Do US News and World Report College Rankings Encourage Grade Inflation?

By Maxim Lott

May 11, 2009

Abstract

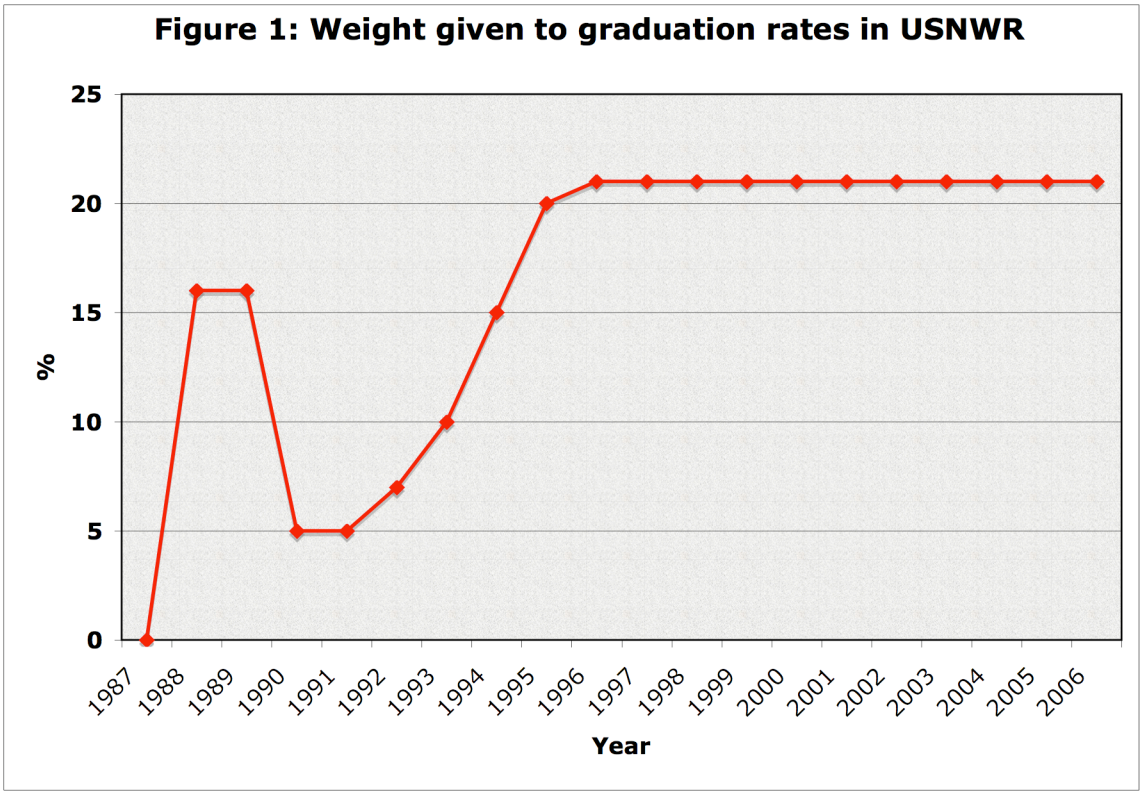
US News and World Report (USNWR) bases a full 21 percent of its influential university rankings formula on graduation rates. By making the assumption that higher graduation rates are good, USNWR implicitly assumes that higher rates are primarily caused by better counseling services and programs that help keep students from dropping out. But is this assumption warranted? Another hypothesis is that universities with higher graduation rates, everything else equal, are just easier on their students. They give B's when others would have given C's, and may have fewer graduation requirements in general.

Following this logic, one might fear that US News's inclusion of graduation rates would give universities a reason to further inflate grades. I test that hypothesis by using a fixed effects model to estimate the effect that changes in the weight given to graduation rates had on grades awarded at 56 universities between 1988 and 2003.

I. The History of USNWR Rankings and Grade Inflation

USNWR's "Best Colleges" ranking began in 1983 as a survey that asked university presidents to rate their peer institutions. It was not until 1988 that USNWR began to base rankings on data – including nearly all the factors that are used today, such as graduation rates, sat scores, acceptance rates, student-to-faculty ratios, and financial resources.

As shown in Figure 1, the weight that USNWR gave to graduation rates changed 7 times between 1988 to 1996.[1] The frequency and size of these changes should give us a lot of statistical power to tell whether changes in the weight correspond with changes in grade inflation.



Why all the changes? USNWR usually justifies them with generic statements like, *"Because 'outcomes' research shows that a school's retention rate – it's ability to retain and*

[1] *Best Colleges*, U.S. News and World Report, volumes 1987 to 2006.

graduate its students – is an important measure of academic quality, we raised the weight of this attribute from 15 to 25 percent."[2] More cynical observers argue that USNWR makes changes in order to shake up the rankings, which gets people to buy the magazine every year.

But USNWR sometimes does pay attention when schools pursue potentially detrimental policies for the purpose of moving up the rankings. In 2003, USNWR eliminated the yield rate in their formula – which only had a weight of 2 percent to begin with – because many worried that it was encouraging universities to reject the best students out of fear that they would go somewhere else. It was also seen to be needlessly encouraging early decision policies, which allow universities to admit only students they know will attend.

"What we've been told is that schools felt they could manipulate their ranking by manipulating the yield number," Sara Sklaroff, USNWR's education editor, explained.[3]

But that manipulation would be done by admissions officials, whereas grade inflation would be done by individual professors. Many are skeptical that professors care about USNWR rankings enough to change the way they grade. As one Institutional Research director told me when I asked for grade distribution data: *"Our faculty are so removed from the rankings and the relationship between grades and graduation rates that if there is a correlation, it is not a causal relationship.*"[4]

But it may not need to be the professors themselves who care about the rankings – an alternative hypothesis is that the rankings cause administration officials to exert less pressure on

---

[2] *The Methodology: How U.S. News determines the rankings in this guide*, U.S. News and World Report, 9/18/1995, p. 123. Note: The new retention measure accounted for 5 percent of the new formula, leaving 20 percent for the graduation rate.

[3] *College Rating By U.S. News Drops Factor In Admissions*, by Jacuqes Steinberg, the New York Times, July 10, 2003

[4] E-mail from Paula Maas, Executive Director of the Center for Institutional Effectiveness at The College of New Jersey - April 27, 2009.

department heads to keep grade inflation in check, or officials may decide not implement anti-inflation policies that they otherwise would have.

Several academic studies show why university administrators care so much about even minor aspects of USNWR's ranking formula.

James Monks and Ronald G. Ehrenberg, for instance, find that slipping six places in the rankings causes a university's yield rate to fall by 1 percent, and the average SAT score falls by 16.6 points.[5] They also find a negative effect on the graduation rate of the entering class that year.

For public institutions, Ginger Zhe Jin and Alexander Whalley conclude in a 2007 paper that an increase in rank leads to a 6.5 percent increase in state funding and a 4 percent increase in the institution's instructional budget.[6]

Given these strong effects, it would be rational for university presidents to care about their institution's rankings and do what they could to make them better.

Grade inflation is of course nothing new, and the start of serious grade inflation is usually attributed to the Vietnam War, when professors wanted to help students pass so they could avoid the draft. But as inflation goes on with no sign of slowing, many worry that it is reaching unacceptable levels (Pomona College, for instance, now has an average GPA of 3.51.)

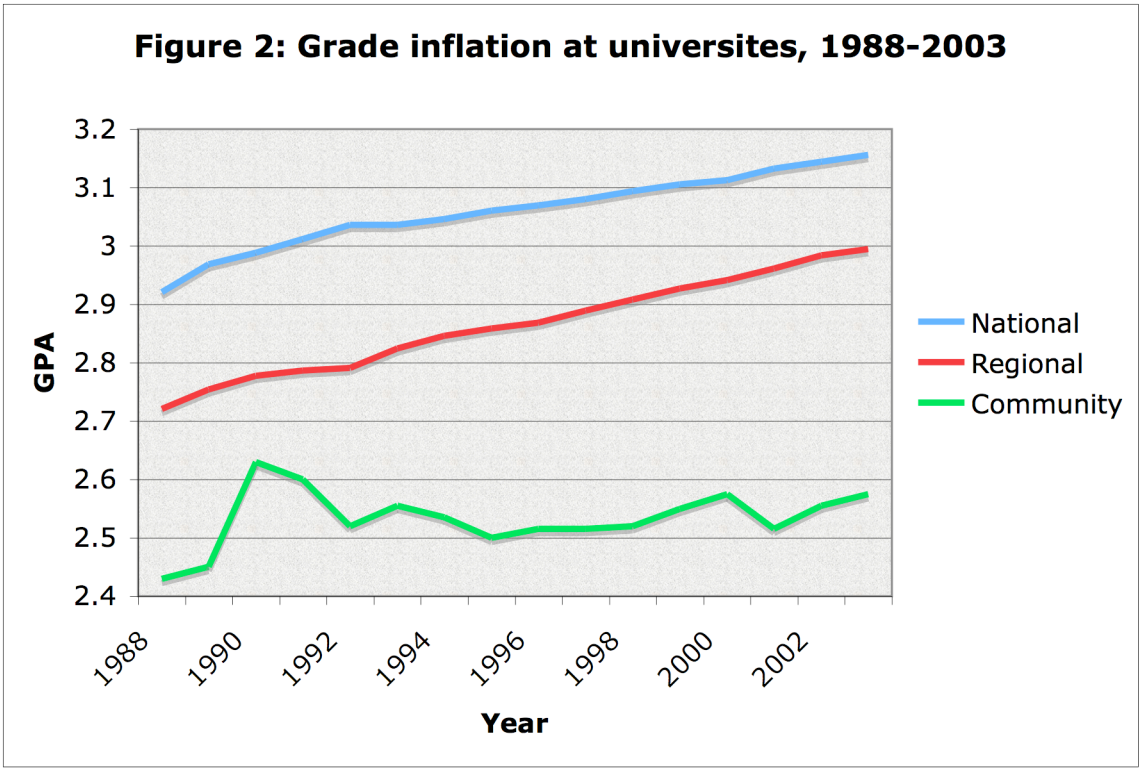
Among the 54 four-year schools we have data for, the average GPA went from 2.84 in 1988 – when USNWR introduced formulaic rankings – to 3.1 now. It grew at a similar rate

---

[5] *The Impact of U.S News & World Report College Rankings On Admission Outcomes and Pricing Decisions at Selective Private Institutions*, by James Monks and Ronald G. Ehrenberg, NBER Working Paper No. 7227 - July 1999.

[6] *The Power of Information: How Do U.S. News Rankings Affect the Financial Resources of Public Colleges?* by Ginger Zhe Jin and Alexander Whalley, working paper, 2007.

between 1960 and 1988 (the average was 2.6 in 1960.)[7] It would also be interesting, if one could get the data, to compare grade inflation the U.S. to that of other countries.



A couple universities have taken active steps to combat grade inflation in recent years. In 2004, Princeton began capping the number of "A's" given by each department to 35 percent, causing the average GPA to fall by .07.[8] In 2006, the University of Colorado began trying to rein in grade inflation by publicizing information about the distribution of grades in departments and

[7] Data is from Stuart Rojstaczer, a retired Duke University professor, and available at gradeinflation.com.

[8] Students dig in as Princeton caps A's: Departments can award top marks to no more than 35 percent, The Associate Press, Jan. 22, 2005 - http://www.msnbc.msn.com/id/6856357/%5Benter%20URL%5D

standardizing grading practices in large lecture classes.[9] And Dartmouth students' transcripts list the average grade awarded in each class taken by the student along with the student's own grade.

But by discouraging grade inflation, universities risk reducing their graduation rate – a number that determines 21 percent of their USNWR ranking. Could the rankings have deterred administrators from trying to keep inflation in check?

[9] CU's fight against grade inflation deflates overall GPA, by Erika Gonzalez, Rocky Mountain News, August 16, 2007 - http://www.rockymountainnews.com/drmn/education/article/0,1299,DRMN_957_5674283,00.html

II. Data Collection and Methodology

To study the effect of changes in the USNWR ranking formula on grade inflation, I used average undergraduate GPAs from 56 institutions of higher education that reported data for all fall semesters from 1988 to 2003.[10] 34 are nationally ranked, 20 are regionally ranked, and two are community colleges. The GPA data on these came from retired Duke professor Stuart Rojstaczer, who collects data on university GPAs.[11]

I also went through past USNWR rankings and coded each school based on whether it was public or private, liberal arts or standard (based on USNWR criteria,) four year or community, and the tier it was given in USNWR.

I collected SAT score data on nationally ranked schools from old USNWR issues in order to control for the degree to which grades increased because of better-prepared incoming classes. My SAT data covers nearly the entire period during which the weight given to graduation rates changed, from 1989 to 1997 – SAT scores were not reported in 1988.[12]

The most important thing to control for in this analysis, however, is the general upward trend in GPA. Because GPA is rising over the whole period, any variable that also trends upward (such as the weight given to graduation rates in USNWR) will be positively correlated. I account for that problem by using university- and year- fixed-effects.

The basic fixed effects specification model has the following form:

(1) $Y_{ut} = \gamma Z_{ut} + \beta X_{ut} + W_u + V_t + \varepsilon_{ut}$

[10] A few schools have a year or two missing at the start or end of the range, but are still usable because the data they do have covers the years when UNSWR's weight was changing.
[11] Data and sources are available on his website, gradeinflation.com.
[12] There were several irregularities in the raw data, such as "recentering" of scores by the College Board and the reporting of ACT scores. See appendix B for details on how these were dealt with.

Where Y is the average grade given in each year for school U, X is the average SAT score, Z is the weight given to graduation rates in USNWR, W is for university fixed effects, and V is for year fixed effects. ε is a mean-zero random error term.

However, the effect of an entering class's SAT score is spread throughout all four years that they are in school. Therefore I also control for SAT scores that have been lagged from one to three years. It may also be the case that it takes time for schools to react to a new USNWR graduation rate weight, so I add a lagged variable for that. The new model has the form:

$$(2)\ Y_{ut} = \gamma Z_{ut} + \gamma_1 Z_{ut-1} + \beta_0 X_{ut} + \beta_1 X_{ut-1} + \beta_2 X_{ut-2} + \beta_3 X_{ut-3} + W_u + V_t + \varepsilon_{ut}$$

Additionally, all regressions are run using robust standard errors to deal with potential heteroskedasticity. I also ran a Lagrange multiplier test to detect autocorrelation: The coefficient on the lagged residuals on the residuals was insignificant (p-value .815,) indicating that there is no serious autocorrelation problem.

Another potential problem that needs to be considered is self-selection in the data. One might suspect that only the schools with the least grade inflation would report scores. On the other hand, schools with the worst grade inflation may be more likely to commission reports to investigate the problem. Whichever effect dominates, self-selection bias is mitigated because Rojstaczer tries to get data from as many institutions as possible, and uses internal university reports as well as public ones. On his website, under the headline "Additional Contributions

Wanted," he asks for anyone with verifiable data to send it to him. He also offers to keep senders anonymous.[13]

A final issue with the data is that Rojstaczer's GPA data is for all undergraduates. With that measure, it would take 4 years for any inflationary policy to be fully realized in the variable; we really want to know the average grade given out in any particular year. I decided to estimate that by deconstructing the undergraduate GPA variable, which is possible based on some reasonable assumptions about the distribution of grades (see Appendix A for details.)[14]

Earlier in the paper, I raised the question of whether grade inflation came from professors, or whether it came more indirectly from administrators. We can attempt to address that question by looking at whether the main increase in grades occurs in the lower grades (a lot of D's turning into C's) or the higher ones (A- grades turning into A's.) The former type of change would likely have more of an effect on graduation rates due to the fact that those in danger of failing out of school likely tend to get grades concentrated in the low grades levels. Therefore, a professor trying to inflate grades for the purpose of increasing his school's ranking should be more likely to increase grades on the lower end of the spectrum. Using data on grade distributions, then, allows me to test whether professors do it for that reason.

But it would not disprove the original hypothesis if inflation were evenly distributed across the grade spectrum – rather, it would seem to indicate that the effect comes from administrators.

For data on this, I e-mailed the institutional research directors at 50 universities. Five of them – the University of Delaware, DePauw University, Victoria College, Georgia Institute of

[13] www.gradeinflation.com

[14] As a robustness check, I also ran the regression using the original undergraduate GPA variable and using 3 lags. The lagged coefficients as a group are statistically significant in a joint-significance f-test.

Technology, and Roanoke College – were willing and able to provide detailed distribution data between 1989 and 2001.

That is a very small sample, but fortunately we have another way to measure this due to the fact that I have data from six schools who report the average GPA for graduating seniors instead of the average undergraduate GPA. If professors are indeed inflating grades to help boost the graduation rate, and as a result there was more of shift in the lower parts of the grade distribution, we would expect there to be more marginal graduates in the graduating class. Then we should expect to see a smaller increase in the GPA of graduating classes than we do for undergraduates as a whole. In some sense, this can serve as a proxy for the actual distribution of grades.

Table 1: Details of Sample (See Appendix B for complete list of schools)

Private Institutions	42%
Liberal Arts Institutions	32%
Four-year Institutions	96%

Table 2: Variables

Variable	Obs.	Mean	St. Dev.	Min	Max
Mean SAT score of nationally ranked institutions	271	1151.5	113.85	970	1400
GPA of nationally ranked institutions	553	3.05	.193	2.47	3.44
GPA of regionally ranked institutions	290	2.88	.180	2.46	3.4
GPA of community colleges	30	2.59	.239	2.31	3.16
GPA of private institutions	367	3.08	.217	2.47	3.44
GPA of public institutions	506	2.90	.184	2.31	3.38
USNWR weight given to graduation rates	873	16.39	6.01	5	21

III. Results

Using the fixed effects model *without* controlling for SAT scores, I find that when USNWR adds one percent to the weight given to graduation rates in their formula, the average grade awarded at 4-year institutions rises by .0089 the same year, and by .0022 in the next year (the next-year effect is insignificant, but a f-test finds that both variables are significant together.) Adding those effects together to get .0111, and multiplying by 21 (for the 21 percent weight given to graduation rates,) we get that the policy as a whole caused a 0.2331 increase in GPA.

However, when one controls for the SAT scores of the entering class, the effect shrinks substantially. USNWR's new same-year effect is .0041 and the next-year USNWR effect is negative .0018. This indicates a total effect of .0023 for every percentage increase in weight, or a .0483 increase in GPA due to the policy as a whole.

These findings also show that USNWR rankings, which come out in September or early October, have an effect on that fall semester. In other words, it does not take much time for the information to be absorbed and acted on – perhaps because USNWR tends to let university officials know in advance about major changes in the formula.

Table 3: Results[15]

Variable	**Simple Model Coefficient (p-value)**	**Model with SAT Control Coefficient (p-value)**
USNWR weight	0.008895 (.001)***	0.0040791 (0.053)*
USNWR (lagged 1 year)	0.0022484 (.337)	-0.0018334 (.453)
SAT score	--	0.000176 (.161)[16]
SAT score (lagged 1 year)	--	0.0001421 (.256)
SAT score (lagged 2 years)	--	0.0000279 (.875)
SAT score (lagged 3 years)	--	0.000335 (.005)**

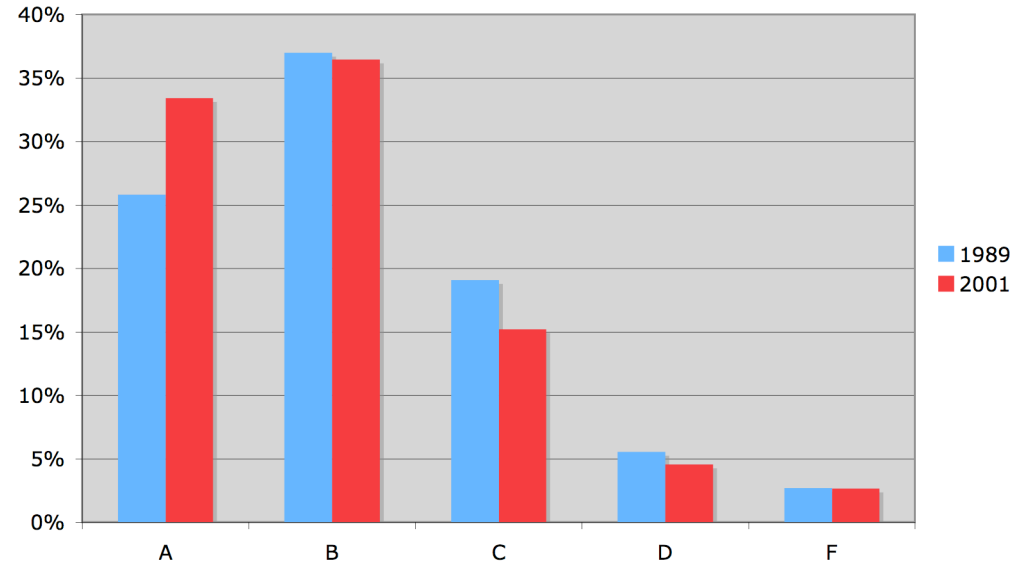
[15] * Valid at the 90% confidence level. ** Valid at the 95% confidence level *** Valid at the 99% confidence level

[16] Fortunately for the College Board, the signs on the SAT controls are all positive and an f-test reveals that they are jointly significant at the 99% confidence level. (P-value of .0093.) A 20 point increase in the average SAT score of an incoming class tends to increase average GPA by .0268 over four years.

It is also interesting to note that, as shown in Figure 2 on page 4, community colleges – which are unranked in USNWR – have hardly experienced grade inflation. There are many possible reasons for that, but it is in line with the finding that USNWR rankings encourage grade inflation. Unfortunately, my data on community colleges is too limited for statistical tests.

The data also indicate that the main increase in GPA comes at the top of the grade distribution. As shown in figure three, the proportion of A's awarded rose by nearly 30% at the 5 schools looked at between 1989 and 2001, while D's and C's each became 20% less common. Despite the small sample of five institutions, the shifts in A's is large enough to be statistically significant from zero at a 90% confidence level.[17]

Figure 3: Change in grade distributions, 1989 to 2001



Using the fixed-effects model described above to determine the difference between the effect on graduate and undergraduate GPAs also confirms the implications of the grade

[17] Calculated using a two-tailed t-test. This assumes a random selection of universities and that changes in the proportion of grades awarded are normally distributed. The changes in the other grades awarded are not statistically significant.

distribution data. The sample here is likely too small to say anything meaningful after controlling for SAT scores (which significantly reduces the degrees of freedom.) When one does not, however, the effect is actually slightly higher than that reported for undergraduate GPAs, indicating a shift primarily in the higher grades.

Table 4: Effect of average GPA of graduating students

Variable	**Simple Model Coefficient (p-value)**	**Model with SAT Control Coefficient (p-value)**
USNWR weight	.0131273 (.002)***	-.0020689 (.849)
USNWR (lagged 1 year)	-.0032364 (.475)	.0080739 (.504)
SAT score	--	-.0009964 (.216)
SAT score (lagged 1 year)	--	.0022615 (.057)*
SAT score (lagged 2 years)	--	.0029107 (.034) **
SAT score (lagged 3 years)	--	.0006547 (.242)

IV. Conclusions

I look at GPA data and changes in the weight given to graduation rates in USNWR between 1988 and 2003 to determine if one causes the other. After controlling for trends and SAT scores, I estimate that USNWR's 21% graduation rate component raised the average grade awarded by .048 between 1989 and 1996. Furthermore, a change in the weight of the graduation rate component causes GPA to rise in the same year, indicating a quick response on the part of university administrators.

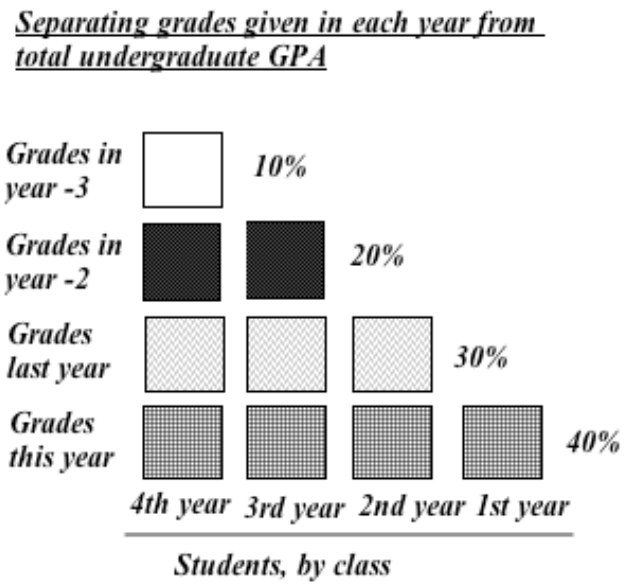
Examining the changes in the distribution of grades reveals that most of the increase has come in the higher levels of the distribution. This is inconsistent with the idea that professors are intentionally inflating grades to boost the graduation rate, but works with the hypothesis that administrators become less hostile to grade inflation because of a desire to do well in USNWR rankings. Regression analysis on the GPA of graduating students also backs up that result.

My findings suggest that USNWR's measure is causing perverse incentives among university administrators. It also indicates that the measure, which is intended to register improved services at universities, is at least to some degree registering relaxed standards instead.

Appendix A: Converting undergraduate GPAs to grades awarded in the current year.

As noted on pages 5 and 6, one problem with a variable for the average GPA for all undergraduates is that it would take 4 years for any inflationary policy to be fully realized in the variable. What we really want to know is the average grade given out in any particular year. I calculate the new variable by assuming that new grades determine 40 percent of the undergraduate GPA every year and solving the resulting simultaneous equations.

The graphic below shows how we can think about new grades, and why we can expect 40 percent of the undergraduate GPA in the current year to be determined by grades awarded in the same year. Each year you go back, the proportion falls by 10 percent, as one of the classes who got grades in that year leaves.



Note that this assumes no dropouts for simplicity. A more complex approach might try to use retention rates to come up with exact proportions. My approach will likely slightly underestimate the magnitude of the change in each year, because assuming there are no dropouts gives too little weight to grades in the current year. However, this would not systematically bias the estimate in either direction – it only decreases the magnitude of changes, which just makes it harder to find correlations when running regressions.

If we further assume that the GPA in the first year of our data applies equally to all classes in that year, we can set up a series of simultaneous equations and solve for average grade awarded in each year.[18]

[18] According to Stuart Rojstaczer's data, grades were relatively flat in the mid- 80's, indicating that this may be a reasonable assumption.

Notation

GPA_y is the undergraduate GPA in the current year. A negative number after the "y" means it is the undergraduate GPA of that many years ago.

Y is the average grade given out in the current year. Negative subscripts indicate previous years.

System of equations, given our assumptions

$$GPA_y = .4Y + .3Y_{-1} + .2Y_{-2} + .2Y_{-3}$$
$$GPA_{y-1} = .4Y_{-1} + .3Y_{-2} + .4Y_{-3}$$
$$GPA_{y-2} = .4Y_{-2} + .6Y_{-3}$$
$$GPA_{y-3} = Y_{-3}$$

Solving

Note: Calculations were checked by making up a GPA series that followed our assumptions and making sure that the formula correctly predicted grades awarded in any particular year.

Solve for Y_{-2}:

$$.4Y_{-2} = GPA_{y-2} - .6 * GPA_{y-3}$$
$$Y_{-2} = (10/4) * (GPA_{y-2} - .6 * GPA_{y-3})$$

Plug in Y_{-2} to Y_{-1} equation and solve for Y_{-1}:

$$GPA_{y-1} = .4Y_{-1} + (3/4) * GPA_{y-2} - (18/40) * GPA_{y-3} + (12/40) * GPA_{y-3}$$
$$.4Y_{-1} = GPA_{y-1} - (3/4) * GPA_{y-2} + (3/20) * GPA_{y-3}$$
$$Y_{-1} = (5/2) * GPA_{y-1} - (15/8) * GPA_{y-2} + (3/8) * GPA_{y-3}$$

Plug Y_{-1}, Y_{-2}, Y_{-3} into GPA_y equation and solve for Y. We end up with:

(3) $Y = (GPA_y - (3/4)GPA_{y-1} + (9/16) * GPA_{y-2} - (1/80) * GPA_{y-3}) / .4$

Formula 3 was then used to calculate grades awarded in each year past the third year (1991.)

Appendix B: Institutions that make up the data set

Appalachian State University
Auburn University
Bucknell University
CSU – San Bernardino
California State University – Fullerton
Carleton
Central Michigan
Centre College
College of William and Mary
Community College of Philadelphia
Cornell
Duke University
Elon
Furman
Georgia Tech
Hampden-Sydney College
Harvard
Hope College
Kenyon
Knox College
Los Angeles Mission College
Messiah College
Middlebury College
Missouri
Montana State University
Northwestern University
Penn State University
Pomona
Princeton
Purdue
Reed College
Roanoke
San Jose State University
Texas A&M
Texas State University
U. Texas – Austin
UC Irvine
UCLA
UNC Chapel Hill
University of Central Florida
University of Georgia
University of Hawaii – Hilo
University of Hawaii – Manoa
University of Michigan - Ann Arbor
University of Michigan – Flint
University of North Carolina - Asheville
University of Northern Iowa
University of Utah
University of Washington
University of Wisconsin – La Crosse
University of Wisconsin – Milwaukee
Wake Forest
Wheaton
Williams
Winthrop University
Wisconsin – Madison

Institutions for which I have detailed distribution data

University of Delaware
DePauw University
Victoria College
Georgia Institute of Technology
Roanoke College

Discussion of SAT score data

There were three problems to deal with before the SAT data given in USNWR rankings could be used.

The first is that the College Board "recentered" scores in 1995 to get the average back up to 1,000. For the average school in our sample, the corresponding 1994 SAT

score would have been 920, so I subtracted 80 points from the scores after 1994 to make the data consistent.

Another issue is that USNWR would often alternate between reporting ACT and SAT scores for midwestern schools in an inconsistent manner. Using ACT to SAT conversion charts yeilded implausible changes in scores between years, probably because most midwestern schools had many students without SAT scores to report. Thus when they were forced by USNWR to provide an SAT average, it was for only a part of their college population.

I attempt to deal with this issue by assuming no change in the SAT scores at an institution during years when USNWR switched to reporting ACT scores for them. If ACT scores changed in years after that, I estimated the corresponding change in SAT scores (using an ACT to SAT conversion chart) and adjusted the data accordingly. This allows us to pick up changes during years when ACT data was reported, while still being consistent with the actual SAT data that we have.

A final thing to note is that USNWR reported mean SAT scores from 1989 – 1994, when it switched to reporting just the 25^{th} and 75^{th} percentiles. Although it is an imperfect estimation, I averaged the 25^{th} and 75^{th} percentile scores to estimate the means.