GRADE INFLATION

A Cross-Institutional Analysis

Marguerite Kreger

I. Introduction

Depending on who you speak to, grade inflation is either a scourge or a myth. Since the 1970s researchers have attempted to determine both whether or not grade inflation existed as well as potential causes or influences. Nearly thirty years later, the resulting data is still conflicting and inconclusive.

Those who see grade inflation as a scourge identify the after-effects of affirmative action and multiculturalism, the increased importance of student evaluations on professors' compensation, and rising tuition as the key causes of the phenomena. Skeptics suggest that students today are smarter than they used to be, or even that the average GPA has decreased rather than increased. While the increase in grade inflation has led to more and more research on the topic, it tends to focus on the effect of student evaluations, but has barely considered the effect of a school's selectivity on its respective grade inflation. Whether grade inflation exists or not, we may learn something about grading practices at four-year colleges and universities by examining grade distributions at a sample of selective four-year colleges and universities. A test for grade inflation relative to institution selectivity while holding other factors constant will hopefully reveal a new indicator of variation in the grades that college students receive.

A number of institutions have taken measures to reduce the number of high grades once studies (Levine and Cureton, 1998) began to show an increase in the percentage of A's from 7% of all grades in 1969 to 26% of all grades in 1993. For instance, Princeton enacted a policy to limit the number of A's that can be awarded by each department to 35%. Students may have protested to the change last year (prior to the crackdown on grades, nearly half the grades awarded had been A's) but many other top colleges are expected to monitor Princeton's results (Mulvihill, 2005).

One of the problems caused by grade inflation is that grade distribution becomes compressed at the top of the grading spectrum. Mansfield (2001) makes the excellent point that this compression makes it difficult to discriminate between the best students and the very good students, the very good and good, etc. etc. Compression at the top of the scale does not allow for the gradation that is necessary in order to accurately rate students. For example, raising grades above what students actually deserve could award an average student a possible B+ for what should be considered C level work. Sabot and Wakeman-Linn (1991), suggest that one of the main reasons for grades is to convey information to students about their relative strengths and weaknesses. If this is accurate, then grade distributions with more dispersion and a lower average will actually be preferable. Differences in grades across departments may cause students to specialize in areas with the highest grades rather than the areas best suited to their individual talents. Compression of the grade distribution at the top of the scale also makes grades a poor indicator of achievement for future employers.

While these are all important issues, I two issues have been generally overlooked, or at least under-studied. First, is the effect of the selectivity of a school on grade inflation. Some experts would argue that more selective schools have higher quality students and higher grades are a reflection of that. If this is the case then grades will exhibit the same informational deficiencies noted above. Grades may not provide accurate information to students about their relative strengths and employers evaluating graduates from the same school may not be able to make judgments about relative skill.

The second issue considered is the variance in grading between public and private schools. The gap in GPA between each type of institution has widened over time, but the trend for both public and private schools is similar in that both have undergone grade inflation for the past 35 years. This paper will attempt to study the relationship between school selectivity and rates of grade inflation.

II. Background

A number of explanations have been advanced to explain the variation in grades over time and across institutions. One influence on grade inflation is the expectations that students have of their school and their grades. Mansfield (2001) contends that Affirmative Action movements and the behavior of professors towards minority groups had some effect on grade inflation. In addition there is some discussion about both the variance in grade inflation between public and private schools, as well as between different departments at the same school (Sabot and Wakeman-Linn 1991). However, the most studied influence on grade inflation is student evaluations. These studies seem to monopolize much of the literature that can be found on grade inflation. By far the most studied cause of grade inflation is student evaluations. Mansfield (2001) comments that professors react to student expectations which are reflected in evaluations: "With an eye to student course evaluations and confounded by the realization that they have somehow lost authority, professors begin from what they think students expect."

Lichty et al. (1978) make a theoretical case for the idea that grade inflation, or at least lax grading standards by professors, lead to higher student ratings for that instructor. "If two professors are equal in every other respect, we would predict that the professor with the highest grading structure… will tend to receive the highest student ratings." The theoretical section of his paper suggests that the changes in student evaluations of professors are a function of the grading patterns and the work effort required by the professor.

Zangenehzadeh (1988) also attributes grade inflation to the introduction of student evaluations of teachers; the results of his studies indicate that ratings of an instructor are very highly significant factor in relation to students' grade expectations. Teachers are said to give higher grades to increase their evaluation ratings by students, and the students are said to evaluate their instructors according to their performances in class. Kanagaretnam et al. (2003) find that the use of student evaluations has led to a reduction in student knowledge and a manipulation of grades, which is consistent with an increase in the number of students receiving As and Bs in the last two decades. Specifically, this study revealed that when professors require a high or excessive level of student ratings, they will turn to soft grading or free points in order to increase student satisfaction.

It is generally accepted that student satisfaction is an increasing function of their grades. Krautmann and Sander (1999) add that if evaluations can be increased by giving higher grades, then they are a flawed instrument for the evaluation of teaching. Therefore this process of student evaluations could contribute to grade inflation in higher education if the faculty have an incentive to increase their evaluations. The findings of Krautmann and Sander's (1999) study indicate that grades affect student evaluations, suggesting that faculty have the ability to 'buy' higher evaluations by lowering their grading standards. One effect is that long-term grade inflation could be partly caused by the increasing importance of student evaluations in teacher compensation.In addition to student evaluations, students' expectations for their own grades have several different effects on how grades are given out and distributed. According to Mansfield (2001), a Harvard professor, Harvard itself has a very inflated grade distribution wherein 25% of the grades given out are A's and another 25% are A-'s. In this case, it is difficult for students to work hard or professors to get them to work hard if their chances of getting an A or A- are 50-50.

Rojstaczer (2003) considers another possibility; students are paying successively higher tuition each year, and in return for this, they want a good grade as a reward for their purchase. In this situation, "professors are not only compelled to grade easier, but also to water down course content. Both intellection rigor and grading standards have weakened." Kanagaretnam et al. (2003) comment on the common practice of offering a 5-10% class participation grade to virtually everyone in the class. Some professors also change their own grade distributions by lowering the cut-off points for grades. This combination of free points and soft grading causes the end grades to be other than a real measure of the students' performance. This leads to inappropriate factors, like a professor's toughness in grading, free points, teaching effort, student quality, a student's level of risk aversion, and class size, having a large effect on grades, more so than actual student effort and comprehension of the material. These few sources demonstrate several of the ways that grade inflation can be caused, or at least influenced by, students' expectations of higher grades.

According to many scholars in the field of education, the goal of teaching is to make students feel capable and empowered. This suggests that grading students strictly could be both cruel and dehumanizing (Mansfield, 2001). Grading is construed as stress inducing and judgmental. Some authors suggest that with the initial affirmative action initiatives professors graded minority students more generously, but to make this fairer, they consequently graded all students easier. These ideas regarding self-esteem as related to education go "hand in hand with multiculturalism or sensitivity to people of diverse races and ethnicities – meaning that professors must avoid offending the identities of victimized groups." (Mansfield, 2001) Reports by Rojstaczer (2003) however dispute this idea because much of the rise in minority enrollments in college occurred during the 1970s and 80s. During this era grade inflation actually waned, so it is unlikely that affirmative action actually had a significant influence on grade inflation.

The humanities tend to have higher grades than the social sciences, which in turn have higher grades than the natural sciences. We cannot assume that a school's best students are simply enrolled in the study of humanities and the worst students study natural sciences. In fact, science students tend to do better in non-science courses than non-science students do in science courses (citation). If they can be considered valid indicators, the fact that there is no significant difference in either SAT scores or grades in other courses among students in different departments suggests a problematic discrepancy in grading practices. The 1970s and 80s saw a decline in the number of graduating science majors as students began to make their course selections in response to the expected grade.

The result was that many universities split into high (Art, English, Philosophy, Political Science) and low (Economics, Chemistry, Math, Psychology) grading departments. (Sabot and Wakeman-Linn, 1991) Many experts suggest that this could be one reason that American universities have fallen behind in technology and science: "A conflict exists between the incentives offered to students and the institutional goal of increased science and math education." There has been a substantial variation in the pace of inflation between departments which led to an increase in the variation of the mean grades across departments.

Much of the current literature available on grade inflation tends to minimize the effect of the selectivity of a university on its respective grade inflation. There are two suggestions offered by the experts to explain the apparent link between higher grades and higher selectivity. First of all there is the idea that selective schools have smarter students who are therefore capable of earning these higher grades. The second is that over time there has been an overall improvement in student quality. However, there is not much evidence to support either claim. For example, Mansfield (2001) suggests that "When bright students take a step up and find themselves with other bright students, they should face a new, higher standard of excellence." He disputes the theory that students today are possibly better than those in the past and therefore deserve higher grades by making the suggestion that we instead raise standards and demand more of current students.

Rojstaczer (2003), on the other hand, argues that selective schools do not have more grade inflation, but that they simply started at a higher baseline GPA. His studies indicate that grade inflation can be found "virtually everywhere" and its rate of change (in terms of change of GPA) is also about the same across universities. The apparent difference in grade inflation is thus caused by the fact that less selective colleges started off with a lower baseline GPA. Rojstaczer (2003) is also skeptical that increases in student quality have caused the increases in GPA. Some studies have tried to show a relationship between SAT and ACT scores that is associated with the changes in grades. This theoretical trend cannot even be explained by local increases in student quality, as there is no evidence that students nationwide have improved since the mid-1980s. Another facet of this, analyzed by Zangenehzadeh (1988), rejects the idea that students are smarter than before because if SAT scores are acceptable indicators of academic ability, they indicate a downward rather than an upward trend in student intelligence.

Much of the currently available research focuses on the effect of student evaluations on grade inflation. We would choose to focus this study more so on admissions selectivity and its influence on higher grades. By controlling for factors such as whether a school private or public, selective, as well as certain administrative policies, we aim to find a correlation between grading patterns and institutional characteristics. One of the major problems of the research and studies done thus far on grade inflation is that there is little analysis of differences in grading across institutions. This study aims to provide at least some preliminary evidence on the causes of differences in grading across institutions.

III. Data and Methods

Previous studies of grade inflation have used time-series data for individual institutions (Rojstaczer 2003). In this case, however, other variables are being considered to determine which factors drive the average GPA higher, rather than whether or not there actually is an increase over time. Hypothetically, the average grade on a 4.0 scale should always be 2.0. While there are certain things that could shift this, for example, the number of credits certain students take [a student taking 15 credits with a GPA of 3.5 is accorded the same weight as a part-time student taking 6 credits with a 2.0 GPA might bias the results of an average GPA measurement], these biases are unlikely to account for all the variation in GPA averages across institutions.

Data for this study was collected through a variety of methods, so it is unfortunately not a true random sample. A survey of a random sample of schools was conducted in order to obtain their average undergraduate GPAs; however, many of the offices contacted were unable to provide this information.[1] The other source used to obtain this grade data was Rojstaczer's website, GradeInflation.com, which catalogs the average GPAs over time from a sample of schools large enough to be considered normally distributed. This resulted in a data set of GPA values pulled from a sample calculated over the years from 1992-2004.

The independent variables for the schools in our sample were pulled from a current database of the *National Center for Education Statistics* and the *2003 Peterson's Guide to Colleges*. In this way we were able to create a large sample with many variables, including institution selectivity (SEL), top 75th percentile of SAT scores of admitted students (SAT), number of undergraduates enrolled (ENRL), endowment per student (ENDW), average financial aid award (FINA), full-time undergraduate in-state tuition (TUIT), whether an institution is public or private (PRIV), freshman retention rate (RR), whether or not the institution is a Doctoral/Research university (RES), and the ratio of students to faculty (STR). These were the factors initially considered as potentially having an effect on an institutions average GPA.

[1] We used a standard institutional database to randomly select one hundred four-year colleges and universities with acceptance rates of 85% or lower. We called each school using a standard script and asked about the average GPA and grade distribution in undergraduate classes. Only ten institutions offered to provide the information.

Different effects were anticipated from each of these characteristics, and not every a priori assumption was correct. It was expected that more selective schools with a smaller selectivity percentage (calculated by dividing the number of students who applied by the number who were accepted) would have a higher GPA. The same was predicted for SAT score (on the typical 1600 point scale), often considered one of the best indicators of future performance in college. The size of a school, measured by the number of undergraduates enrolled, also had the potential to skew the average GPA; a large number of students could lead professors to grade more harshly, lowering the GPA. It may be easier to fail a student you do not know in a class of two hundred than the student you know in a class of ten.

The endowment per student, calculated by dividing endowment by enrollment and measured in millions, was considered because schools with a higher proportional endowment may be less dependent on tuition as a source of revenue. If schools are less tuition dependent, they are more likely to fail low performers. A higher average financial aid award could motivate worthy students to put more effort into their studies, again raising their GPA. Tuition (full-time undergraduate in-state amount for our purposes) was considered after Rojstaczer (2003) mentioned that increases in tuition could lead to student expectations of higher grades with a 'get what you pay for' mentality. The public vs. private factor (measured with a dummy variable where PRIV = 1 means the institution is private, and PRIV = 0 indicates a public institution) was also expected to be a significant factor; it was expected that private institutions would award higher grades, thereby driving up their average GPA. Private schools are also generally more tuition dependent and they may give out higher grades simply to keep their "customers" happy.

Freshman retention rate was a prospective factor because it was less likely that students achieving higher grades in college would be as inclined to transfer. Because some studies (Sabot and Wakeman-Linn 1991) have shown that grade inflation is less prominent in the science and math studies it was possible that a school being classified as a Doctoral/Research University (a dummy variable where RES = 1 shows that an institution is a Doctoral/Research University) would drive the GPA downward. The student-faculty ratio was expected to have similar results to the student enrollment. Larger classes could lead to less intensive grades and coursework or have the opposite effect if professors are less interested in seeing all of the students in a large class excel.

IV. Results and Analysis

Regression results are reported in Table 2. Tests showed no evidence of heteroskedasticity, but a variance inflation factor test revealed that some of the independent variables were highly correlated. For instance, SAT score was very highly correlated with institution selectivity and whether or not an institution was a Doctoral/Research university. Another correlation that should not have been surprising was the correlation between tuition, financial aid, and the public vs. private dummy variable. Due to this link, the significance of each of these variables improved considerably when the others were eliminated from the regression. Because there were questions about the randomness of our sample, we calculated robust standard errors in each of the regression procedures.

Because many of the independent variables were correlated we ran a series of regressions, deleting some variables and adding others to reduce the collinearity problem.

GPA was first regressed on all of the variables we considered, and then selections of those as their effects became more apparent, as well as their interactions with one another. Each regression successively made an attempt to avoid correlation among the independent variables as well as maintain manageable VIF values, while at the same time creating an equation with a reasonably predictive R^2 and significant F-value.

Regression 1 (see Table 2) regressed GPA on every variable previously discussed, but was acknowledged as having the potential to be misleading. In addition to the constant intercept, SAT scores, whether or not the school was a Research institution, selectivity, and retention rate appeared to be significant. Not all of the significant variables had the expected effect however. The higher an institution's accepted SAT scores, the higher the GPA will be, and a Doctoral/Research university will tend to have a lower GPA (about 0.1 points lower on a 4.0 scale) than a Baccalaureate/Liberal Arts or Masters institution, all as expected. Yet less selective schools initially appear to have a higher GPA; a higher percentage selectivity (meaning less selective) will drive up the average GPA. Also, the coefficient on retention rate is negative, suggesting a negative relationship in that students are less inclined to return to a school with a higher average GPA.

It was also here in Regression 1 that the correlation between Tuition, Financial Aid, and Public vs. Private became evident (VIF of 21.98, 14.89, and 16.20 respectively). Specifically, in Regression 1 higher tuition is indicating a lower GPA, and none of these three variables are significant. This is surprising considering our expectations. It is a possibility that these variables are being affected by others in the regression, and their signs could change dependent on the other independent variables included in each regression.

In Regression 2, all insignificant variables (enrollment, endowment per student, and student-teacher ratio) were dropped from the regression except for the Private/Public binary variable. Without the correlation of private/public with tuition and average financial aid, the private variable becomes significant at the 5% level and indicates that private schools are more likely to have an average GPA of about 0.13 points higher than public institutions. SAT scores also remained highly significant (if the score increases by 100 points, the institutions average GPA increases by almost 0.1 points on a 4.0 scale). However, this regression also dropped the importance of the research, selectivity, and retention rate variables were all insignificant. On the positive side, the sign of the coefficient attached to selectivity now matches our assumptions.

SAT score was dropped from Regression 3 to correct for the effect of its correlation with selectivity. Regressing GPA against selectivity, retention rate, and the private/public variables yielded impressive results. Every variable was significant, and each had the coefficient sign that common sense indicated it should. To be more explicit, the less selective a school is, the lower its average GPA is anticipated to be (estimate of SEL = -0.26147), while private institutions and those with relatively higher retention rates will see increases in the average undergraduate GPA. Once again, students in private institutions have higher GPAs.

In Regression 4 we tested the significance of tuition without the simultaneous effect of the private/public factor. While tests for the usefulness of the model (R^2 = 0.667 and F-value = 30.71) are almost identical to those for Regression 2, the only variable that emerges as significant in this case is SAT score (again, an increase in SAT score by 100 points raises the average GPA by about 0.1 on a 4.0 scale). Selectivity and Research institutions maintain their negative coefficients but lack significance. Also, the retention rate has reverted to having a negative effect on GPA which seems highly illogical. Tuition, however, is insignificant.

Our final equation, Regression 5, again retained the tuition variable without the private variable, and also eliminated the Research aspect and SAT score which we suspect are highly correlated with selectivity. These adjustments had the predicted effects of making both selectivity and tuition significant, as well as correcting the sign on the coefficient for the retention rate. In general, the results of Regression 5 are consistent with the other regressions. Decreases in selectivity (increases in the selectivity variable value) will drive down an institutions average undergraduate GPA and increases in the retention rate will also be related to higher grades. Finally, an increase in tuition of one thousand dollars will increase the average GPA by 0.007 points on a 4.0 scale.

V. Conclusion

Trends in grade point average warrant our attention because of its use as an indicator of student achievement. If a students GPA is no longer a valid gauge, institutions and employers will be hard-pressed to find other measurements to differentiate students. Some colleges and universities have taken measures to correct these more recent inaccuracies and unequal grading distributions, either by restricting the number of A's that can be awarded to students, or extending their grade scale to 4.3 to restore a more normal grade curve to the grading distribution.

The goal of this study was essentially to determine some of the factors that do drive up an institution's average grade point average. Based on our results, the most important determinants appear to be the cost of tuition and SAT scores of incoming students. Some other variables that proved to be influential in varying degrees dependent on which other variables were present in specific regressions were the freshman retention rate, whether or not a school is a Doctoral/Research University, and school selectivity. The difficulty in this study lay in the very strong correlation among the variables used.

The results may face the argument that the institutions with higher GPAs do not suffer from grade inflation, but merely more intelligent students. Yet the disadvantages of a system where student grades are so compressed certainly has its drawbacks. This subject is worth further study once you recognize that the average GPA in our study was 3.051, more than 50% higher than an "average" 2.0 C grade.

REFERENCES

"IPEDS Peer Analysis System." *National Center for Education Statistics.* Ed.gov. Data year for data: 2004.

Kamber, R., ----, "Grade Inflation as Epidemic."

Kanagaretnam, K., R. Mathieu, and A. Thevaranjan, 2003, "An Economic Analysis of the Use of Student Evaluations: Implications for Universities," *Managerial and Decision Economics*, 24, 1-13.

Krautmann, A.C. and W. Sander, 1999, "Grades and student evaluations of teachers," *Economics of Education Review*, 18, 59-63.

Lichty, R.W., D.A. Vose, and J.M. Peterson, 1978, "The Economic Effects of Grade Inflation on Instructor Evaluation: An Alternative Interpretation," *Journal of Economic Education (Pre-1986)*, 10, 3-10.

Mansfield, H.C., 2001, "Grade Inflation: It's Time to Face the Facts," *The Chronicle of Higher Education*, 4/6/2001, B24.

Mulvihill, Geoff, 2005, "Princeton cracks down on high grades," *Asbury Park Press: The Associated Press*, 1/23/2005, A13.

Nelson, J.P. and K.A. Lynch, 1984, "Grade Inflation, Real Income, Simultaneity, and Teaching Evaluations," *Journal of Economic Education (Pre-1986)*, 14, 21-37.

Peterson's … 4 Year Colleges. 33rd ed. Lawrenceville, NJ: 2003.

Rojstaczer, S., 2003, "National Trends in Grade Inflation: American Colleges and Universities," *GradeInflation.com*, 1-7.

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

Zangenehzadeh, H., 1988, "Grade Inflation: A Way Out," *Journal of Economic Education (1986-1998)*, 19, 217-227.

Table 1 Means and Standard Deviations

Variable	Mean	Standard Deviation	Minimum	Maximum
GPA	3.051	0.226	2.51	3.47
SEL	0.589	0.242	0.104	0.979
ENDW	0.185	0.436	0	2.748
ENRL	10.080	8923.22	443	36049
FINA	12.127	6564.09	2400	26168
PRIV	0.416	0.496	0	1
RES	0.487	0.503	0	1
RR	0.833	0.120	.44	.98
SAT	1306.74	162.976	950	1580
STR	14.300	4.789	4	23
TUIT	12.631	11589.97	0	31260

GPA: average institutional grade point average on a 4.0 scale.
SEL: selectivity expressed as a percentage calculated by dividing students admitted/applied.
ENDW: endowment per student calculated as endowment ($ millions) / undergraduates enrolled.
ENRL: number of undergraduate students enrolled in thousands.
FINA: average financial aid award in $ thousands.
PRIV: private (PRIV = 1) versus public (PRIV = 0) institutions.
RES: Doctoral/Research Universities (RES = 1) or Baccalaureate/Masters Colleges (RES = 0).
RR: retention rate of returning freshmen.
SAT: combined math and verbal SAT score.
STR: student:faculty ratio.
TUIT: in-state, full-time undergraduate tuition in $ thousands.

Table 2 Regression results for undergraduate GPA as a measure of grade inflation.

Regression	(1)	(2)	(3)	(4)	(5)
Variable					
Constant	1.58186*** (0.4540)	2.16684*** (0.3391	2.82359*** (0.1990)	2.14008*** (0.3650)	2.90699*** (0.1934)
RES	-0.09689* (0.0589)	-0.04912 (0.0436)		-0.06708 (0.0447)	
SEL	0.02938* (0.1517)	-0.12238 (0.1209)	-0.26147*** (0.0940)	-0.16205 (0.1236)	-0.31076*** (0.1004)
RR	-0.68028* (0.3786)	-0.30864 (0.2365)	0.37042** (0.1880)	-0.36804 (0.2812)	0.29986 (0.1867)
ENRL	0.00324 (0.0066)			-0.00114 (0.0033)	
ENDW	0.02679 (0.0475)				
FINA	0.01571 (0.0106)				
TUIT	-0.00870 (0.0054)			0.00259 (0.0033)	0.00718*** (0.0028)
PRIV	0.10906 (0.1826)	0.12757** (0.0585)	0.17905*** (0.0435)		
STR	0.00353 (0.0108)				
SAT	0.00141*** (0.0005)	0.00090*** (0.0003)		0 .00102*** (0.0003)	
	n = 44 R^2 = 0.722 F = 17.44 Pr > F = 0.0000	n = 61 R^2 = 0.678 F = 36.13 Pr > F = 0.0000	n = 72 R^2 = 0.567 F = 32.54 Pr > F = 0.0000	n = 61 R^2 = 0.667 F = 30.71 Pr > F = 0.0000	n = 72 R^2 = 0.542 F = 29.54 Pr > F = 0.0000

Standard errors in parentheses.
*** = significant at 0.01
** = significant at 0.05
* = significant at 0.10