty members by setting a “target” average grade for the department.

Figure 4: Change in grading patterns over time.

Notes: The long (short) dashed line plots the grade point average within all courses with average grades below (above) average grades in their corresponding division-level pair in the pre-information period across semesters. The dash-dot line at GPA of 3.31 is the college-wide average grade in the pre-information period and the vertical line falls immediately after Fall 2012, the last semester for which grades were assigned without the Grade Comparison Sheets.

the previously high grading courses declined. This would correspond to a positive sign for β_2 and a negative sign for β_3 . In the next section, I estimate these values more formally with appropriate controls.

The identifying assumption for estimating the causal impact of information provision on grades with model [1] is that there should be no differential changes in the way grades are determined between the pre- and post-information periods, other than the information treatment. By controlling for student characteristics such as high school standardized test scores, we can help mitigate a potential issue like a sudden change in the quality of students in a particular department at the same time that the Grade Comparison Sheets were provided to instructors. There is also the possibility of differential sorting across departments, whereby the treatment of information induces grading changes that may be amplified by the

migration of grade-sensitive students to different departments. Utilizing the rich micro-data on student and class characteristics as controls for each grade assigned, however, helps to mitigate such issues.

4 Empirical Results

The results for the estimation of model [1] are presented in columns 1–4 of Table 4. In all specifications, the dependent variable is the numerical value of the grade. Column 1 provides no controls, estimating the simple differential change in grades following the introduction of the Grade Comparison Sheets for courses that were below and above their corresponding division-level grade average in the pre-information period. These estimates correspond to the effects depicted graphically in Figure 4. The second column introduces controls for student demographics like test scores and dummy variables for ethnicity, first-generation college students, and legacy students. Additional course-level controls such as the number of enrolled students, units, and dummies for the course level and division are included as well. In the third column, department and student fixed effects are added. The student fixed effects absorb the demographic controls and the department fixed effects absorb the division dummy variables from column 2, controlling for all fixed observable and unobservable differences across students and departments. Finally, in column 4, semester fixed effects are added to all controls from column 3. The introduction of semester fixed effects necessarily absorbs one of the treatment interaction terms. Thus, I omit the interaction term between information treatment and the dummy variable indicating the grade was assigned in an above average course in the pre-period. This specification implicitly assumes that grades did not change in the above average courses, so that the estimated coefficient on the interacted treatment-below average group variable measures the impact of the information on grades in the below average courses relative to the above average courses.

Across the first three columns in Table 4, there is evidence that the Grade Comparison Sheets raised grades more in the courses that were below their division-level average than for those above. This difference in differences is at least 0.0796 across specifications and statistically significant at the 0.1% level. Estimating the separate effects of the grade information on grades in below and above average grading courses depends on the controls included. In the first two columns, I find strong evidence that below average courses saw their grades increase by at least 0.04 grade points, whereas for above average courses grades decreased by at least 0.05 grade points following the

Table 4: Impact of information on grades in full sample.

	(1)	(2)	(3)	(4)
Below	−0.431**** (0.0386)	−0.396**** (0.0204)	−0.321**** (0.0232)	−0.319**** (0.0228)
<i>T</i> · Below	0.0490*** (0.0169)	0.0388** (0.0173)	0.0780*** (0.0221)	0.0804**** (0.0185)
<i>T</i> · Above	−0.0484**** (0.0128)	−0.0515*** (0.0144)	−0.00155 (0.0179)	
Class size		−0.00276*** (0.000932)	−0.00269** (0.00106)	−0.00259** (0.00108)
Test score/100		0.0971**** (0.0102)		
First generation		−0.0485**** (0.0124)		
Legacy		−0.00206 (0.0143)		
<i>T</i> · Below − <i>T</i> · Above	0.0974**** (0.0185)	0.0903**** (0.0189)	0.0796**** (0.0188)	0.0804**** (0.0185)
Demographics	N	Y	N	N
Other controls	N	Y	Y	Y
Department fixed effects	N	N	Y	Y
Student fixed effects	N	N	Y	Y
Semester fixed effects	N	N	N	Y
Observations	69635	69017	69635	69635
Adjusted <i>R</i> ²	0.088	0.168	0.455	0.456

Notes: The dependent variable in all columns is the numerical grade score. *T* is a dummy variable equal to 1 in semesters after the introduction of the Grade Comparison Sheets, beginning Spring 2013. *Below* (*Above*) is a dummy variable indicating that the grade was assigned in a course with an average grade that was below (above) the average grade in the corresponding division-level pair in the pre-information period. In column 2, “Demographics” includes maximum test score (composite SAT Math and Verbal and SAT concordance with ACT score), dummy variables for all ethnicity and race categories, and dummy variables indicating legacy and first-generation students. The sample size is lower in column 2 because not all students have a record of taking either the SAT or ACT. The “Other Controls” in columns 2–4 include the number of students enrolled in the course, the number of units assigned to the course, and dummy variables for the level of the course and the division. The department fixed effects in columns 3–4 absorb the division dummy variables included in “Other Controls.” The semester fixed effects in column 4 absorb the *T* · *Above* control. The final row is computed by the difference in rows 1 and 2. Robust standard errors clustered by department are in parentheses.

p* < 0.10, *p* < 0.05, ****p* < 0.01, *****p* < 0.001.

information provision. These effects are consistent with the graphical evidence in Figure 4. When student demographics are included in column 2, we find the intuitive results that lower test scores and being a first-generation college student both reduce grades, while being a legacy student does not have a statistically significant impact on grades. Note also that all specifications that include class size as a control show it to have a statistically significant and negative impact on grades assigned in that class, with an additional 10 students lowering grade points by about 0.03, or just under 1% on average. Larger class sizes may allow instructors to more readily assign a low grade without the consequence of pulling down the course grade average significantly.

The estimated impact of the information provision on below and above average grading courses is sensitive to the inclusion of student and department fixed effects. With these additional controls in column 3, we now observe that below average graders increased grades by 0.078 grade points, while there is no statistically significant change in grading behavior for the above average graders.¹⁹ Note that with pre-information average grades of 3.12 (3.53) in the below (above) average groups, the increase of 0.0780 grade points in below average courses corresponds to a 2.5% increase in average grades for this group.²⁰ This result underscores the importance of controlling for unobservable compositional changes in the types of students across courses and within the college. Thus, although there is still strong evidence of a reduction in grade dispersion, the source of this reduction stems primarily from increases in grades for previously low grading courses, with little or no change for previously high grading courses. This effect of the policy is beneficial with respect to reducing the potential distortions induced by widely divergent grading practices that can (as discussed in the Introduction) impact students' class, effort, and major choices.

A potential concern with this analysis is that by construction, courses with above average grades in the pre-information period may not be able to increase grades as dramatically as those with below average grades, since they already issue higher grades and grades are bounded above. It may therefore not be

19 The results are robust to alternative sample restrictions, including the exclusion of small classes or observations from the Fall 2012 semester when the Grade Comparison Sheets were in process of being implemented.

20 The magnitude effects are smaller than the 5.2% reduction for average grades at Wellesley in departments impacted by the policy capping average grades at 3.33 in 100- and 200-level courses (Butcher et al. 2014). The relatively larger shift in grades at Wellesley generates changes in student enrollment and major decisions. I cannot precisely detect such changes at Occidental due to the relatively weaker information treatment.

surprising to observe below average grading courses increasing their grades more than above average grading courses. In fact, almost 13% of above average grading courses in the pre-information period assign A's to every student. However, there is still room for the vast majority of classes to increase their grades (the 50th (75th) percentile of GPAs in above average courses in the pre-information period is 3.64 (3.85)). Thus, the upper limit on GPAs is not a binding constraint for most courses and a priori it is entirely possible to observe increases or decreases in the average grading behavior of previously high grading courses. The fact that I find no evidence for a change in grading behavior for previously high graded courses with the introduction of the grade information is not a necessary implication of the data structure.

To verify the inference that the estimated effects are in fact due to the information treatment, I perform an additional robustness check. In particular, the model estimated in column 3 of Table 4 does not control for time-varying shifts in grading behavior distinct from the information treatment. For example, grades may be trending or systematically differ in fall versus spring semesters. To address this, I supplement the model with semester fixed effects in column 4. As noted previously, with the inclusion of semester fixed effects I cannot disentangle the level treatment effects for below and above average grading groups, but I can still estimate the differential response of below relative to above average groups. The introduction of semester fixed effects does not change the results from columns 1 to 3 that grades increased in below average grading courses by approximately 0.08 more than in above average grading courses. To better understand the potentially heterogeneous responses in grading behavior across the academic population, I next turn to a subsample analysis of the model, adopting the specifications in columns 3 and 4 of Table 4 as my preferred models throughout the remainder of the paper.

Table 5 presents the estimates from the main model on subsamples by course level and division. Recall that the average effect of the information provision for below average courses is 0.078 and for the above average courses is statistically insignificant at conventional levels. In Panel A, which matches the analysis of column 3 in Table 4, we similarly observe for the below average course group a statistically significant increase in average grade points at all course levels, with the effect increasing in the course level, from 0.0510 to 0.107 (although these are not statistically different). Across divisions, only the sciences do not show evidence of increasing grades in their courses with below average grades. Also consistent with the aggregate result, there is no evidence of a change in grading behavior for above average grading courses across any subsample. With the inclusion of semester fixed effects we observe across all subsamples a consistent differential effect in grading behavior

Table 5: Impact of information on grades by course subsamples.

	100 level (1)	200 level (2)	300 level (3)	Arts and humanities (4)	Sciences (5)	Social sciences (6)
Panel A:						
no semester FE						
Below	-0.328*** (0.0273)	-0.363*** (0.0285)	-0.348*** (0.0310)	-0.320*** (0.0339)	-0.363*** (0.0418)	-0.254*** (0.0230)
T · Below	0.0510* (0.0285)	0.100*** (0.0317)	0.107** (0.0426)	0.0922* (0.0444)	0.0473 (0.0442)	0.0645* (0.0307)
T · Above	-0.00791 (0.0232)	-0.0149 (0.0319)	0.0131 (0.0241)	-0.0289 (0.0257)	0.0178 (0.0259)	-0.0308 (0.0507)
Adjusted R ²	0.451	0.464	0.511	0.432	0.516	0.463
Panel B: semester FE						
Below	-0.324*** (0.0269)	-0.361*** (0.0287)	-0.347*** (0.0321)	-0.318*** (0.0331)	-0.360*** (0.0426)	-0.258*** (0.0217)
T · Below	0.0593** (0.0257)	0.122*** (0.0287)	0.0867* (0.0456)	0.119*** (0.0264)	0.0287 (0.0318)	0.0955* (0.0415)
Adjusted R ²	0.452	0.466	0.512	0.433	0.517	0.464
Observations	29,641	22,413	17,581	24,725	24,559	20,351

Notes: The dependent variable in all regressions is the numerical grade score. The column labels refer to the restricted subsample on which the model is estimated. T is a dummy variable equal to 1 in semesters after the introduction of the Grade Comparison Sheets, beginning Spring 2013. Below (Above) is a dummy variable indicating that the grade was assigned in a course with an average grade that was below (above) the average grade in the corresponding division-level pair in the pre-information period. All regressions include the following controls: the number of students enrolled in the course, the number of units assigned to the course, and department and student fixed effects. Level fixed effects are included in columns 4–6. In Panel B, the semester fixed effects absorb T · Above. Robust standard errors clustered by department are in parentheses. *p < 0.10, **p < 0.05, ***p < 0.01, ****p < 0.001.

between below and above average grading groups following the information provision, with the sciences again as the only statistically insignificant case. Thus, the aggregate effects of information provision on grading behavior are widespread throughout most segments of the college course offerings.²¹

The sciences stand out among the subsamples for not exhibiting the same statistically significant responsiveness to the information provision as measured in the aggregate. One possibility is that grades in the sciences did not in fact change in response to the information provision. With plausibly more quantitative means of student assessment relative to other disciplines, science instructors may possess a firmer sense of appropriate grading norms, independent of the grading behavior of their peers. Alternatively, grades in the sciences, the subsample that assigns the lowest grades in the college, may have increased uniformly across all courses following the information provision. In such a case, we would not expect to find differential grading behavior across previously high- and low-graded courses in the sciences, as is evidenced in panel B of Table 5. We would, however, expect to see statistically significant increases in panel A for both the below- and above-average grading groups. Although the effect for below average graded science courses is not statistically different from zero, it is also not statistically different from the positive effects in the other divisions. And the effect for the above average graded science courses, while also not statistically different from zero, is the only division-specific effect with a positive estimate. This discussion is only suggestive of a false negative and not conclusive about the grading behavior in the sciences. It is clear, however, that the sciences did not exhibit a dramatic change in grading behavior, which could have helped to align traditionally low science grades more with grading norms across the college.

I also question whether or not the grade information provision generated differential effects on grades across student demographic characteristics, namely race and ethnicity. In Table 6 I estimate my preferred specification with semester fixed effects, but supplement the model with a set of triple

²¹ The results are also robust to a subsample analysis by class size. Sorting grade observations into top, bottom, and interquartile groups by class size (more than 30 students, fewer than 16 students, and otherwise, respectively), I again find across all course size groups statistically significant increases in grades for the below average grading courses and no statistically significant changes for the above average grading courses. The inclusion of semester fixed effects shows increases in the differential change in below relative to above average grading groups (0.134 for small courses and 0.129 for large courses, both significant at the 1% level), although there is no statistically significant effect for medium-sized courses.

Table 6: Impact of information on grades by student race/ethnicity.

$r =$	1	Race Categories		
		Black or African-American	Hispanic or Latino	Asian
$r \cdot \text{Below}$	−0.290**** (0.0219)	−0.131**** (0.0351)	−0.110**** (0.0193)	−0.0250** (0.0105)
$r \cdot T \cdot \text{Below}$	0.0648*** (0.0214)	0.0942** (0.0409)	0.0440** (0.0190)	0.00450 (0.0250)

Observations: 69,635
Adjusted R^2 : 0.457

Notes: The dependent variable for the single model estimated in this table is the numerical grade score. T is a dummy variable equal to 1 in semesters after the introduction of the Grade Comparison Sheets, beginning Spring 2013. *Below* is a dummy variable indicating that the grade was assigned in a course with an average grade that was below the average grade in the corresponding division-level pair in the pre-information period. The main regressors of interest are $T \cdot \text{Below}$ and separate interactions of $T \cdot \text{Below}$ with each race/ethnicity dummy variable. There are four race/ethnicity dummy variables: for Black or African-American, Hispanic or Latino, Asian, and Other – White is omitted. Additional controls include *Below*, along with separate interactions of *Below* with each race/ethnicity dummy variable, the number of students enrolled in the course, the number of units assigned to the course, and student, department, level, and semester fixed effects. Robust standard errors clustered by department are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.01$.

interaction terms between each race/ethnicity group, T , and *Below*.²² The race/ethnicity groups are Black or African-American, Hispanic or Latino, Asian, and other, with White omitted to serve as the benchmark. The estimated coefficients on these triple interaction regressors measure the differential effect of the information provision on grades in previously low grading courses relative to high grading courses for non-White students relative to their White peers. I find this differential effect to be positive and statistically significant for Black and Hispanic students. The increase in grades in below relative to above average grading groups increased by 0.0942 and 0.0440 grade points more for Black and Hispanic students, respectively, than for White students at the college. These are large effects, considering that the total differential effect across all students was 0.0804. To obtain additional context for these magnitudes, it is instructive to note that in the pre-information period among low grading courses, the GPA for Black students was 2.80 and for Hispanic

²² Consistent with the previous analyses, I omit the $T \cdot \text{Above}$ interaction term, which is absorbed with the inclusion of semester fixed effects, as well as any interaction term between race/ethnicity and $T \cdot \text{Above}$.

students it was 2.92, relative to the higher GPAs of 3.12 for Asian students and 3.20 for White students. There is no similar statistically significant differential impact for Asian students relative to White students. This evidence suggests that Black and Hispanic students gained the most from this policy, which induced instructors in low grading courses to raise grades particularly for students from the racial groups that earn the lowest grades on average.

Finally, to better understand the impact of the grade information on the distribution of grades assigned, I estimate a series of linear probability models, where the outcomes are dummy variables equal to 1 if the grade is equal to a particular value and 0 otherwise. The results are summarized in Table 7, with the

Table 7: Impact of information on distribution of grades.

	Any A	Any B	Any C	Below C–
	(1)	(2)	(3)	(4)
Panel A: no semester FE				
Below	–0.236**** (0.0206)	0.150**** (0.0205)	0.0752**** (0.00796)	0.0113**** (0.00205)
<i>T</i> · Below	0.0633**** (0.0137)	–0.0448**** (0.00951)	–0.0192*** (0.00647)	0.000783 (0.00185)
<i>T</i> · Above	0.00370 (0.0137)	–0.0101 (0.0110)	0.00346 (0.00602)	0.00298 (0.00197)
Adjusted <i>R</i> ²	0.342	0.156	0.183	0.149
Panel B: semester FE				
Below	–0.233**** (0.0204)	0.146**** (0.0206)	0.0755**** (0.00780)	0.0116**** (0.00212)
<i>T</i> · Below	0.0584**** (0.0156)	–0.0322** (0.0137)	–0.0236**** (0.00439)	–0.00262 (0.00202)
Adjusted <i>R</i> ²	0.343	0.157	0.184	0.149
Observations	69,635	69,635	69,635	69,635

Notes: The dependent variable in all columns is a dummy variable equal to 1 if the grade is equal to the grade listed in the respective column heading. The model is estimated as a linear probability model. *T* is a dummy variable equal to 1 in semesters after the introduction of the Grade Comparison Sheets, beginning Spring 2013. *Below* (*Above*) is a dummy variable indicating that the grade was assigned in a course with an average grade that was below (above) the average grade in the corresponding division-level pair in the pre-information period. In all specifications, controls include the number of students enrolled in the course, the number of units assigned to the course, and department, level, and student fixed effects. In panel B the semester fixed effects absorb *T* · *Above*. Robust standard errors clustered by department are in parentheses.

p* < 0.10, *p* < 0.05, ****p* < 0.01, *****p* < 0.001.

cases of the dependent dummy variable indicating any A, B, or C, and any grade strictly less than a C-. There is a statistically significant increase in A's assigned and a corresponding decrease in B's and C's for courses below their division-level grade average, although as consistent with the previous analyses, there is no statistically significant effect for the above average courses. There is no statistically significant impact on very low grades below a C-. This finding is robust to the inclusion of semester fixed effects in panel B, where again we observe that the assignment of A's (B's and C's) increased (decreased) in below average grading courses relative to above average grading courses.²³ To obtain a better sense of the magnitude of these effects, 33 % of grades were some A in the pre-information period among the below average grading courses, so a 6.3 percentage point increase in some A for low grading courses corresponds to a 19 % increase in some A assigned. This evidence therefore suggests that the grade information served to further exacerbate the compression of grades at the top of the distribution, with little effect on the assignment of very low grades.

5 Conclusion

With a unique and comprehensive data set from a selective liberal arts college, I have shown the impact of a policy intervention that provides instructors with information about average grades in their own classes, department, division, and for the college overall. In particular, by changing the information set available to instructors, I find that average grades increased by 0.08 points more in courses that assigned average grades below their corresponding division-level pair in the pre-information period relative to those that assigned above average grades. Therefore, for institutions concerned with grade disparities, the provision of grade information to instructors may be effective in reducing the dispersion in grades across the college.

While the estimated differential effect of the information policy between high and low graders at Occidental College is robust, there are some limitations to this analysis. First, the result may not be externally valid, especially at other institutions of higher education that are dissimilar to a private liberal arts college. In addition, I have only estimated relatively short-run effects of the policy, ignoring potentially longer term dynamics. Finally, I do not investigate

²³ The results from instead estimating a logit model are presented in Table 8 and similarly show that the information treatment increased A's and decreased all lower grades in the below average courses relative to the above average courses.

whether the information policy was effective in reducing the overall level of grades in the absence of a compelling external control.

This analysis contributes both to a growing literature on the impact of information provision as a policy tool in general, and specific concerns in higher education related to combating the potentially distortionary impact of nonstandardized grading practices. As schools consider adopting policies to change instructors' incentives for grade assignments, this paper provides evidence for the effects of one type of such policy.

Acknowledgment: The author thanks Robby Moore, Mary Lopez, and two anonymous referees for helpful comments, and Jim Herr for collecting and providing access to the data.

Appendix

Table 8: Impact of information on distribution of grades: logit.

	Any A (1)	Any B (2)	Any C (3)	Below C– (4)
Below	–1.336**** (0.123)	0.705**** (0.103)	1.299**** (0.0972)	1.182**** (0.131)
<i>T</i> · Below	0.378**** (0.0914)	–0.165** (0.0658)	–0.485**** (0.114)	–0.467** (0.183)
Other controls	Y	Y	Y	Y
Department fixed effects	Y	Y	Y	Y
Student fixed effects	Y	Y	Y	Y
Semester fixed effects	Y	Y	Y	Y
Observations	64,538	66,170	38,328	10,226
Pseudo <i>R</i> ²	0.157	0.0555	0.159	0.138

Notes: The dependent variable in all columns is a dummy variable equal to 1 if the grade is equal to the grade listed in the respective column heading. The model is estimated as a logit and values in table are (untransformed) estimated coefficients. *T* is a dummy variable equal to 1 in semesters after the introduction of the Grade Comparison Sheets, beginning Spring 2013. *Below* is a dummy variable indicating that the grade was assigned in a course with an average grade that was below the average grade in the corresponding division-level pair in the pre-information period. The “Other Controls” include the number of students enrolled in the course, the number of units assigned to the course, and dummy variables for the level of the course. The number of observations varies across models because students with all grades equal or all grades not equal to the dependent variable are dropped in the estimation. Robust standard errors clustered by department are in parentheses. **p* < 0.10, ***p* < 0.05, ****p* < 0.01, *****p* < 0.001.

Table 9: Impact of information on grades with course GPAs relative to department-level pair GPAs.

	(1)	(2)	(3)	(4)
Below	−0.348**** (0.0298)	−0.304**** (0.0196)	−0.283**** (0.0208)	−0.282**** (0.0204)
$T \cdot \text{Below}$	0.0477*** (0.0152)	0.0400** (0.0172)	0.0799**** (0.0205)	0.0851**** (0.0150)
$T \cdot \text{Above}$	−0.0529**** (0.0138)	−0.0534**** (0.0141)	−0.00524 (0.0179)	
Class size		−0.00196 (0.000932)	−0.00231* (0.00106)	−0.00218 (0.00108)
Test score/100		0.0915**** (0.00927)		
First generation		−0.0469*** (0.0128)		
Legacy		−0.00184 (0.0139)		
$T \cdot \text{Below} - T \cdot \text{Above}$	0.1006**** (0.0168)	0.0934**** (0.0190)	0.0851**** (0.0154)	0.0851**** (0.0150)
Demographics	N	Y	N	N
Other controls	N	Y	Y	Y
Department fixed effects	N	N	Y	Y
Student fixed effects	N	N	Y	Y
Semester fixed effects	N	N	N	Y
Observations	69,635	69,017	69,635	69,635
Adjusted R^2	0.055	0.137	0.455	0.456

Notes: The dependent variable in all columns is the numerical grade score. T is a dummy variable equal to 1 in semesters after the introduction of the Grade Comparison Sheets, beginning Spring 2013. *Below* (*Above*) is a dummy variable indicating that the grade was assigned in a course with an average grade that was below (above) the average grade in the corresponding department-level pair in the pre-information period. In column 2, “Demographics” includes maximum test score (composite SAT Math and Verbal and SAT concordance with ACT score), dummy variables for all ethnicity and race categories, and dummy variables indicating legacy and first-generation students. The sample size is lower in column 2 because not all students have a record of taking either the SAT or ACT. The “Other Controls” in columns 2–4 include the number of students enrolled in the course, the number of units assigned to the course, and dummy variables for the level of the course and the division. The department fixed effects in columns 3–4 absorb the division dummy variables included in “Other Controls.” The semester fixed effects in column 4 absorb the $T \cdot \text{Above}$ control. The final row is computed by the difference in rows 1 and 2. Robust standard errors clustered by department are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Table 10: Impact of information on grades with department-level pair GPAs relative to division-level pair GPAs.

	(1)	(2)	(3)	(4)
Below	−0.282**** (0.0526)	−0.252**** (0.0287)	−0.0950**** (0.0150)	−0.0891**** (0.0152)
<i>T</i> · Below	0.0419* (0.0213)	0.0264 (0.0193)	0.0825*** (0.0244)	0.0398* (0.0224)
<i>T</i> · Above	−0.00796 (0.0104)	−0.0160 (0.0115)	0.0439** (0.0164)	
Class size		−0.00339** (0.00149)	−0.00349* (0.00170)	−0.00334* (0.00171)
Test score/100		0.0927**** (0.0100)		
First generation		−0.0483*** (0.0145)		
Legacy		−0.00673 (0.0145)		
<i>T</i> · Below − <i>T</i> · Above	0.0499* (0.0243)	0.0424* (0.0232)	0.0386* (0.0225)	0.0398* (0.0224)
Demographics	N	Y	N	N
Other controls	N	Y	Y	Y
Department fixed effects	N	N	Y	Y
Student fixed effects	N	N	Y	Y
Semester fixed effects	N	N	N	Y
Observations	73057	72383	73057	73057
Adjusted <i>R</i> ²	0.039	0.126	0.420	0.421

Notes: The dependent variable in all columns is the numerical grade score. *T* is a dummy variable equal to 1 in semesters after the introduction of the Grade Comparison Sheets, beginning Spring 2013. *Below* (*Above*) is a dummy variable indicating that the grade was assigned in a department-level pair with an average grade that was below (above) the average grade in the corresponding division-level pair in the pre-information period. In column 2 “Demographics” includes maximum test score (composite SAT Math and Verbal and SAT concordance with ACT score), dummy variables for all ethnicity and race categories, and dummy variables indicating legacy and first-generation students. The sample size is lower in column 2 because not all students have a record of taking either the SAT or ACT. The “Other Controls” in columns 2–4 include the number of students enrolled in the course, the number of units assigned to the course, and dummy variables for the level of the course and the division. The department fixed effects in columns 3–4 absorb the division dummy variables included in “Other Controls.” The semester fixed effects in column 4 absorb the *T*·*Above* control. The final row is computed by the difference in rows 1 and 2. There are an additional 3,422 observations in this analysis since it is unnecessary here to drop grades assigned in courses that do not exist in the pre-information period. Robust standard errors clustered by department are in parentheses.

p* < 0.10, *p* < 0.05, ****p* < 0.01, *****p* < 0.001.

Table 11: Impact of information on grades with department GPAs relative to college-wide GPA.

	(1)	(2)	(3)	(4)
Below	−0.283**** (0.0409)	−0.271**** (0.0369)		
<i>T</i> · Below	0.0409** (0.0185)	0.0204 (0.0153)	0.0750**** (0.0197)	0.0168 (0.0197)
<i>T</i> · Above	0.00955 (0.0150)	−0.00214 (0.0152)	0.0597*** (0.0207)	
Class size		−0.00407** (0.00171)	−0.00361** (0.00174)	−0.00343* (0.00174)
Test score/100		0.0939**** (0.0102)		
First generation		−0.0533**** (0.0134)		
Legacy		−0.00380 (0.0148)		
Demographics	N	Y	N	N
Other controls	N	Y	Y	Y
Department fixed effects	N	N	Y	Y
Student fixed effects	N	N	Y	Y
Semester fixed effects	N	N	N	Y
Observations	73,057	72,383	73,057	73,057
Adjusted <i>R</i> ²	0.041	0.120	0.419	0.421

Notes: The dependent variable in all columns is the numerical grade score. *T* is a dummy variable equal to 1 in semesters after the introduction of the Grade Comparison Sheets, beginning Spring 2013. *Below* (*Above*) is a dummy variable indicating that the grade was assigned in a department with an average grade that was below (above) the average grade in the college in the pre-information period. *Below* is absorbed by department fixed effects in columns 3 and 4. In column 2 “Demographics” includes maximum test score (composite SAT Math and Verbal and SAT concordance with ACT score), dummy variables for all ethnicity and race categories, and dummy variables indicating legacy and first-generation students. The sample size is lower in column 2 because not all students have a record of taking either the SAT or ACT. The “Other Controls” in columns 2–4 include the number of students enrolled in the course, the number of units assigned to the course, and dummy variables for the level of the course and the division. The department fixed effects in columns 3 and 4 absorb the division dummy variables included in “Other Controls.” The semester fixed effects in column 4 absorb the *T* · *Above* control. There are an additional 3,422 observations in this analysis since it is unnecessary here to drop grades assigned in courses that do not exist in the pre-information period. Robust standard errors clustered by department are in parentheses.

p* < 0.10, *p* < 0.05, ****p* < 0.01, *****p* < 0.001.

References

- Abaluck, J. 2011. "What Would We Eat if We Knew More: The Implications of a Large-Scale Change in Nutrition Labeling." Working Paper, MIT.
- Abaluck, J., and J. Gruber. 2011. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review* 101 (4):1180–210.
- Babcock, P.. 2010. "Real Costs of Nominal Grade Inflation? New Evidence from Student Course Evaluations." *Economic Inquiry* 48 (4):983–96.
- Bar, T., V. Kadiyali, and A. Zussman. 2009. "Grade Information and Grade Inflation: The Cornell Experiment." *Journal of Economic Perspectives* 23 (3):93–108.
- Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu. 2009. "The role of simplification and information in college decisions: Results from the H&R block FAFSA experiment." NBER working paper series w15361.
- Boleslavsky, R., and C. Cotton. 2015. "Grading Standards and Education Quality." *American Economic Journal: Microeconomics* 7 (2):248–79.
- Bollinger, B., P. Leslie, and A. Sorensen. 2011. "Calorie Posting in Chain Restaurants." *American Economic Journal: Economic Policy* 3 (1):91–128.
- Butcher, K. F., P. J. McEwan, and A. Weerapana. 2014. "The Effects of an Anti-Grade-Inflation Policy at Wellesley College." *Journal of Economic Perspectives* 28 (3):189–204.
- Chan, W., L. Hao, and W. Suen. 2007. "A Signaling Theory of Grade Inflation." *International Economic Review* 48 (3):1065–90.
- Chetty, R., A. Looney, and K. Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review* 99 (4):1145–77.
- Chetty, R., and E. Saez. 2013. "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients." *American Economic Journal: Applied Economics* 5 (1):1–31.
- Duflo, E., and E. Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *The Quarterly Journal of Economics* 118 (3):815–42.
- Dupas, P.. 2011. "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics* 3 (1):1–34.
- Hastings, J. S., and J. M. Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *The Quarterly Journal of Economics* 123 (4):1373–414.
- Hoxby, C., and S. Turner. 2013. "Expanding College Opportunities for High-Achieving, Low Income Students." SIEPR discussion paper no. 12-014.
- Jensen, R.. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics* 125 (2):515–48.
- Johnson, V. E.. 2003. *Grade Inflation: A Crisis in College Education*, 2003 ed. New York: Springer.
- Kling, J. R., S. Mullainathan, E. Shafir, L. C. Vermeulen, and M. V. Wrobel. 2012. "Comparison Friction: Experimental Evidence from Medicare Drug Plans." *The Quarterly Journal of Economics* 127 (1):199–235.
- Kuh, G. D., and S. Hu. 1999. "Unraveling the Complexity of the Increase in College Grades from the Mid-1980s to the Mid-1990s." *Educational Evaluation and Policy Analysis* 21 (3):297–320.

- Liebman, J. B., and E. F. P. Luttmer. 2015. "Would People Behave Differently if They Better Understood Social Security? Evidence from a Field Experiment." *American Economic Journal: Economic Policy* 7 (1):275–99.
- Millman, J., S. P. Slovacsek, E. Kulick, and K. J. Mitchell. 1983. "Does Grade Inflation Affect the Reliability of Grades?." *Research in Higher Education* 19 (4):423–9.
- Nguyen, T. 2008. "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar." MIT PhD thesis.
- Ostrovsky, M., and M. Schwarz. 2010. "Information Disclosure and Unraveling in Matching Markets." *American Economic Journal: Microeconomics* 2 (2):34–63.
- Ouazad, A., and L. Page. 2013. "Students' Perceptions of Teacher Biases: Experimental Economics in Schools." *Journal of Public Economics* 105:116–30.
- Rojstaczer, S., and C. Healy. 2010. "Grading in American Colleges and Universities." *Teachers College Record*, March 4.
- Rojstaczer, S., and C. Healy. 2012. "Where A Is Ordinary: The Evolution of American College and University Grading, 1940–2009." *Teachers College Record* 114 (7):1–23.
- Sabot, R., and J. Wakeman-Linn. 1991. "Grade Inflation and Course Choice." *Journal of Economic Perspectives* 5 (1):159–70.
- Summary, R., and W. L. Weber. 2011. "Grade Inflation or Productivity Growth? An Analysis of Changing Grade Distributions at a Regional University." *Journal of Productivity Analysis* 38 (1):95–107.
- van Ewijk, R.. 2011. "Same Work, Lower Grade? Student Ethnicity and Teachers' Subjective Assessments." *Economics of Education Review* 30 (5):1045–58.
- Wiswall, M., and B. Zafar. 2015. "Determinants of College Major Choice: Identification Using an Information Experiment." *The Review of Economic Studies* 82 (2):791–824.
- Zheng, Y., E. W. McLaughlin, and H. M. Kaiser. 2013. "Salience and Taxation: Salience Effect versus Information Effect." *Applied Economics Letters* 20 (4–6):508–10.