2010

Revitalizing the Signaling Power of Class Rank at Colby College

Nicholas Van Niel
Colby College

Revitalizing the Signaling Power of Class Rank at Colby College[1]

Nicholas Van Niel

Colby College

May 2010

Abstract

Consistent with trends at American colleges and universities nationwide, grades have been monotonically increasing at Colby College over the past decade while controlling for student aptitude. A rich data set that describes every Colby student over three cohorts is used to predict student performance. By comparing the mean predicted grade to the mean actual grade in a department, it is seen that some departments award mean grades that are significantly higher than predicted grades warranted by explainable factors. With some departments giving significantly higher grades than other departments, the current grading system is susceptible to awarding unwarranted higher grades to students who major in those departments. Therefore, a normalized system of class rank is proposed in which performance is standardized against the mean grade at that particular class level in the respective department. The alternative system of class rank appears to mitigate the detrimental effects that large discrepancies in mean grades between departments engender.

[1] I would like to thank Professor Philip Brown and Professor Michael Donihue for their support and recommendations. I would also like to thank Bill Wilson for providing data for such commentary.

I. Introduction

Average grades have increased at such a drastic pace over the past half century that nine in ten grades awarded at Colby College are in the A or B range. Such findings are not unique as evidenced by the literature, which demonstrates that grades have increased significantly at American universities and colleges since the 1960s.[2] For example, Juola (1976, 1980) find that the average GPA rose by 0.432 points, on average, between 1960 and 1974 in their sample of over 180 colleges (Table 1). Sabot and Wakeman-Linn (1991) find that the mean grade in introductory courses of eight large departments at Williams College rose from 2.49 in 1962-1963 to 2.93 in 1985-1986. In the other eight institutions in their sample, the mean grade in introductory courses of the same eight departments rose from 2.38 in 1962-1963 to 2.91 in 1985-1986. In a survey of 4,900 undergraduates from all types of institutions between 1969 and 1993, Levine and Cureton (1998) find that the percentage of A's awarded increased from 7 percent to 26 percent while the share of C's decreased from 23 percent to 9 percent. Kuh and Hu (1999) compare the GPAs of 52,256 students from 112 institutions to find that student grades rose from 3.07 in the mid-1980s to 3.34 in the mid-1990s. Gose (1997) finds that the mean GPA at Duke University rose from 2.7 in 1969 to 3.3 in 1996; at Lehigh University, it rose from 2.60 in 1972 to 2.90 in 1996; at Pacific Lutheran University it rose from 2.99 in 1974 to 3.20 in 1996; at University of California, Berkeley it rose from 2.95 in 1986 to 3.10 in 1996; and at the University of Washington, it rose from 2.31 in 1964 to 3.12 in 1996.[3][4] McSpirit and Jones (1999) also find a consistent climb—more than 0.10 grade points every five years—in the

[2] There is a large debate in the literature concerning whether the increases in grades can be labeled as 'grade inflation.' In order to remain agnostic, I avoid the term.
[3] Eckert (1988), Alexander (1993), Cole (1993), Shea (1994), Johnson (1997), and Healy (2001) also report conclusive evidence that support steady increases in the average GPA during the latter half of the 20th century.
[4] More recent evidence of increases in grades can be found at http://cs.furman.edu/~chealy/stats/. A summary of this data was published on April 18, 2010 in the New York Times: *Data; A as the New B.*

average graduating GPA between 1983 and 1996 while controlling for aptitude, institutional, and other demographic factors.

While the evidence presented above indicates the prevalence of rising grades across a wide range of educational institutions, Rosovsky and Hartley (2002) report that increases in grades are especially prevalent at elite, selective colleges and universities; 91 percent of Harvard University's Class of 2001 graduated with Latin honors. Indeed Lambert (1993) reports that the proportion of undergraduate grades that were an A- or higher Harvard University increased from 22 percent to 43 percent between 1966 and 1991. Valle (1993) finds that 88 percent of grades awarded at Stanford University were either in the A or B range between 1991 and 1993. Archibold (1998) reports that 83 percent of the grades awarded at Princeton University between 1992 and 1997 were between the A+ and B- range, compared to only 69 percent between 1973 and 1977. In addition, the median GPA at Princeton University rose from 3.08 for the Class of 1973 to 3.42 for the Class of 1997. These increases in average grades awarded over the past half century may be warranted if student ability has increased proportionally over the same period; however, Bowen and Bok (1998) and McSpirit and Jones (1999) find that when considered alongside multiple indices of student achievement, the increases in GPA do not appear to be justified.

Birnbaum (1977) and Winzer (2002) note that there are institutional pressures that help explain some of the increases in grades over the past half century. Rosovsky and Hartley (2002) point to the Vietnam War and the turmoil of the 1960s as the primary driver of increases in grades, for professors knew that awarding poor grades caused students to drop out of school and made the students subject to military service in Vietnam. In addition, there have been curriculum adjustments that have engendered increases in grades. During the turmoil of the 1960s, schools

provided their students with the option to drop classes late in the semester or to ex post change a letter grade to a pass/fail mark. Institutions also believed that awarding their students higher grades would improve retention and attract applicants.

Increases in grades have also led to significant differences in mean grades between departments. Sabot and Wakeman-Linn (1991) find that Art, Music, English, Philosophy, Political Science, and Psychology were high-grading departments at Williams College, with a mean grade in introductory courses of 3.03, while Chemistry, Economics, and Math were low-grading departments with a mean grade in introductory courses of 2.67 in 1985-1986. Consistent with the evidence provided by Sabot and Wakeman-Linn (1991), Quinones (2008) reports that grades are higher in the humanities at Princeton University: between 2001 and 2004, A's accounted for 55.6 percent of student grades in the humanities, 43.3 percent in the social sciences, and 37.2 percent in the natural sciences. Johnson (2003) also finds that grades at Duke University tend to be higher in the humanities than in the natural and social sciences.

These differences would be warranted if the quality of students was different across disciplines, but Sabot and Wakeman-Linn (1991) find that there is no significant difference in either SAT scores of students or the grades for the same student in other courses across different departments. Consistent with this analysis, Zirkel (1999) finds that there is no evidence that classes in the humanities generally attract superior students, which would explain their higher mean grades.

With student ability being similar across departments, Dickson (1984) argues that departments award higher grades in order to maintain high student enrollments and to avoid reductions in their size and in the resources available for their faculty members. Freeman (1999) believes that high grades in the humanities are a compensating differential for lower post-

graduation salaries that humanities majors may receive. Bar and Zussman (2009) and Achen and Courant (2009) argue that two conditions are necessary in order to award relatively low average grades: students' enrollment demand must be high and inelastic; and the cost for professors of assigning low grades must be relatively low, which is the case when there are objective methods of assessment. Bar and Zussman (2009) hypothesize that both of these requirements are more likely to hold in the natural sciences than in the humanities.

These departmental differences may be exacerbated by the pressures on faculty to award high grades. Dickson (1984), Wallace and Wallace (1998), and Eiszler (2002) find that professors who award higher grades receive higher student evaluations. Therefore since student evaluations heavily influence a professor's promotion, tenure decision, and merit-based pay increases, professors have an incentive to award higher grades. Any of these factors may have triggered higher grades, but the upward trend in grades naturally perpetuates itself. For example, younger faculty members who received higher grades while undergraduates may award higher mean grades because their frame of reference is based on a higher mean GPA as compared to older faculty. In addition, low-grading professors will often be forced to conform to the policies of their high-grading peers in order to maintain enrollments.

The recent trends in grades have marginalized the effectiveness of grades as defined by Rosovsky and Hartley (2002), who argue that the purpose of grades is to inform students about how well/poorly they understand the content of their courses; to inform students of their strengths, weaknesses, and areas of talent; and to provide information to internal audiences, such as faculty and administration, and external audiences, such as graduate schools and employers.

With the upper constraint of grades at an A+, McSpirit and Jones (1999) and Winzer (2002) find that lower-ability students have experienced the highest rate of grade increase.

Therefore, as grades continue to rise at a faster rate for lower-quality students, the difference between good students and excellent students becomes blurred. This creates problems for students in determining how well they understand the content of their courses. Felten (2004) argues that a grade conveys information by telling us how a student's performance compares to that of other students. Felten (2004) extends this analysis to find that a Princeton grade conveys about 11% less information in 1997-2002 than it did in 1973-1977.

The discrepancy in mean grades across departments also has detrimental effects on students' ability to comprehend their academic strengths and weaknesses. Both Sabot and Wakeman-Linn (1991) and Johnson (2003) show that grading disparities between departments bias students' course and major choices: Johnson (2003) estimates that if the disparities in grading practices between departments were eliminated, the average undergraduate student at Duke University would probably take fifty percent more natural science and mathematics electives than he or she did. Romer (2002), Johnson (2003), and Bar and Zussman (2009) go as far as to purport that the dearth of students majoring in the natural sciences is one factor adversely affecting economic growth in the United States.

Sabot and Wakeman-Linn (1991), Shea (1994), Becker (1997), Edwards (2000) and Johnson (2003) find that large discrepancies in the grading practices between departments leads to internal, institutional allocation problems. Since Edwards (2000) reports that students flock to high-grading departments and since higher-enrolled departments receive more funding, on average, low-grading departments are faced with inadequate enrollments and funding. For example, Bruno (2007) reported that once Cornell class averages were posted online, students searched for the classes where the median grade was an A, and those courses experienced large increases in enrollments. In addition to internal problems, unwarranted increases in grades also

lead to employers putting less weight on their job candidates' academic performance: the percentage of human resource officers who agreed that transcripts of college grades ought to be included with an applicant's resume fell from 37.5 percent to 20 percent between 1978 and 1995 (Rosovsky and Hartley 2002).

With the detrimental effects that increases in grades have engendered, it is important to determine whether or not these increases have been due to explainable factors. In this paper, we see that grades have increased at Colby College consistently over the past decade. Similar to the increases in GPAs observed nationwide, the median GPA at Colby College rose from 2.38 in 1968 to 3.16 in 1996, at a rate of 0.14 points every five years on average. More recently, the rate of increase has slowed to 0.12 points every five years, with the median GPA increasing from 3.16 in 1996 to 3.37 in 2005 (Figure 1).[5] These increases have led to a compression of grades in the A or B range. Currently, nine in ten grades awarded at Colby College are either in the A or B range; moreover, nearly half (44%) of all grades awarded at Colby College are in the A range (Table 2). With such a narrow distribution of grades, employers and graduate schools may find it difficult to distinguish between excellence and average performance. Recent increases in grades have also resulted in certain departments awarding grades significantly higher and lower than values warranted by explainable factors. The increases in grades and discrepancies in grades across departments have caused grades at Colby College to lose much of their signaling power. In order to combat this problem, an alternative measure of class rank that measures students against their peers is proposed. By maintaining the current GPA-based grading system but altering class rank to be a relative measure of performance, grades are able to provide a clearer and more accurate signal to internal and external audiences, as well as to students.

[5] The recent decrease in the rate of grade increases may be due to the upper compression of an A+.

In Section II, data and variables are defined, and the theory of the regression models underlying the analysis is explained. In Section III, general results are addressed. In Section IV, grades are predicted for each department, and an analysis of the discrepancies between department mean grades is provided. In Section V, an alternative, standardized system of class rank is proposed. Section VI concludes the paper.

II. Data and Variables

Data come from student transcripts and admission forms.[6] The data set tracks 1,224 students through three graduating classes over their four-year matriculation: the Class of 2002 includes 390 students; the Class of 2003 includes 426 students; and the Class of 2004 includes 408 students.[7][8][9] Departments in which fewer than 100 grades were recorded over the six year period are excluded. The data include the courses that each student at Colby College completed during his or her enrollment; classes taken off-campus are not included. All told, 34,558 grade records are included in the data set. Cumulatively over the three class years, the mean grade awarded is 3.28; and the median grade awarded is 3.30. The mean and median cumulative GPA upon graduation are 3.28 and 3.30, respectively. In the final data set, the Class of 2002 has a median GPA of 3.26; the Class of 2003 has a median GPA of 3.30; and the Class of 2004 has a median GPA of 3.35 (Table 3). Even over the course of three class years, monotonic increases in cumulative median GPA can be seen.

[6] Data are provided by Dr. William P. Wilson, Director of Institutional Research, Colby College, Waterville, ME.
[7] Students for whom SAT scores and/or Dean's ratings were not available were omitted from the analysis, as were classes taken in 2005-2006 and those taken on a credit/non-credit or pass/fail basis and those taken for 6 or 8 credits. In addition, observations from departments which had fewer than 100 observations were omitted from the analysis. Lastly, non-U.S. citizens who did not report high school type are omitted.
[8] Observations that did not report school type and are international citizens are labeled as private-school observations.
[9] I fail to reject the hypothesis that the mean grade after such observations are omitted is different than the original mean grade at the 99% confidence level.

In order to obtain a more nuanced analysis of the increases of grades at Colby College, it is important to determine if the increases are due to explainable factors of student aptitude. A parsimonious OLS regression model is used that includes only exogenous factors of student performance:

$$GRADE = \beta_0 + \boldsymbol{REGION}\,\beta_1 + \beta_2\,International + \beta_3\,Female + \beta_4\,White + \beta_5\,Private + \beta_6\,SAT + \beta_7\,Early\ Decision + \beta_8\,Dean\ Rating + \beta_9\,Fall + \boldsymbol{YEAR}\,\beta_{10} + u$$

where the dependent variable *GRADE* is the grade awarded to a student in a particular course. Grades are converted to their respective GPA equivalent.[10] Region dummies are included because Colby considers geographic residence when considering an applicant's candidacy.[11] 8% of students are from the south, 8% are from the midwest, 9% are from the west, and 71% are from the northeast (Table 4).[12] All priors are seen in Table 5. Bar and Zussman (2009) do not find that foreign students to perform any differently when compared to U.S. citizens, but a measure of citizenship is included to confirm these findings: 7% of the sample is comprised of non-U.S. citizens. Adelman (1995) finds that gender is an important determinant of grades, for females generally earn higher grades than males; 54% of the sample is female. Vars et al. (1998) find that white students earn higher grades than their minority peers. 88% of students in the sample are white. Sabot and Wakeman-Linn (1991) find that high school type is a good proxy for high school quality, with private high schools producing higher-achieving students than public high schools. 35% of the sample attended private high schools. Sabot and Wakeman-Linn (1991) report that SAT is a good proxy for cognitive ability and find that SAT scores are positively correlated with grades. The mean SAT score in the sample is 1300. In order to

---

[10] A+: 4.3; A: 4.0; A-: 3.7; B+: 3.3; B: 3.0; B-: 2.7; C+: 2.3; C: 2.0; C-: 1.7; D+: 1.7; D: 1.0; D-: 0.7; F: 0.

[11] http://www.colby.edu/administration_cs/ir/upload/cds2006.pdf

[12] Colby strives to represent a large proportion of states and regions in order to increase the diversity of students. Colby College is most likely to accept students from non-northeastern regions if the students are of equal ability, ceteris paribus.

increase the sample size, students who submitted ACT scores but not SAT scores are provided with SAT-equivalent scores.[13] Jensen (2009) finds that first term and first year GPAs for students admitted under Early Decision I are about 0.5 points below those admitted under regular decision, but over the longer term, these differences become statistically insignificant. Early decision is included in order to clarify these results. 38% of Colby students are admitted under early decision. Dean rating is a rating that each student receives by the Office of Admissions in order to predict academic success at Colby based on pre-enrollment characteristics; scores range from 1, most likely to succeed academically, to 5, least likely to succeed academically. The mean dean rating assigned to incoming students is 3.08. Bar and Zussman (2009) find evidence of a "senioritis effect", indicating seniors receive lower grades than freshman. If students do shirk, then this effect might extend to semester as well. Under this hypothesis, grades received in the fall would be higher than in the spring. 51% of grades are received during the fall semester. Consistent with the upward trend in grades over time, Bars and Zussman (2009) find a strong positive time trend in grades. 4% of grades are awarded in 1998, 14% of grades are awarded in 1999, 22% of grades are awarded in 2000, 24% of grades are awarded in 2001, 20% of grades are awarded in 2002, 12% of grades are awarded in 2003, and 4% of grades are awarded in 2004.

The parsimonious model (1) of Table 6 fails to account for course characteristics and does not include any variables that are determined once enrolled at Colby College, including potentially endogenous factors that influence grades awarded such as major. Ex ante, students who double major may have more motivation/ambition due to the hypothesis that they are able to take on a greater work load, on average; 24% of students double major at Colby College. Bar and Zussman (2009) find that course credits are negatively associated with course grades, and since

[13] When students reported only ACT scores, SAT equivalents were substituted from http://www.act.org/aap/concordance/index.html.

independent studies average 2.2 credits at Colby and classes average 3.9 credits, a variable for class type is included. 5% of all grades received are in independent studies. Sabot and Wakeman-Linn (1991) indicate that students should earn higher grades in classes in their major, leading to a positive relation between taking courses in his or her major and grade received. 41% of all grades received are in classes in a student's major. Bar and Zussman (2009) find that compared to 100-level courses, grades are lower in 200-level courses but higher in 300- and 400-level courses. In our data, 46% of all grades received are in 100-level courses; 29% are in 200-level courses; 17% are in 300-level courses; and 8% are in 400-level courses. Consistent with Sabot and Wakeman-Linn (1991), Johnson (1997) observes grades to be lower in the social sciences and natural sciences compared to in the humanities. In our data, 5% of grades awarded are in interdisciplinary areas; 37% are in the humanities; 23% are in the natural sciences; and 34% are in the social sciences.

III. General Results

The parsimonious OLS specification reported in equation (1) is significant at the 99% confidence level with an F-statistic of 100.32 and an R^2 of 0.122, i.e., 12.2% of the variation in grades can be explained by the explanatory variables in the parsimonious model. Robust standard errors clustered by department are reported in parentheses in Table 6. Grades awarded to students from the midwest are insignificantly different than those received by students from the northeast. Students from the south receive grades 0.04 points lower than their northeastern peers, and this effect is significant at the 95% confidence level. Students from the west are awarded grades 0.03 points lower than students from the northeast, on average, and this effect is significant at the 90% confidence level. International students receive grades 0.09 points higher

than their U.S-citizen classmates; this effect is significant at the 99% confidence level. Females are awarded grades 0.118 points higher than males, on average, and this effect is significant at the 99% confidence level. White students, on average, receive grades 0.133 points higher than minorities; this effect is significant at the 99% confidence level. Students who attended private high schools receive grades 0.02 points lower than their public school peers; this effect is significant at the 90% confidence level. A 100-point increase in SAT scores is associated with a 0.046 point increase in grade received, on average; this effect is significant at the 99% confidence level. Early decision applicants receive grades that are not significantly different from other students. A 1-point increase in dean rating is associated with a 0.155 point decrease in grade received, on average; this effect is significant at the 99% confidence level. Grades received in the fall semester are insignificantly different than those received in the spring semester, on average. Last, grades awarded in 1999 are insignificantly different from those awarded in 1998. Compared with 1998, grades awarded in the year 2000 are 0.096 points higher, and this is significant at the 90% confidence level. Grades awarded in the year 2001 are 0.20 points higher than in 1998; grades awarded in the year 2002 are 0.28 points higher than in 1998; grades awarded in the year 2003 are 0.35 points higher than in 1998; and grades awarded in the year 2004 are 0.41 points higher than grades awarded in 1998. The differences for years 2001, 2002, 2003, and 2004 are all significant at the 99% confidence level. When controlling for student aptitude, the large magnitude and robustness of the differences in grades between 1998 and 2004 indicate that these increases in grades are not warranted by explainable factors.

However, the parsimonious model fails to include variables that are determined once enrolled at Colby College, and the literature suggests that these variables are significant in determining an accurate model that predicts grades. When course characteristics are included as

predicting factors of grades (model 2), the signs of the variables do not change, and the statistical significance of some coefficients change only slightly; the results remain extremely similar to the previous specification. Equation (2) is significant at the 99% confidence level with an F-statistic of 440.62, and the R^2 of the model is 0.178, i.e., 17.8% of the variation in grades can be explained by the reported student and course characteristics. Grades awarded to students who double major are 0.09 points higher than to students with a single major; this effect is significant at the 99% confidence level. Grades awarded in independent studies are 0.41 points higher than grades awarded in standard classes; this effect is significant at the 99% confidence level. Students perform 0.07 points better when taking classes in their major, significant at the 99% confidence level. Grades received in the humanities are not significantly different than grades received in interdisciplinary areas. However, grades received in the natural sciences and social sciences are 0.29 points and 0.24 points lower than grades received in interdisciplinary areas, respectively; this effect is significant at the 99% confidence level. Grades awarded in courses at the 200 level are 0.06 points higher than grades awarded at the 100 level; this effect is significant at the 90% confidence level. Grades awarded in 300-level courses and 400-level courses are 0.09 and 0.14 points higher than grades awarded in 100-level courses, respectively; both of these effects are significant at the 99% confidence level.

While the fully specified OLS model improves upon the parsimonious model by controlling for factors of course characteristics, it also contains predicted values above and below the GPA-bound constraints of 4.3 and 0. Therefore, a tobit model is specified in order to provide further robustness to the results (model 3). Equation (3) is significant at the 99% confidence level with an F-statistic of 406.54. All coefficients remain similar to the fully specified OLS model (2), except year 2002 for which the effect is now significant at the 99% confidence level.

Based on these results, it is evident that certain course characteristics lead to higher grades for students. In an effort to further reduce unobserved heterogeneity, student fixed-effects are added to the OLS model (model 4). Equation (4) is significant at the 99% confidence level with an F-statistic of 116.34, and the R^2 of the model is 0.105. In the student fixed-effects model, a student earns the highest predicted grades in the fall semester, enrolled in an independent study, taking a class in his or her major, taking a class in an interdisciplinary area, and taking a 400 level course. All of these effects are significantly different from zero at the 99% confidence level. Another important conclusion from the fixed effects model is that the coefficients for the year variables continue to increase monotonically.[14] Since course level is controlled for, the year variable captures the increase in grades over time due to unexplainable factors. The more recently a student has taken a class, the higher the grade that he or she receives. For example, a student who receives a grade 0.41 grade points in 2004 than he or she would have received in 1998.

Although the student fixed-effects model helps reduce some of the unobserved heterogeneity, a department fixed-effects model is also useful in order to control for the differences in grading practices across departments. Equation (5) provides such an analysis and is significant at the 99% confidence level with an F-statistic of 244.94, and the R^2 of the model is 0.167. Even when controlling for specific departments, results remain extremely similar to previous specifications, and there is still a positive trend in grades awarded over time; on average, a given department awards grades 0.23 points higher in 2004 than in 1998. It is clear that grades have been rising monotonically on an all-college level, but it is important to determine whether these increases have resulted in some departments awarding mean grades that are significantly

[14] Replacing year of course taken with graduation year provides entirely consistent results.

different than explainable, warranted values. In order to provide such an analysis, we first examine the differences in the grading practices between disciplines.

IV. Differences in Department Grades

Table 7 indicates that the disciplines with lower mean grades have higher standard deviations. As mean grades in disciplines increase, standard deviations consistently decrease. This is evidence of grade compression, possibly indicating that departments in the humanities and interdisciplinary areas have a harder time differentiating between the excellent and mediocre students. Although it is clear that some disciplines award higher grades than others, it is unclear whether these differences are warranted by student ability. In addition, it is uncertain whether these differences occur across all departments within the discipline. In order to address both of those issues, the previously-specified tobit model is used to predict grades based on the student and course characteristics included in regression model 3 of Table 6. Using the tobit model to predict grades that should be awarded based on explainable factors, a one-tailed t-test on the equality of the mean department grade awarded and the mean predicted grade for each department is performed. For departments in which the null can be rejected, and the actual mean grade awarded is significantly greater than the predicted mean grade in that department, the department is labeled as high-grading (represented by positive t-statistics in Table 8). Conversely, departments in which the null fails to be rejected, and the actual mean grade is significantly lower than the predicted mean grade, are labeled as low-grading (represented by negative t-statistics in Table 8).

In Table 8, the mean and standard deviation of both the predicted and actual grades are reported.[15] Also consistent with grade compression, the departments that award the highest mean grades have the lowest standard deviations. Departments 24, 27, 30, and 28 have four out of the five highest mean grades for all departments, and these four departments have the lowest standard deviations. Grades are not as effective a signal in these departments.

Similar to the large differences in mean grades between disciplines are the large differences in mean grades between departments. Departments 11 and 20 award mean grades of 3.03 and 3.06, while departments 5 and 30 award mean grades of 3.64 and 3.69, respectively. Departments 3, 11, and 21 stand out as having the largest negative gap (lower than predicted values) between actual and predicted grades with differences of -0.17, -0.25, and -0.25, respectively. Departments 2, 3, 6, 7, 8, 11, 13, 18, 20, 21, 22, and 23 emerge as low-grading departments, awarding mean grades significantly lower than predicted mean grades at the 99% confidence level. Departments 24, 26, and 30 stand out as being the departments with the largest positive (higher than predicted values) difference between the actual grade and the predicted grade with gaps of 0.34, 0.20, and 0.28, respectively. Departments 5, 9, 10, 15, 17, 19, 24, 26, 27, 29, and 30 emerge as high-grading departments, awarding mean grades that are significantly higher than the predicted mean grades at the 99% confidence level.

To think about the implications of the differences between mean department grades, assume two students take one third of their classes in his or her respective major. Also, let student A, who majors in department 11, and student B, who majors in department 30, receive average grades (B+) in classes that are not in his or her major. By performing on average within their major, student B will have a final GPA of 3.43 and student A will have a final GPA of 3.21,

[15] Departments have been rendered anonymous per request of the Head of Institutional Research in order to respect the confidentiality of such data.

a 0.22 point difference. If these two students were in the Class of 2003, this is a difference between being in the top 35 percent of the class versus being in the top 60 percent of the class. Under the current measure of class rank, there is no way of correcting this inequality.

Large discrepancies in mean department grades not only skew class ranks, but they also provide inaccurate information to students when deciding in which departments they comparatively excel. In addition, departments that award higher-than-predicted mean grades may experience enrollment increases due to their high mean grades. In as much as resources made available to departments reflect enrollments, enrollment fluctuations may force a skewed allocation of resources as students shy away from low-grading departments. If students and internal/external audiences are able to consider more than just absolute GPA as a measurement of achievement, then the merit of grades as a signaling device may increase. That is, if a relative measure of performance is incorporated, then class rank may be able to serve as more than just a basis for honors distinctions.

V. An Alternative Grading System

Given that some departments award significantly higher/lower-than-predicted grades, a class ranking system that measures relative performance is worth testing to determine whether measuring relative performance may help to mitigate the detrimental signaling effects of majoring in low-grading departments.

One strategy to improve the signaling power of grades is evaluated by Bar, Kadiyali, and Zussman (2009), who find that administrators at Cornell, in an attempt to revitalize the signaling power of grades, decided to provide median course grades on the internet. However, this system

engendered students to ‘shop’ for lenient courses, courses with high median grades. Surprisingly, the rate of the increases in grades actually increased after such a policy was implemented.

Another more prominent and effective attempt in improving the signaling power of grades is the Achievement Index (AI), proposed by Johnson (1997). The basic premise of the AI is to measure a student’s academic performance using comparisons of how that student did relative to his or her classmates. Pedersen (1997) explains that under the AI, professors would grade students regularly, but the results would be run through complex algorithms to adjust for levels of class difficulty. The AI measures students against real-world classmates, not a 4.0 ideal. It yields the greatest rewards to students who perform well in classes in which there is a wide distribution of grades. If a student earns a B+ in an economics class where the average grade was a B-, her AI rises. If she earns a B in a physics class where the average grade was a B+, her AI falls. However, the AI was not implemented due to its complexity and difficulty of interpretation.

To maintain the spirit of Johnson’s AI index without the complexity, I propose a measure of class rank that measures performance against the mean grade in classes at the same class level in the same department. Specifically, a student’s grade is based on the *z*-score measuring deviations from the mean grade in the department for a given course level. This can be represented by the following equation:

$$\mathrm{z} \equiv \sum_{h=1}^{N}\sum_{i=1}^{30}\sum_{j=100}^{400}\frac{Grade_{hij}-\overline{Grade}_{ij}}{\sigma_{ij}}$$

where *h* is the student and *i* is the department and *j* is the class level of the course. For example, if the mean grade in a 100-level Economics course is a B+, then every student who earns a grade above a B+ will receive a positive *z*-score. Conversely, every student who receives a grade below

a B+ will receive a negative *z*-score.[16] A student only receives a positive (negative) *z*-score if she performs better (worse) than the mean grade in that department at that course level. The *z*-score is then weighted based on the amount of credit hours for each student in order to obtain a final, cumulative number. The weighting occurs by multiplying the *z*-score by the amount of credit hours of that course. All of the *z*-scores for all of the student's courses are summed together and divided by the total number of credits that the student completed. Intuitively, the final, cumulative *z*-score is the average amount of standard deviations a student performed above/below the mean grade.

After ranking the cumulative *z*-score values, the alternative class rank system can be compared to the existing, GPA-based class rank system. Figure 2 displays the comparisons of class rank between the current system and the alternative system. There is a strong linear correlation between the two class rank systems, especially for students ranked near the top and bottom of their class. Students ranked in the middle of the class appear to experience the largest fluctuations in rank. For the Classes of 2002, 2003, and 2004, over 90% of the variation in status quo class rank can be explained by the variation in the alternative measure of class rank. This strong correlation is also seen by the mobility matrix in Table 9. 84% of students who are in the top 10th percentile under the status quo class rank system remain in the top 10th percentile under the alternative class rank system. Similarly, 85% of students who were in the 100th percentile in the status quo class rank system remain in the 100th percentile in the alternative class rank system. Thus, the alternative system neither penalizes the top performers nor rewards the bottom-feeders. However, the middle percentiles experience relatively large fluctuations in class rank: only 38% of students who were in the 50th percentile in the status quo class rank system remain in the 50th percentile under the alternative class rank system.

[16] Naturally, every student who receives a B+ in a 100-level Economics course receives a *z*-score of 0.

In order to examine which departments experience these relatively large fluctuations in class rank, all three class years are aggregated to measure the differences between the class rank systems by major. Each observation represents one student. Table 10 shows that students who major in high-grading departments generally experience large decreases in class rank, while students who major in low-grading departments experience marginal increases. For example, students who major in high-grading departments 24 or 30, on average, worsen in class rank by 16 and 15 percentage points, respectively, while students who major in low-grading departments 11 or 21, on average, improve in class rank by 5 and 3 percentage points, respectively.

Returning to the example of Student A and Student B, Student B (who majors in high-grading department 30) will experience a 16 percentage point decrease in class rank. This would place Student B in the top 51 percent of his or her class. Conversely, Student A (who majors in low-grading department 11), will experience a 5 percentage point increase in class rank such that Student A would now be in the top 55 percent of his or her class. What was once a 25 percentage point difference in class rank under the status-quo system is now only a 4 percentage point difference. It is important to note, however, that there are exceptions to the rule. For example, students who major in department 28, which awards grades insignificantly different from its predicted grades, on average lose 11 percentage points. One possible explanation arises if these students perform poorly in classes outside of the major.

An alternative system of class rank that measures relative performance against the mean grade may have several implications on the grading system. First, since grades are measured against all courses within a department at a particular course level, the alternative system of class rank encourages faculty to discuss (and possibly unify) their grade distributions on an intra-departmental level. As professors begin comparing grade distributions within departments, they

will naturally learn more about the grading practices of their departments. Second, if there is more uniformity in the grade distributions within departments, instructors may become curious about how departmental grade distributions compare to other departments within the college. This may lead to departments altering grade distributions, for distributions are vital factors in student course choice, which subsequently determines enrollments and funding decisions. Third, under an alternative system of class rank, a student will have to excel both absolutely and relatively in order to receive honors. Whether a student majors in a high- or low-grading department will have less of an impact on honors distinctions. Fourth, professors may also begin to employ grade distributions with more variance, for they may want to reward the best students. Grade distributions may be modified under the alternative proposed system because instructors now penalize excellent-performing students by awarding high mean grades with small variances. Last, a relative measure of class rank provides additional information to employers and graduate schools. External audiences will be able to determine how a student performs relative to his or her peers. By having the status-quo grading system and the relative class rank measure, grades will be a better signal of academic performance for both students and employers/graduate schools.

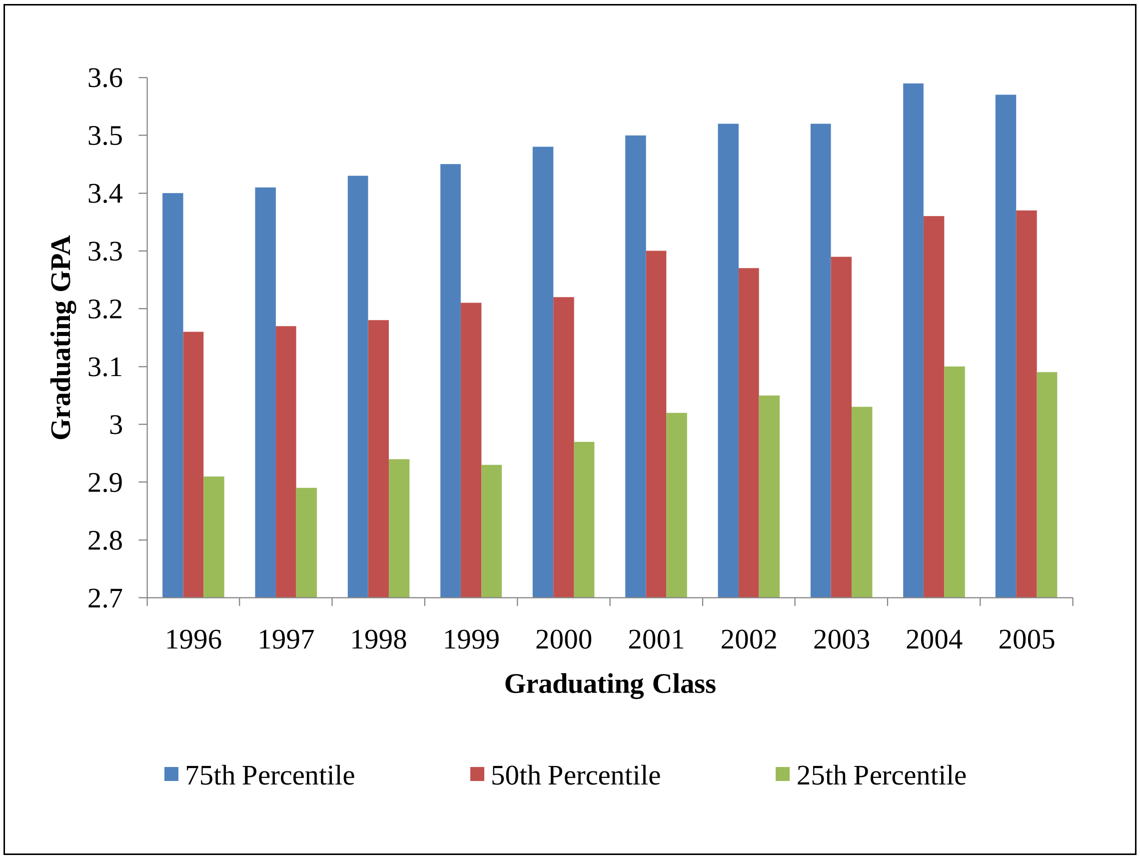
VI. Conclusion

Grades have been monotonically increasing at Colby College without proportionate increases in student ability. In addition, departments in humanities and interdisciplinary areas award grades that are significantly higher than those awarded in the social and natural sciences. The unexplainable increases in grades and the unwarranted discrepancies in grades across departments have diminished the signaling power of grades as grades are compressed, as

employers put less weight on academic performance, as students flock toward high-grading departments, and as students' course and major choices are biased by such disparities. Considering these detrimental effects, an alternative relative measure of class rank is proposed. This relative system of class rank improves the signaling power of grades by rewarding students who major in low-grading departments and by penalizing students who major in high-grading departments.

The implications of such a relative measure of class rank are significant. First, a relative measure of class rank provides different types of signaling power for both GPA and class rank. GPA becomes a signal of absolute performance, while class rank becomes a measure of relative performance. Future employers and graduate programs would then be able to attach importance to both measures. Second, a relative measure of class rank provides a relative measure of performance while still maintaining the status quo grading system. Pedersen (1997) notes that Duke undergraduates rejected the AI because it would make competition more cut-throat and leave the future of students at the mercy of an arcane formula no other college uses. Pedersen (1997) also notes that the AI failed because it was difficult to interpret and indicated an overhaul of the current system. By supplementing the current grading system with a different class rank system, the status quo is retained while additional information is provided. To ease initial implementation, it might be beneficial to provide students with the option of including the relative measure of class rank or the absolute, status quo measure of class rank on their transcripts. While the strictly relative measure of class rank may be able to be improved upon with more technically-rich methods, it could serve as a stepping stone for further evaluations of academic performance due to its ease of interpretation and ease of implementation.

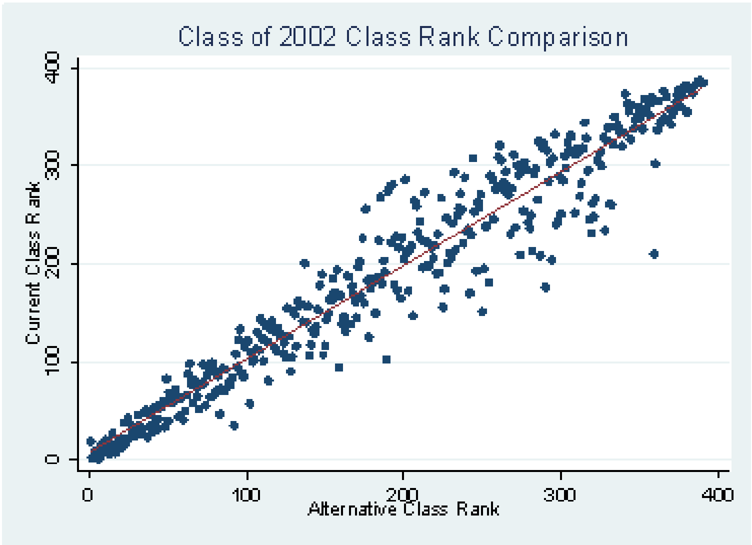
Figure 1: ***Colby College GPAs and Percentiles—Classes of 1996-2005***



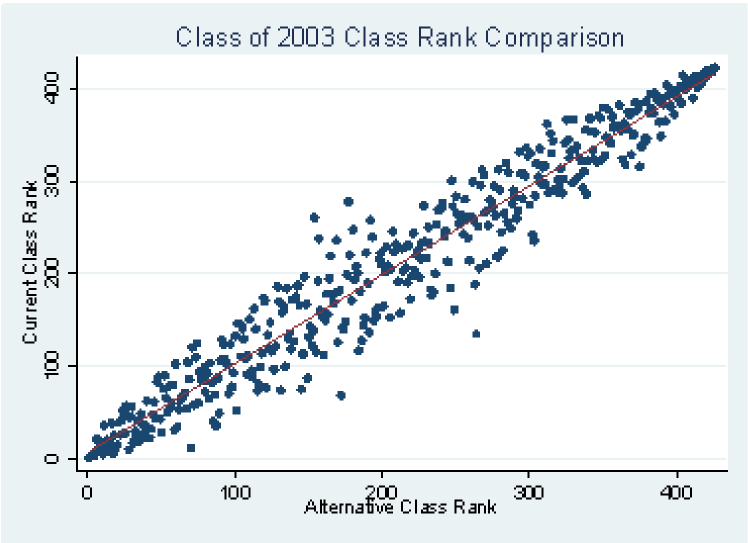
Note: The median (50th percentile) GPA in 1962 was 2.38 and in 1996 was 3.16.
The 25th percentile GPA in 1996 was 2.91 and in 2005 it was 3.09.
The 50th percentile GPA in 1996 was 3.16 and in 2005 it was 3.37.
The 75th percentile GPA in 1996 was 3.40 and in 2005 it was 3.57.

Source: Head of Institutional Research, Colby College

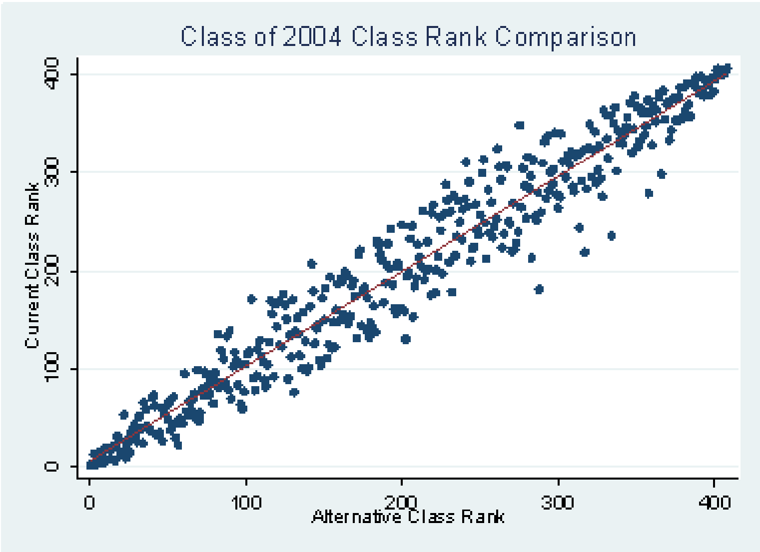
Figure 2: *Class Rank Comparisons*



R^2: 0.928



R^2: 0.942



R^2: 0.946

Table 1: ***Prior Studies***

Author(s)	Years Studied	Sample	Findings
Juola (1976)	1960-1974	180 colleges with graduate programs	Between 1960 and 1974, the average GPA rose by 0.432 points aggregated over all colleges.
Sabot and Wakeman-Linn (1991)	1962-1963; 1985-1986	Williams College; Amherst College, Duke University, Hamilton College, Haverford College, Pomona College, University of Michigan, University of North Carolina, and the University of Wisconsin	At Williams College, the mean grade in the introductory courses of eight large departments at Williams rose from 2.49 in 1962-1963 to 2.93 in 1985-1986. In the other eight institutions, the mean grade in the introductory courses of the same eight departments rose from 2.38 in 1962-1963 to 2.91 in 1985-86.
Levine and Cureton (1998)	1969; 1976; 1993	Data from survey of 4,900 undergraduates at all institutional types	At all institutional types, the number of A's increased nearly four times (from 7 percent to 26 percent) between 1969 and 1993 and the number of C's decreased by 66 percent (from 25 percent in 1969 to 9 percent in 1993).
Kuh and Hu (1999)	1984-1987; 1995-1997	52,256 student surveys from the Colleges Student Experiences Questionnaire (CSEQ) from 112 institutions	College grades increased over time in every institutional type on average from 3.07 to 3.34 between 1984-1987 and 1995-1997.
McSpirit and Jones (1999)	1983-1996	1,986 graduating seniors that entered the university as full time freshmen, between 1983 and 1992 at a large public University.	The average graduating GPA rose, on average, more than 0.1 grade points every five years between 1983 and 1996 while controlling for aptitude, institutional and other demographic factors.
Lambert (1993)	1966; 1991	Harvard University Undergraduates	At Harvard University, the proportion of undergraduate grades of A- or higher increased from 22 percent to 43 percent between 1966 and 1991.
Johnson (1997); Archibold (1998)	1968; 1994. 1988; 1998.	Dartmouth College Undergraduates	At Dartmouth the mean GPA rose to 3.23 in 1994, up from 3.06 in 1968. At Dartmouth in 1998, A's and A-'s comprised 44 percent of grades, up from 37 percent in 1988 and B-'s or below accounted for only 20 percent.

Table 1 (continued): ***Prior Studies***[17]

Author(s)	Years Studied	Sample	Findings
Archibold (1998)	1973-1977; 1992-1997	Princeton Dean's Report: A comprehensive review of undergraduates' grades over 24 years.	83 percent of the grades awarded at Princeton between 1992 and 1997 were between A+ and B-, compared with 69% between 1973 and 1977. The amount of C's fell from 15% to 9% over the same period. The median GPA at Princeton for the Class of 1997 was 3.42 compared with 3.08 for the Class of 1973.
Archen and Courant (2009)	1992-2008	University of Michigan's College of LSA, which has 16,000 students of its own covering 25 departments from Fall 1992 through Winter 2008	The median grade in lower-division courses at University of Michigan's College of LSA in 1992-1994 was 3.13 and in 2005-2007 were 3.30. The median grade in upper-division courses at University of Michigan's College of LSA was 3.24 in 1992-1994 and 3.40 in 2005-2007.[18]

[17] Stuart Rojstaczer has collected a fairly large database that reports trends in grades for over 70 American colleges and universities. Institutions included in the database are chosen strictly because they have either published their data or have sent their data to Stuart Rojstaczer. This data can be accessed at www.gradeinflation.com.

[18] Lower-division courses are mostly introductory. Higher-division courses are mostly taken by majors and others who have shown interest and/or aptitude for a field of study.

Table 2: ***Grades Awarded at Colby College***

Grade	Frequency	Percent	Cumulative Percent
A+	646	1.87%	1.87%
A	7,230	20.92%	22.79%
A-	7,115	20.59%	43.38%
A's	**14,991**	**43.38%**	**43.38%**
B+	6,529	18.89%	62.27%
B	6,256	18.10%	80.37%
B-	3,018	8.73%	89.10%
B's	**15,803**	**45.73%**	**89.10%**
C+	1,595	4.62%	93.72%
C	1,192	3.45%	97.17%
C-	536	1.55%	98.72%
C's	**3,323**	**9.62%**	**98.72%**
D+	145	0.42%	99.14%
D	168	0.49%	99.63%
D-	58	0.17%	99.80%
D's	**371**	**1.07%**	**99.80%**
F’s	**70**	**0.20%**	**100.00%**
Total	**34,558**	**100%**	**100%**

Grades shown were awarded to students of the Class of 2002, 2003, and 2004. Grades are from all departments on campus and all classes taken for a letter grade at Colby College. No transfer credits or credits taken abroad are included.
Source: Head of Institutional Research, Colby College

Table 3: ***Median Graduating GPAs***

	Class of 2002	Class of 2003	Class of 2004
Median Final GPA	3.26	3.30	3.35

Table 4: ***Summary Statistics***

	Mean	Std. Dev.	Minimum	Maximum	Observations
Grade Received	3.29	0.65	0	4.3	34,558
International	0.07	0.26	0	1	34,558
1 if non-USA citizen					
Female	0.54	0.50	0	1	34,558
1 if female, 0 if male					
White	0.88	0.32	0	1	34,558
1 if white, 0 if minority					
Private	0.35	0.48	0	1	34,558
1 if from Private High School 0 if from Public, Charter, or Other HS					
SAT	1300.6	100.9	820	1590	34,558
Includes ACT equivalents true score					
Early Decision	0.38	0.49	0	1	34,558
1 if accepted Early Decision					
Dean Rating	3.08	0.90	1	5	34,558
1 is most likely to succeed academically 5 is least likely					
Fall Semester	0.51	0.50	0	1	34,558
Double Major	0.24	0.43	0	1	34,558
1 if double majored, 0 if one major					
Independent Study	0.95	0.22	0	1	34,558
1 if Ind. Study, 0 if Class					
Major Match	0.41	0.49	0	1	34,558
1 if majored in dept. that class was taken					
Interdisciplinary	0.05	0.23	0	1	34,558
1 if class was taken in Interdisciplinary area					
Humanities	0.37	0.48	0	1	34,558
1 if class was taken in Humanities area					
Natural Sciences	0.23	0.42	0	1	34,558
1 if class was taken in Natural Sciences					
Social Sciences	0.34	0.47	0	1	34,558
1 if class was taken in Social Sciences					
Course Level 100	0.46	0.50	0	1	34,558
1 if course taken at 100 level					
Course Level 200	0.29	0.45	0	1	34,558
Course Level 300	0.17	0.38	0	1	34,558
Course Level 400	0.08	0.27	0	1	34,558

Note: Grade is measured as a GPA equivalent to letter grade. A+: 4.3; A: 4.0; A-: 3.7; B+: 3.3…; D-: 0.7; F: 0. Other controls include the year the class was taken (1998-2004), the course level (100-400), and region (Northeast, South, Midwest, and West).

Table 5: ***Priors***

Variable	Prior	Source
Midwest	-	
South	-	
West	-	
International	?	Bar and Zussman (2009)
Female	+	Adelman (1995)
White	+	Vars et al. (1998)
Private	+	Sabot and Wakeman-Linn (1991)
SAT	+	Sabot and Wakeman-Linn (1991)
Early Decision	-	Jensen (2009)
Dean Rating	-	
Fall	+	Bar and Zussman (2009)
More Recent Years	+	Bar and Zussman (2009)
Double Major	+	
Independent Study	+	Bar and Zussman (2009)
Major Match	+	Sabot and Wakeman-Linn (1991)
Humanities	?	
Natural Sciences	-	Johnson (1997)
Social Sciences	-	Johnson (1997)
Higher Course Levels	+	Bar and Zussman (2009)

Table 6: ***Regression Results***

	OLS Regressions		Tobit	Fixed Effects Student	Fixed Effects Department
	(1)	(2)	(3)	(4)	(5)
Midwest	0.0280	0.0319*	0.0308*		0.0317*
	(0.0209)	(0.0169)	(0.0172)		(0.0164)
South	-0.0423**	-0.0486***	-0.0494***		-0.0589***
	(0.0177)	(0.0167)	(0.0170)		(0.0155)
West	-0.0311*	-0.0236*	-0.0254*		-0.0201
	(0.0162)	(0.0133)	(0.0137)		(0.0132)
International	0.0913***	0.107***	0.111***		0.128***
	(0.0272)	(0.0222)	(0.0233)		(0.0191)
Female	0.118***	0.103***	0.103***		0.0903***
	(0.0187)	(0.0178)	(0.0181)		(0.0179)
White	0.133***	0.135***	0.135***		0.133***
	(0.0248)	(0.0226)	(0.0228)		(0.0224)
Private	-0.0212*	-0.0288***	-0.0304***		-0.0235***
	(0.0117)	(0.00771)	(0.00809)		(0.00761)
SAT	0.0457***	0.0398***	0.0400***		0.0420***
	(0.00903)	(0.00669)	(0.00678)		(0.00611)
Early Decision	-0.00855	-0.00603	-0.00668		-0.00628
	(0.0122)	(0.0107)	(0.0111)		(0.0108)
Dean Rating	-0.155***	-0.159***	-0.163***		-0.168***
	(0.0122)	(0.00936)	(0.00970)		(0.0101)
Fall	0.0148	0.0116	0.0111	0.0323***	0.0103
	(0.0133)	(0.0140)	(0.0145)	(0.00655)	(0.0123)
Year 1999	0.0448	0.0180	0.0178	0.0502***	0.0301
	(0.0287)	(0.0269)	(0.0273)	(0.0185)	(0.0267)
Year 2000	0.0958**	0.0421	0.0429	0.112***	0.0567
	(0.0414)	(0.0394)	(0.0398)	(0.0205)	(0.0382)
Year 2001	0.201***	0.110**	0.112***	0.214***	0.128***
	(0.0467)	(0.0415)	(0.0416)	(0.0218)	(0.0392)
Year 2002	0.282***	0.150***	0.153***	0.271***	0.164***
	(0.0582)	(0.0508)	(0.0509)	(0.0237)	(0.0494)
Year 2003	0.353***	0.196***	0.199***	0.353***	0.207***
	(0.0531)	(0.0483)	(0.0488)	(0.0257)	(0.0497)
Year 2004	0.405***	0.223***	0.231***	0.400***	0.230***
	(0.0518)	(0.0452)	(0.0462)	(0.0296)	(0.0448)
Double Major		0.0895***	0.0914***		0.0810***
		(0.0183)	(0.0189)		(0.0178)
Independent Study		0.410***	0.423***	0.333***	0.358***
		(0.0857)	(0.0889)	(0.0252)	(0.0554)
Major Match		0.0662**	0.0666**	0.0665***	0.117***
		(0.0262)	(0.0265)	(0.00916)	(0.0255)

Table 6 (continued): ***Regression Results***

Humanities		-0.0940	-0.0914	-0.121***	
		(0.0630)	(0.0639)	(0.0158)	
Natural Sciences		-0.290***	-0.285***	-0.349***	
		(0.0771)	(0.0792)	(0.0191)	
Social Sciences		-0.240***	-0.239***	-0.263***	
		(0.0795)	(0.0802)	(0.0156)	
Course Level 200		0.0673*	0.0664*	0.0311***	0.0344
		(0.0360)	(0.0367)	(0.00977)	(0.0377)
Course Level 300		0.0943***	0.0921***	0.0345***	0.0624*
		(0.0301)	(0.0308)	(0.0121)	(0.0324)
Course Level 400		0.144***	0.143***	0.0646***	0.127***
		(0.0509)	(0.0525)	(0.0178)	(0.0418)
Constant	2.787***	3.443***	3.464***	3.569***	2.848***
	(0.124)	(0.132)	(0.136)	(0.0339)	(0.102)
Observations	34,558	34,558	34,558	34,558	34,558
R-squared	0.122	0.178		0.105	0.167
Number of students				1,224	
Number of departments					30

For models 1, 2, and 3, robust standard errors clustered by department are given in parentheses.
For models 4 and 5, robust standard errors are given in parentheses.
SAT is measured in 100 point increments.
*** p<0.01, ** p<0.05, * p<0.1
Course-level dummies are measured against 100-level courses. Year dummies are measured against 1998. Discipline dummies are measured against interdisciplinary subjects. Region dummies are measured against the northeast.

Table 7: ***Grades Awarded Within Disciplines***

	Mean	Std. Dev.	N
Natural Science Grades	3.195	0.739	7,998
Social Science Grades	3.235	0.652	11,815
Humanities Grades	3.383	0.582	12,891
Interdisciplinary Grades	3.489	0.530	1,854

Table 8: ***Results for Differences in Mean Department Grades***

	N	Predicted Grade Mean	Predicted Grade Std. Dev.	Actual Grade Awarded Mean	Actual Grade Awarded Std. Dev.	Actual - Predicted Mean	t-statistic for equality of means
Dept. 1	683	3.186	0.219	3.239	0.703	0.053	2.03**
Dept. 2	2,334	3.259	0.255	3.186	0.691	-0.073	-5.87***
Dept. 3	778	3.329	0.216	3.155	0.649	-0.174	-8.38***
Dept. 4	1,514	3.295	0.248	3.274	0.632	-0.021	-1.39*
Dept. 5	1,620	3.534	0.298	3.636	0.590	0.102	8.06***
Dept. 6	1,582	3.296	0.259	3.227	0.756	-0.069	-4.18***
Dept. 7	1,213	3.338	0.233	3.254	0.596	-0.084	-5.26***
Dept. 8	554	3.262	0.277	3.147	0.775	-0.115	-3.76***
Dept. 9	971	3.214	0.264	3.405	0.626	0.191	10.19***
Dept. 10	496	3.335	0.266	3.450	0.525	0.115	5.12***
Dept. 11	2,344	3.276	0.283	3.030	0.728	-0.246	-18.26***
Dept. 12	565	3.470	0.270	3.508	0.500	0.038	1.92**
Dept. 13	1,160	3.353	0.259	3.207	0.667	-0.146	-8.32***
Dept. 14	1,007	3.181	0.243	3.215	0.669	0.034	1.77**
Dept. 15	1,200	3.219	0.221	3.396	0.478	0.177	14.32***
Dept. 16	599	3.091	0.250	3.113	0.713	0.022	0.78
Dept. 17	3,613	3.408	0.254	3.444	0.478	0.036	4.86***
Dept. 18	475	3.413	0.269	3.317	0.639	-0.096	-3.72***
Dept. 19	2,342	3.212	0.249	3.294	0.604	0.082	7.05***
Dept. 20	1,258	3.215	0.285	3.060	0.842	-0.155	-7.05***
Dept. 21	312	3.580	0.266	3.334	0.608	-0.246	-8.02***
Dept. 22	1,897	3.210	0.257	3.139	0.571	-0.071	-6.03***
Dept. 23	1,438	3.389	0.239	3.313	0.510	-0.076	-5.85***
Dept. 24	125	3.273	0.191	3.614	0.388	0.341	9.74***
Dept. 25	315	3.410	0.244	3.394	0.611	-0.016	-0.50
Dept. 26	796	3.255	0.252	3.451	0.49	0.196	11.83***
Dept. 27	467	3.505	0.268	3.620	0.435	0.115	6.07***
Dept. 28	195	3.510	0.243	3.527	0.457	0.017	0.56
Dept. 29	2,246	3.161	0.261	3.304	0.738	0.143	9.81***
Dept. 30	459	3.412	0.259	3.690	0.444	0.278	13.41***

Departments with fewer than 100 observations have been omitted.
Departments that are shaded are award grades significantly higher than predicted values at the 99% confidence level.
Positive t-statistics note high-grading departments; conversely, negative t-statistics note low-grading departments.
* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level

Table 9: ***Mobility Matrix***

Alternative System (Percentile)

Current System (Percentile)	10th	20th	30th	40th	50th	60th	70th	80th	90th	100th
10th	84%	14%	2%	0%	0%	0%	0%	0%	0%	0%
20th	15%	63%	19%	2%	1%	0%	0%	0%	0%	0%
30th	0%	24%	49%	25%	2%	0%	0%	0%	0%	0%
40th	0%	0%	26%	41%	28%	3%	2%	0%	0%	0%
50th	0%	0%	4%	26%	38%	24%	6%	2%	0%	0%
60th	0%	0%	0%	4%	24%	41%	21%	7%	2%	1%
70th	0%	0%	0%	1%	4%	25%	38%	25%	7%	0%
80th	0%	0%	0%	0%	2%	5%	31%	44%	17%	1%
90th	0%	0%	0%	0%	0%	0%	2%	21%	60%	17%
100th	0%	0%	0%	0%	0%	0%	0%	0%	15%	85%

The 10^{th} percentile indicates that a student is in the top 10 percent of his or her class. Conversely, the 100^{th} percentile indicates that a student is in the bottom 10 percent.

Table 10: ***Comparison of Class Rank Percentiles***

		N	Mean	Min.	Median	Max.
Major	**1**	8	3.3	-1.2	2.6	8.9
Major	**2**	164	3.8	-11.5	3.1	25.1
Major	**3**	24	0.3	-18.0	1.1	12.7
Major	**4**	43	-2.0	-13.4	-2.0	8.6
Major	**5**	20	-8.0	-29.2	-4.0	2.2
Major	**6**	65	-3.3	-17.2	-2.7	6.2
Major	**7**	37	0	-14.1	0	10.6
Major	**8**	29	1.2	-6.8	1.0	21.3
Major	**9**	37	-5.1	-30.3	-2.5	7.2
Major	**10**	14	-6.3	-15.7	-7.4	7.1
Major	**11**	157	4.9	-11.3	4.5	22.6
Major	**12**	37	-5.3	-18.5	-4.4	5.1
Major	**13**	27	-2.5	-17.6	-2.1	11.5
Major	**14**	29	-0.8	-13.6	-1.2	23.7
Major	**15**	38	-4.7	-25.4	-4.7	2.0
Major	**16**	12	3.7	-6.9	2.8	14.7
Major	**17**	152	-7.3	-38.2	-5.3	4.5
Major	**18**	12	-0.7	-9.3	0.9	4.0
Major	**19**	98	0.1	-14.6	-0.5	15.7
Major	**20**	39	1.7	-13.8	0.7	15.4
Major	**21**	36	3.1	-8.5	2.6	12.0
Major	**22**	129	3.7	-10.6	3.6	15.4
Major	**23**	63	-0.9	-14.9	-1.2	17.2
Major	**24**	6	-16.2	-38.2	-15.6	3.8
Major	**25**	5	0.5	-5.9	0.7	10.0
Major	**26**	17	-4.3	-13.2	-3.5	2.6
Major	**27**	13	-5.3	-13.6	-6.6	4.9
Major	**28**	12	-11.4	-25.4	-10.4	0
Major	**29**	67	1.1	-10.3	1.0	23.7
Major	**30**	11	-15.3	-26.2	-13.5	-10.3

Only majors that have corresponding departments with more than 100 observations in Table 8 are included.
Shaded majors are the high-grading departments from Table 8.

References

Adelman, C. *The new college course map of transcript files: changes in course-taking and achievement, 1972-1993.* 1995. Washington, DC: Office of Educational Research and Improvement, US Department of Education. http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/19/35/15.pdf

Achen, A.C. and P. N. Courant. *What Are Grades Made Of?* Journal of Economic Perspectives. 2009. Vol. 23, No. 3. p. 77-92.

Alexander, S. *Trophy transcript hunters are finding professors have become an easy mark.* The Wall Street Journal. April 27, 1993. p. B1, B12.

Archibold, R. C. *Just because the grades are up, are Princeton students smarter.* New York Times. February 12, 1998. http://www.nytimes.com/1998/02/18/nyregion/just-because-the-grades-are-up-are-princeton-students-smarter.html?pagewanted=1

Bar, T., V. Kadiyali, and A. Zussman. *Grade Information and Grade inflation: The Cornell Experiment.* Journal of Economic Perspectives. 2009. Vol. 23, No. 3, p. 93-108. http://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.23.3.93

Bar, T. and A. Zussman. *Partisan Grading.* August 21, 2009. http://findarticles.com/p/articles/mi_qa3673/is_3_120/ai_n28770301/

Bar, T. and A. Zussman. *Quest for knowledge and pursuit of grades: Information, course selection, and grade inflation at an Ivy League school*. 2005. http://digitalcommons.ilr.cornell.edu/workingpapers/53/

Becker, W. E. *Teaching economics to undergraduates.* Journal of Economic Literature. September, 1997. p. 1347–1373.

Birnbaum, R. *Factors Related to University Grade Inflation.* The Journal of Higher Education, September/October, 1977. Vol. 48, No. 5, p. 519-539. http://www.jstor.org/sTable/1981595

Bowen, W. G. and D. Bok. *The Shape of the River: Long-term Consequences of Considering Race in College and University Admissions*, Princeton University Press, 1998. Princeton, NJ.

Bruno, L. *Princeton leads in grade deflation.* USA Today. March 27, 2007. http://www.usatoday.com/news/education/2007-03-27-princeton-grades_N.htm

Cole, W. *By rewarding mediocrity we discourage excellence.* The Chronicle of Higher Education. January 6, 1993. p. B1-B2.

Cool, K. *Fear of Inflation.* Colby Magazine. http://www.colby.edu/colby.mag/issues/85n3/faculty1.htm.

Dickson, V. *An economic model of faculty grading practices.* Journal of Economic Education. 1984. Vol. 15, No. 3, p. 197-203.

Eckert, E. K. *The day I realized the game was over.* The Chronicle of Higher Education. June 1, 1988. p. A48.

Edwards, C. H. *Grade inflation: The Effects on Education Quality and Personal Well Being.* Education. 2000. http://findarticles.com/p/articles/mi_qa3673/is_3_120/ai_n28770301/

Eiszler, C. F. *College Students' Evaluations of Teaching and Grade Inflation.* Research in Higher Education. August 2002. Vol.43, No. 4, p 483-501.

Ellenberg, J. *Don't Worry about Grade Inflation.* Slate. October 2, 2002. http://slate.msn.com/default.aspx?id=2071759

Felten, E. W. *How Much Information is Conveyed by a Princeton Grade?* Princeton University Department of Computer Science. 2004. http://www.cs.princeton.edu/~felten/grading.pdf

Freeman, D. G. *Grade Divergence as a Market Outcome.* Journal of Economic Education. 1999. Vol. 31, No. 4. p. 344-351.

Gose, B. *Efforts to curb grade inflation get an F from many critics.* Chronicle of Higher Education. July 25, 1997. A41.

Goldman, L. *The Betrayal of the Gatekeepers: Grade Inflation.* Journal of General Education 1985. Vol. 37, No. 2, p. 97-121.

Healy, P. *Harvard's Honors Fall to the Merely Average.* Boston Globe. October 8, 2001

Healy, P. *Harvard's Quiet Secret: Rampant Grade Inflation.* Boston Globe. October 7, 2001

Jensen, E. J. and S. Wu. *Early Decision and College Performance.* Hamilton College. January 2010. http://academics.hamilton.edu/economics/swu/Early_Decision.pdf

Johnson, V. E. *An alternative to traditional GPA for evaluating student performance.* Statistical Science. 1997. Vol. 12, No. 4, p. 251-278.

Johnson, V. E. *Grade Inflation: A Crisis in College Education.* New York: Springer Publishers. 2003. http://www.aaup.org/AAUP/pubsres/academe/2004/JF/BR/perr.htm

Juola, A. E. *Grade inflation in higher education. Is it over?* March 1979. http://eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED189129&ERICExtSearch_SearchType_0=no&accno=ED189129

Juola, A. E. *Grade inflation in higher education: What can or should we do?* April 1976. http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/36/59/7d.pdf

Kuh, G. and S. Hu, *Unraveling the Complexity of the Increase in College Grades from the Mid-1980s to the Mid-1990s*. Educational Evaluation and Policy Analysis. Fall 1999, p. 297–320.

Lambert, C. *Desperately Seeking Summa*. Harvard Magazine. May-June 1993. p. 37.

Levine, A. and J. S. Cureton, *When Hope and Fear Collide: A Portrait of Today's College Student*. San Francisco: Jossey-Bass, 1998.

Millman, J. and S. P. Slovacek and E. Kulick and K. J. Mitchell. *Does grade inflation affect the reliability of grades?* Research in Higher Education. December, 1983. Vol. 19, No. 4.

McSpirit, S. and K. E. Jones, *Grade inflation rates among different ability students, controlling for other factors.* Education Policy Analysis Archives, 1999.

Pedersen, D. *When an A is average: Duke takes on grade inflation*. Newsweek. March 3, 1997. http://www.newsweek.com/id/95432

Quinones, E. *University's Grading Policy Change Marks Further Progress.* Princeton Weekly Bulletin, 2008. Vol. 88, No. 3.

Rojstaczer, S. *Grade Inflation at American Colleges and Universities.* Last updated March 2, 2009. www.gradeinflation.com.

Rojstaczer, S. *Grade Inflation at My Old Stomping Ground.* March 23, 2009. http://fortyquestions.blogspot.com/2009/03/grade-inflation-at-my-old-stomping.html.

Romer, P. *Growth Hinges on Science Education.* Stanford Business Magazine. 2002. Vol. 70, No. 3.

Rosovsky, H. and M. Hartley. *Evaluation and the Academy: Are We Doing the Right Thing?* American Academy of Arts and Sciences. 2002.

Sabot, R. and J. Wakeman-Linn. *Determinants of Performance in Introductory Courses In Economics and Seven Other Disciplines.* April 1991. http://www.williams.edu/wpehe/DPs/DP-7.pdf.

Sabot, R. and J. Wakeman-Linn. *Grade Inflation and Course Choice.* The Journal of Economic Perspectives. Winter 1991. Vol. 5, No. 1, p. 159-170.

Shea, C. *Grade Inflation's Consequences.* The Chronicle of Higher Education. January 5, 1994. p. A45.

Shoemaker, J. and M. DeVos. *Are we a gift shop? A perspective on grade inflation?* Journal of Nursing Education. 1999. Vol. 38, No. 9, p. 394-398.

Snare, C. E. *Implications of Considering Students as Consumers.* College Teaching, 1997, Vol. 45, No. 4.

Sonner, B. S. *A Is for 'Adjunct': Examining Grade Inflation in Higher Education.* Journal of Education for Business, September/October 2000. Vol. 76, No. 1.

Summerville, R. M., D. R. Ridley, and T. L. Maris. *Grade inflation: The case of urban colleges and universities.* College Teaching. Winter, 1990. Vol. 38, No. 1. p. 33-38. http://www.jstor.org/sTable/27558385

Valle, D. *Grade Inflation Hurts Best Stanford Students.* Stanford Review. October 1993. Vol. 11, No. 4. http://stanfordreview.org/old_archives/Archive/Volume_XXVII/Issue_7/1991-1993/1991-19932.shtml

Vars et al. *The Black-White test score gap.* Washington, DC, US: Brookings Institution Press. 1998, p. 457-479. http://psycnet.apa.org/psycinfo/1998-06583-012

Wallace, J. J. and A. Wallace. *Why the costs of student evaluations have long exceeded their value.* Issues in Accounting Education. 1998. p. 443-447.

Wilson, B. P. *The Phenomenon of Grade Inflation in Higher Education.* National Forum, September 22, 1999.

Winzer, Dr. M. *Grade Inflation: An Appraisal of the Research.* May, 2002. http://people.uleth.ca/~runte/inflation/Inflationb.pdf

Zirkel, P. A. *Grade inflation: A leadership opportunity for schools of education?* Teachers College Record. Winter 1999. Vol. 101, No. 2, p. 247-260.