



Received: 19 October 2015
Accepted: 27 November 2015
Published: 25 January 2016

*Corresponding author: John Robst,
Department of Mental Health Law &
Policy/Department of Economics, Florida
Mental Health Institute, University of
South Florida, 13301 Bruce B Downs
Blvd., Tampa, FL 33612, USA
E-mail: jrobst@usf.edu

Reviewing editor:
Caroline Elliott, Huddersfield University,
UK

Additional information is available at
the end of the article

GENERAL & APPLIED ECONOMICS | RESEARCH ARTICLE

The relationship between faculty characteristics and the use of norm- and criteria-based grading

John Robst^{1*} and Jennifer VanGilder²

Abstract: Norm-based grading has been associated with a reduction in student incentives to learn. Thus, it is important to understand faculty incentives for using norm-based grading. This paper used two waves of the National Study of Postsecondary Faculty to examine faculty characteristics related to the use of norm-based grading. Results suggest that norm-based grading is more likely when faculty and departments are more research oriented. Faculty who are at lower rank, male, younger, in science and social science departments are more likely to use norm-based grading, while faculty who feel that teaching should be the primary promotion criterion use criteria-based grading.

Subjects: Economics; Educational Research; Higher Education; Microeconomics

Keywords: grading standards; value of time; student effort

1. Introduction

Numerous studies have examined the assessment of student learning and the incentives for students to learn. The most common form of assessment is the assignment of grades by teachers with studies finding that more rigorous assessment (i.e. assigning low grades) has positive effects on student learning. Grades can be assigned using criterion-based measures or norm-based measures (Glaser, 1963). Criterion-based grading depends on whether the student demonstrates a given level of knowledge. Norm-based grading depends on the student's relative performance compared to other members of a pre-specified group (Glaser, 1963). The group may be a nationally representative

ABOUT THE AUTHORS

John Robst is a research associate professor with the Department of Mental Health Law and Policy and Department of Economics at the University of South Florida, and a research associate with the Institute for Study of Labor (IZA) in Bonn, Germany. His research interests include issues related to health economics and the economics of education, with a recent focus on treatment and outcomes for youth with severe emotional disturbance.

Jennifer VanGilder is Associate Professor of Business and Economics at Ursinus College. Jennifer's research focuses on facial symmetry and skin tone discrimination and has appeared in numerous journal including *Economic Letters*, *International Journal of Sport Finance*, and the *Journal of Sports Economics*. This research received national attention in numerous media outlets including the *WSJ*, *Inside the NFL*, and was featured during an episode of *NFL Films Presents*.

PUBLIC INTEREST STATEMENT

Education policy is often interested in student learning and factors that influence student effort. The method used to determine grades may influence how hard students work. For example, some teachers curve grades, while others do not. Curves can lead to a reduction in student effort because students only need to know more than other students in the class and not necessarily give their maximum effort. This paper sought to understand why some faculty at colleges and universities use curves. We found that faculty are more likely to use curves when they face greater incentives and pressure to perform research. This suggests that curves, in addition to reducing student learning effort, may be associated with a reduction in faculty teaching effort as well. In order to reduce the use of curves, colleges and universities should place greater emphasis on teaching in tenure and promotion decisions.

group (e.g. for scoring IQ tests) or simply be other members of a class (e.g. when curving grades in a class). The use of curves is one type of norm-based or relative grading because performance is evaluated compared to the rest of the class (Aviles, 2001; Becker & Rosen, 1992). Norm-based grading rewards students that do better than other students in the class, while criterion-based grading rewards students that achieve a level of performance regardless of how well (or poorly) other students perform. How that level of performance is defined varies considerably across higher education institutions and departments (Sadler, 2005).

Norm-based grading can be operationalized in different ways. For example, IQ tests typically compare student performance to a national average. In such cases, all scores within a specific class could conceivably be above or below the norm score of 100. In other cases, a desired grade distribution within a class may be pre-determined and grades curved to fit the distribution. Thus, grades or outcomes can depend on the selection of the “norm.” However, it is also possible for a curve to simply adjust criterion-based grading by giving students extra points without final grades fitting any pre-determined distribution. While not fitting a specific grade distribution, this would still result in relative grades not specifically linked to absolute performance or knowledge learned. Thus, as Aviles (2001) notes, “norm-referenced grading is ordinarily called grading on the curve” and despite some definitional differences, we treat norm-based grading and the use of curves as interchangeable throughout the manuscript.

A number of studies have argued that norm-based grading reduces student effort (e.g. Becker & Rosen, 1992; Betts, 1995; Bishop, 1996, 1997, 1999; Covington, 2000; Guskey, 2000; Landeras, 2009). Covington (2000) argued that norm-based grades reduce incentives for students to work and learn since the use of norm-based grades limits student opportunities to earn high grades. When the opportunity for high grades is limited, self-worth is maintained through a reduction in effort. In other words, when grades are curved, less able students will not be incentivized to work for high grades, because they are at a competitive disadvantage to more able students. When all students have the ability to earn high grades, then all students have an incentive to put forth effort to attain the highest grade. Betts (1995) claimed that the use of curves was akin to a tournament model. Students only have to outperform other students in the class. As a result, the incentives for smarter students in the class are reduced since their best work is not necessary to get a good grade. Bishop (1999) argued that norm-based grades reduce average student effort because students can receive the same grade with less effort. As a result, peer pressure is applied to limit overall effort by students. Thus, while Covington (2000) and Betts (1995) suggested, the use of norm-based grades affect incentives at different parts of the ability spectrum, and Bishop (1999) focused on average effort, each argues that incentives are negatively affected by the use of curves.

However, given that the curving of grades can reduce student incentives to learn, it is important to understand the incentives for faculty to use norm-based grading. This paper used two waves of the National Study of Postsecondary Faculty (1993 and 1999) to examine the use of norm-based grading and determine how their use varied based on faculty characteristics and attitudes about teaching, and institutional characteristics. Results suggest that grades are curved when faculty are more research-oriented, while faculty who are teaching-oriented tend to use criterion-based grading. The results suggest that institutions looking to enhance learning and teaching effort might increase the importance of teaching quality in tenure and promotion decisions.

2. Background

This section presents a basic model, based on the work of Mason, Steagall, and Fabritius (2003), to examine why faculty may use norm- or criterion-based grading. Faculty utility is a positive function of student knowledge acquired (K) and student grades (G), and a negative function of time spent on teaching activities (T):

$$F = U(K,G,T)$$

However, unlike Mason et al. (2003), we view teaching effort based on the distribution of time between teaching and research. Most people would prefer to spend less time on teaching effort while controlling for the amount of knowledge gained by the student. However, preferences regarding the distribution of time are likely to vary considerably across faculty. For example, a faculty member that feels teaching is the primary component of his or her job may prefer spending most of their time on teaching. A faculty member at a research institution, particularly junior faculty, may prefer spending a greater proportion of their time on research.

The use of curves is expected to affect faculty utility because norm-based grading can affect student knowledge and faculty work activities. Prior research has argued that norm-based grading reduces student incentives to learn. As such, the use of norm-based grading would seem to reduce faculty utility. However, if we assume that the use of norm-based grading enables faculty members to spend more time on research (this assumption will be examined in more detail later in the paper), then faculty members that place a greater value on research are expected to maximize utility by the use of norm-based grading. Faculty members that place a greater emphasis on teaching are expected to maximize utility by using criterion-based grading.

The use of norm-based grading may or may not affect faculty utility through student grades. Faculty members that use norm-based grading may give difficult tests, and thus a curve may be necessary to bring the grade distribution to a level the faculty member views as acceptable. Similarly, a faculty member that uses criterion-based grading may give easy tests, and have very high grades. An alternative view is that norm-based grading can be used to inflate grades. In other words, a curve may represent a reduction in grading standards. Thus, while some faculty may use norm-based grading as a mechanism to give high grades, there is not a clear association between the use of norm-based grading and grades, and we do not assume such a relationship.

While the relationship between the use of norm-based grading and grading standards is not clear, prior research has argued that norm-based grading reduces average student effort because students can receive the *same* grade with less effort (Bishop, 1999). The ability to receive the same grade with less effort suggests that the use of norm-based grading can be viewed as a lowering of grading standards. Faculty may reduce grading standards in an effort to receive better student evaluations (Krautmann & Sander, 1999; McPherson, Jewell, & Kim, 2009). By giving higher grades, faculty members continue to receive positive student evaluations regardless of how much students learn. Student evaluations are particularly important in tenure and promotion decisions at some institutions, suggesting that junior faculty who need better student evaluations have a greater incentive to reduce grading standards.

The use of norm-based grading is expected to vary across departments as well. Freeman (1999) examined the divergence in grades across departments. Professors may attempt to attract more students through lowering grading standards so the department receives additional funding. Freeman (1999) argued that higher grading departments tend to have lower potential incomes and thus use grades to attract students, while lower grading departments are in fields that offer higher potential incomes. Winzer (2002) suggested many of these factors and also included the role that large class sizes, and multiple teaching, service, and scholarly commitments may play in reduced grading standards.

Finally, the institutional environment may also be a factor in the decision to use relative and criterion-referenced grading. Warning and Welzel (2005) developed a model that examines institutional incentives to promote a reduction in grading standards. For example, with a duopoly with two public universities, both have an incentive to inflate grades with higher government subsidies leading to greater grade inflation. In the case where a public university is competing with a private university, the public university has an incentive to inflate grades to increase enrollment and state funding. While not discussed in their paper, the private institution may still have incentives to reduce grading standards if such actions lead to an increase in their applicant pool. Schools have an incentive to encourage applications, even if they don't reduce admission standards, so the school can report a low acceptance rate.

Overall, the model and prior research suggest that grading on a curve is more likely when faculty have an incentive to spend more time on research and/or reduce grading standards.

3. Methodology

3.1. Data

The data used for this study are from the 1993 and 1999 National Study of Postsecondary faculty. The NSOPF provides a profile of faculty in the US, what they do, and why many aspects of the profession have changed. The sample is restricted to faculty at institutions from the Carnegie classifications of Research, Doctoral, Comprehensive, and Liberal Arts.

The 1993 and 1999 NSOPF collected information on the use of norm-based grading. For example, the 1999 NSOPF survey asked the following question: “In how many of the undergraduate courses that you taught for credit during the 1997 Fall Term did you use: Grading on a curve?” The survey question has three possible responses for grading on a curve: all, some, or none. For those responding they use curves some of the time, the data do not enable us to examine factors that would differentiate between when they do and do not use a curve. Consequently, the “all” and “some” responses are combined. The NSOPF is available for 2004 as well, but the survey does not contain questions on the use of norm-based grading.

The survey also collected information on faculty demographic characteristics, academic rank, and academic department. Institutional characteristics such as whether the institution is public or private, the student–faculty ratio and Carnegie classification are also available. Two survey questions asked whether the respondent agreed that teaching should be the primary criterion for promotion, and that research should be the primary criterion for promotion. Respondents could strongly agree, agree, disagree, or strongly disagree. Strongly agree and agree are combined, as are disagree and strongly disagree to create a dichotomous variable. Finally, the survey collected information on whether the faculty member is satisfied with their salary and with their job.

3.2. Analytical methods

For this study, a binary logistic regression is used to examine factors associated with the use of norm-based grading. The specification estimates the probability of using norm-based grading as a function of faculty member, department, and institutional characteristics:

$$\Pr(\text{Curve}_{ij}) = X_{ij} \cdot \beta + \mu_{ij} \quad (1)$$

where X includes faculty demographic characteristics (age, gender, and race), other faculty characteristics (job satisfaction, whether research/teaching should be primary criterion for promotion, and rank: instructor, assistant professor, associate professor, reference: full professor), department (science, social science, other; reference is humanities), and institutional characteristics (private vs. public; student–faculty ratio; Carnegie classification: doctoral/research, reference is comprehensive/liberal arts).¹ Regressions are estimated for the total sample, and for research-oriented (research/doctoral) and teaching-oriented (comprehensive, liberal arts) institutions separately. Marginal effects are computed for each variable with all other variables held at their mean.

4. Results

4.1. Descriptive statistics

The sample contains 10,776 faculty members from the 1993 survey and 8,438 from the 1999 survey. Descriptive statistics are provided in Table 1. Twenty-six percent of respondents report grading on a curve with 12% using norm-based grading in all classes and 14% in some classes. The sample is 62% male and 83% white with an average age of 48 years old. The majority of faculty members are satisfied with their job, although only about half are satisfied with their salary. Seventy-three percent

Table 1. Descriptive statistics

	Mean	Std. Dev.
Graded using a curve	.257	.437
<i>Faculty demographics</i>		
Age	48.4	9.5
Gender—Male	.623	.485
Race—White	.827	.378
<i>Other faculty characteristics</i>		
Satisfied with job	.825	.380
Satisfied with salary	.519	.500
Promotion criterion: teaching	.734	.442
Promotion criterion: research	.359	.479
Rank: Instructor/lecturer	.081	.272
Rank: Assistant	.298	.457
Rank: Associate	.286	.452
<i>Department</i>		
Science	.226	.418
Social Science	.369	.483
Humanities	.203	.402
Other	.201	.401
<i>Institution</i>		
Private	.356	.479
Student-faculty ratio	14.8	6.7
Carnegie: Research/doctoral	.482	.499
Observations	19,214	

agree that teaching should be the primary criterion for promotion, and 36% feel that research should be the primary criterion. These are separate questions, thus some respondents agreed with both criteria.

One assumption discussed earlier in the paper is that faculty using norm-based grading dedicate less time to teaching and more time to research. The NSOPF data contain information on faculty time allocation between teaching, research, and service activities. Faculty that use curves in all classes spend 57.5% of their time on teaching activities compared to 62.8% for faculty that do not use curves.² On the other hand, faculty that use norm-based grading in all classes spend 21.5% of their time on research activities compared to 16.4% for faculty that do not use norm-based grading. Both differences are significant at the $p < .05$ level. One possible way to reduce the time spent on grading is to give multiple choice examinations. Indeed, 37% of faculty that curved grades in all classes use only multiple choice exams compared to 25% of faculty that do not use curves. Thus, while the data do not enable a detailed analysis of the time allocation between teaching and research, the data suggest that such a difference exists.

4.2. Associations between faculty characteristics and research preferences

Given the model discussed above, the associations between using norm-based grading and many of the independent variables can be predicted. Faculty members that place a greater emphasis on research are expected to use norm-based grading, while faculty that place a greater emphasis on teaching are less likely to use curves. Thus, for each variable we examined the proportion of time actually spent on research and the proportion of time the faculty member prefers to spend on

research. It is important to examine preferences because actual time spent on research may be constrained by teaching and service responsibilities.

Among the demographic characteristics (age, gender, and race), men report spending a greater percentage of their time on research than women (22.7 vs. 17.9%, $t = 8.40$) and prefer to spend a greater percentage of their time on research (27.5 vs. 22.0%, $t = 10.5$). Thus, men place a greater emphasis on research, and are expected to be more likely to use norm-based grading. This is consistent with research that has found men publish more than women (Ginther & Kahn, 2004), and women tend to be rated as slightly better teachers than men (Feldman, 1993). Similarly, faculty ages 35–44 report spending a greater percentage of their time on research than faculty ages 55–64 (22.8 vs. 19.9%, $t = 4.09$) and prefer to spend a greater percentage of their time on research (27.3 vs. 24.4%, $t = 4.92$). Thus, younger faculty members are expected to be more likely to use curves. Racial differences are also evident as non-white faculty report spending a greater percentage of their time on research than white faculty (20.4 vs. 24.5%, $t = 3.82$) and prefer to spend a greater percentage of their time on research (28.8 vs. 24.9%, $t = 4.04$). Thus, non-white faculty members are expected to be more likely to use norm-based grading.

Among other faculty characteristics are whether research/teaching should be primary criterion for promotion, rank: instructor, assistant professor, associate professor (reference: full professor), and job satisfaction. As would be expected, faculty members that feel research should be the primary criteria for promotion spend more time on research (29.4 vs. 16.0%, $t = 23.0$) and prefer to spend more time on research (35.5 vs. 20.1%, $t = 28.5$). Faculty members that feel teaching should be the primary criteria for promotion spend less time on research (16.8 vs. 31.9%, $t = 23.3$) and prefer to spend more time on research (21.0 vs. 38.8%, $t = 33.1$). Thus, faculty members that place a greater emphasis on research in tenure decisions are expected to use norm-based grading, while faculty that place a greater emphasis on teaching quality in tenure decisions are less likely to use norm-based grading. The relationship between rank and research focus is somewhat complex. Junior faculty members who need publications for tenure may be more research focused. On the other hand, full professors often achieve their rank because of their research focus. Indeed, the data suggest the full professors spend more time on research (23.3 vs. 21.4%, $t = 2.43$) and prefer to spend more time of research (30.1 vs. 28.0%, $t = 2.87$). Job satisfaction, while not necessarily linked to an emphasis on research or teaching, is expected to be associated with overall work effort. Consequently, we control for job satisfaction because faculty members that are more satisfied with their job are expected to put forth more effort and are thus less likely to use norm-based grading.

Research emphasis is also expected to vary based on department (science, social science, other; reference is humanities), and institutional characteristics (private vs. public; student–faculty ratio; Carnegie classification: doctoral/research, reference is comprehensive/liberal arts). Faculty in science and social science departments tend to be more research focused than faculty in humanities departments. Faculty in science departments spend 24.9% of their time on research and prefer to spend 28.8% of their time on research activities compared to 16.5 and 22.8% in humanities departments. Faculty in social science departments spend 19.5% of their time on research and prefer to spend 24.8% of their time on research. Thus, faculty in science and social science departments are expected to use curves more often than faculty in humanities departments. Faculty in public institutions report spending marginally more time on research (21.2 vs. 20.1%, $t = 1.84$), but do not prefer to spend more time on research (25.4 vs. 25.2%, $t = .35$) than faculty at private institutions. Faculty at institutions with student–faculty ratios below the mean (14.5) spend more time on research than faculty at institutions with higher student–faculty ratios (24.3 vs. 16.6%, $t = 9.29$) and also prefer to spend more time on research (28.6 vs. 21.4%, $t = 10.4$). However, the research effect may be partly (or completely) offset by the efforts to minimize the additional teaching effort that comes with larger student–faculty ratios. Thus, faculty at institutions with large student–faculty ratios may grade using norm-based grading as a means to manage a large teaching load. As would be expected, faculty at research/doctoral institutions spend more time on research (28.4 vs. 14.0%, $t = 23.8$) and prefer to spend more time on research (33.4 vs. 20.5%, $t = 23.4$) than faculty at comprehensive/liberal arts institutions. Thus, faculty members at research institutions are expected to be more likely to use curves.

4.3. Logistic regression results

The logistic regression results and estimated marginal effects are provided in Table 2. The likelihood ratios indicate the models have the necessary predictive power to examine which faculty use curves. Among the demographic variables, male faculty members are 7.4% points more likely to use norm-based grading than female faculty, and norm-based grading is more common among younger faculty. The results for gender and age are consistent with the hypothesized relationship between the use of curves and the desire to spend a greater proportion of time on research. The results suggest that men tend to place more effort on research, and place a smaller emphasis on research. The results for age suggest that younger faculty also place greater emphasis on research, although the marginal effect is rather small. A 20-year increase in age would only result in a 2% point decline in the use of curves. As noted earlier, age is associated with the preferred time spent on research. However, it remains possible that the results for age may indicate an overall reduction in work effort (both research and teaching) as faculty become older. Race was not associated with the likelihood of using curves.

Among the other faculty characteristics, the faculty member's opinions regarding promotion criteria are associated with norm-based grading, with those agreeing that teaching should be the

Table 2. Logistic regression results

	Total sample				Research/doctoral institutions				Comprehensive/liberal arts institutions			
	Coef.	Std. Err.	p value	Marg. Eff	Coef.	Std. Err.	p value	Marg. Eff	Coef.	Std. Err.	p value	Marg. Eff
Intercept	-1.057	.157	<.001	-	-1.538	.238	<.001	-	-.935	.207	<.001	-
Year 1993	-.185	.035	<.001	-.034	-.331	.052	<.001	-.057	-.092	.048	.055	-.018
<i>Faculty demographics</i>												
Age	-.006	.002	.009	-.001	-.005	.003	.123	-.001	-.007	.003	.015	-.001
Male	.415	.039	<.001	.074	.401	.060	<.001	.067	.448	.051	<.001	.084
White	.003	.045	.951	.001	.040	.067	.552	.007	-.028	.061	.644	-.005
<i>Job satisfaction/opinions</i>												
Satisfied job	-.166	.046	<.001	-.031	-.219	.068	.001	-.039	-.116	.062	.063	-.022
Satisfied salary	-.104	.036	.004	-.019	-.176	.054	.001	-.031	-.042	.048	.385	-.007
Teaching primary	-.106	.042	.011	-.019	-.100	.053	.058	-.017	-.139	.072	.053	-.027
Research primary	.017	.038	.646	.003	-.001	.052	.989	-.0001	.042	.055	.446	.008
<i>Rank</i>												
Instructor	.027	.077	.726	.005	.126	.124	.310	.022	-.043	.099	.669	-.008
Assistant Prof	.138	.053	.009	.025	.123	.082	.135	.022	.147	.069	.034	.028
Associate Prof	.145	.046	.002	.027	.115	.068	.092	.020	.160	.062	.010	.031
<i>Department</i>												
Science	.826	.052	<.001	.167	.911	.082	<.001	.175	.779	.069	<.001	.164
Social Science	.280	.049	<.001	.052	.376	.080	<.001	.066	.229	.061	<.001	.044
Other depts.	-.242	.060	<.001	-.042	-.284	.091	.002	-.048	-.088	.084	.292	-.017
<i>Institutional characteristics</i>												
Private institution	-.039	.039	.305	-.007	.021	.061	.732	.003	-.054	.051	.286	-.010
Student-faculty ratio	.008	.003	.002	.001	.030	.005	<.001	.004	.0002	.003	.955	.0001
Research school	-.239	.040	<.001	-.044	-	-	-	-	-	-	-	-
Observations	19,214				9,226				9,988			
Likelihood ratio	862.7		<.001		548.0		<.001		352.8		<.001	

Note: Dependent variable: Graded using a curve.

primary criterion for promotion being 1.9% points less likely to curve grades. Once again, faculty that place a greater emphasis on teaching are less likely to use curves. Rank is also associated with the use of curves as assistant professors and associate professors are 2.5 and 2.7% points more likely to curve grades than full professors. Assistant and associate professors place a greater emphasis on research given that most institutions require research and publications for promotion. However, similar to the effect of age, the results for rank may indicate an overall reduction in work effort once faculty become full professors. Workers satisfied with their job are expected to put forth more effort in their job. Consequently, faculty members who are more satisfied report being 3.1% points less likely to use norm-based grading. Job satisfaction is not necessarily associated with research emphasis, and is essentially used as an attempt to control for overall worker effort. Thus, the findings for most worker characteristics are consistent with expectations.

Compared to humanities departments, faculty in science and social science departments are 16.7 and 5.2% points more likely to use curves. It is challenging to make conclusions regarding emphasis across departments. For example, it is conceivable that faculty in science and social science departments spend more time on research activities and place less emphasis on teaching. Alternatively, it is also possible that faculty in certain departments tend to give more difficult tests, and a curve is necessary to achieve a desired grade distribution; or that humanities courses tend to require an absolute level of performance for grading. Faculty members at institutions with larger student-faculty ratios are more likely to use norm-based grading. Faculty members at institutions with a larger student-faculty ratio may use curves to manage their teaching work load. There was no significant difference between public and private institutions in the use of curves. One exception to the association between norm-based grading and research emphasis is the finding that faculty at institutions in the doctoral/research Carnegie classifications are 4.4% points less likely to use norm-based grading than faculty at comprehensive/liberal arts institutions. This finding may reflect the different course mix at doctoral/research institutions (e.g. graduate level classes), student quality, and institutional standards and policies.

In order to investigate this result further, the sample was divided based on whether the institution is a research/doctoral and comprehensive/liberal arts institution. Most of the associations found for the overall sample hold for faculty at research and teaching institutions. Among research/doctoral institutions, faculty are more likely to use curves if male, an associate professor (compared to full professor), in a science or social science department, less satisfied with their job, or employed at an institution with a higher student-faculty ratio. Among comprehensive or liberal arts institutions, faculty are more likely to use curves if younger, male, an assistant or associate professor (compared to full professor), in a science or social science department, or less satisfied with their job. Thus, within an institutional type, grading on a curve is more common among faculty members who have an incentive or simply a desire to focus on research.

5. Conclusion

The use of norm-based grading has been relatively unexplored in the economics literature. What little research exists has focused on the effect of relative grading on student incentives. Given prior research has argued that the use of norm-based grading is associated with a reduction in student effort, it is important to understand faculty incentives for using curves. This paper examined the relationship between using norm-based grading and faculty and institutional characteristics. In general, we found faculty members who are focused more on research are more likely to use norm-based grading.

Overall, the relationships between faculty characteristics, departments, and the probability of grading on a curve suggest that grades were more likely to be curved when faculty and departments were more research oriented. For example, younger faculty and males might be considered more research oriented as they tend to publish more than older faculty and females. Assistant and associate professors were more likely to grade on a curve (compared to full professors) as were faculty in social science and science departments (compared to humanities). Norm-based grading was less

likely among faculty who feel that teaching should be the primary criterion for advancement. Workers satisfied with their job were expected to put forth more effort in their job.

The results have several policy implications. A number of studies have assessed the link between norm-based grading and student effort. The majority of research suggests that student effort may decline when grades are curved. Thus, in order to maximize student effort and student learning, there should be an emphasis on criterion-based grading. Such grading assesses absolute student performance and knowledge instead of relative performance, and ensures that students passing a class have learned the class material. However, to accomplish this goal, the trade-off between teaching and research effort must be addressed. In order to increase such teaching effort, higher education institutions should increase the role of teaching in tenure and promotion decisions. Rewarding teaching excellence to a greater degree would increase the relative benefit to teaching effort. Faculty more satisfied with their job are less likely to report using norm-based grading. Thus, worker satisfaction may lead to increased teaching effort, and a reduction in the use of curves. Higher student-faculty ratios are also associated with grading on a curve. Thus, institutions with higher student-faculty ratios may promote greater teaching effort by reducing class sizes. Obviously, enhancing worker satisfaction and reducing student-faculty ratios can have significant financial costs, and thus institutions must weigh the benefit to potential increases in teaching effort with the costs.

There are several limitations to this study. The use of norm-based grading is self-reported. There may be a stigma associated with grading on a curve because of the potential for being seen as having soft standards. Thus, the use of norm-based grading may be under-reported. Alternatively, some faculty may manipulate grades but feel as though their grading system does not fall under the terminology “grading on a curve.” While some of the discussion in this manuscript assumes that grading on a curve represents a reduction in grading standards, we are unable to explicitly test this assumption with the data. It is also possible that norm-based grades are simply a consequence of giving challenging exams. Finally, despite the use of two years of data, the study is essentially a cross-sectional study, and it is difficult to assign any causation to the results. It is possible that there are unobserved effects that may be biasing the results. Using the multiple years of data to establish a longitudinal study was not practical because faculty identifiers were not available. However, given the relative consistency of the results, such unobserved effects would need to be strongly correlated with most of the variables in the specification in order to change the conclusions of the study.

Funding

The authors received no direct funding for this research.

Author details

John Robst¹

E-mail: jrobst@usf.edu

Jennifer VanGilder²

E-mail: jvangilder@ursinus.edu

¹ Department of Mental Health Law & Policy/Department of Economics, Florida Mental Health Institute, University of South Florida, 13301 Bruce B Downs Blvd., Tampa, FL 33612, USA.

² Department of Business and Economics, Ursinus College, Collegeville, PA 19426, USA.

Citation information

Cite this article as: The relationship between faculty characteristics and the use of norm- and criteria-based grading, John Robst & Jennifer VanGilder, *Cogent Economics & Finance* (2016), 4: 1127746.

Notes

1. The decision to create a dichotomous-dependent variable by combining the responses of “used norm-based all of the time” and “used norm-based some of the time” may be questioned by some readers. We also ran the regressions as ordered logit regressions to determine whether estimating the regression with three

categories would matter to the results. The results were qualitatively similar to those reported in the paper and to conserve space was not reported.

2. Descriptive statistics on time spent on teaching, time spent on research, and publications are from the DAS-T Online Version 5.0 website maintained by the National Center for Education Statistics.

References

- Aviles, C. B. (2001). Grading with norm-referenced or criterion-referenced measurements: To curve or not to curve, that is the question. *Social Work Education*, 20, 603–608. <http://dx.doi.org/10.1080/02615470120072869>
- Becker, W. E., & Rosen, S. (1992). The learning effect of assessment and evaluation in high school. *Economics of Education Review*, 11, 107–118. [http://dx.doi.org/10.1016/0272-7757\(92\)90002-K](http://dx.doi.org/10.1016/0272-7757(92)90002-K)
- Betts, J. (1995). *Do grading standards affect the incentive to learn?* (Working Paper). San Diego, CA: University of California.
- Bishop, J. (1996). The impact of curriculum-based external examinations on school priorities and student learning. *International Journal of Education Research*, 23, 653–752.
- Bishop, J. (1997). The effect of national standards and curriculum-based exams on achievement. *American Economic Review*, 87, 260–264.
- Bishop, J. (1999). Are national exit examinations important for educational efficiency? *Swedish Economic Policy Review*, 6, 349–398.

- Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology*, 51, 171–200.
<http://dx.doi.org/10.1146/annurev.psych.51.1.171>
- Feldman, K. A. (1993). College students' views of male and female college teachers: Evidence of students' evaluations of their classroom teachers. *Research in Higher Education*, 34, 151–211.
<http://dx.doi.org/10.1007/BF00992161>
- Freeman, D. (1999). Grade divergence as a market outcome. *Journal of Economic Education*, 30, 344–351.
<http://dx.doi.org/10.1080/00220489909596091>
- Ginther, D. K., & Kahn, S. (2004). Women in economics: Moving up or falling off the academic career ladder? *Journal of Economic Perspectives*, 18, 193–214.
<http://dx.doi.org/10.1257/0895330042162386>
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes. *American Psychologist*, 18, 510–522.
- Guskey, T. R. (2000). Grading policies that work against standards ... and how to fix them. *NASSP Bulletin*, 84, 20–29. <http://dx.doi.org/10.1177/019263650008462003>
- Krautmann, A. C., & Sander, W. (1999). Grades and student evaluations of teachers. *Economics of Education Review*, 18, 59–63. [http://dx.doi.org/10.1016/S0272-7757\(98\)00004-1](http://dx.doi.org/10.1016/S0272-7757(98)00004-1)
- Landeras, P. (2009). Student effort: Standards vs. tournaments. *Applied Economics Letters*, 16, 965–969.
<http://dx.doi.org/10.1080/13504850701222129>
- Mason, P. M., Steagall, J. W., & Fabritius, M. M. (2003). The changing quality of business education. *Economics of Education Review*, 22, 603–609.
[http://dx.doi.org/10.1016/S0272-7757\(03\)00025-6](http://dx.doi.org/10.1016/S0272-7757(03)00025-6)
- McPherson, M. A., Jewell, R. T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35, 37–51.
<http://dx.doi.org/10.1057/palgrave.eej.9050042>
- Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education*, 30, 175–194.
<http://dx.doi.org/10.1080/0260293042000264262>
- Warning, S., & Welzel, P. (2005). *A note on grade inflation and university competition*. Paper presented at the EARIE 2005 conference. Retrieved from http://www.fep.up.pt/conferences/earie2005/cd_rom/Session%20III/III.L/warning.pdf
- Winzer, M. (2002). *Grade inflation: An appraisal of the research*. University of Lethbridge. Retrieved July 8, 2008, from <http://people.uleth.ca/~runte/inflation/Inflationb.pdf>



© 2016 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format
Adapt — remix, transform, and build upon the material for any purpose, even commercially.
The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.
You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



Cogent Economics & Finance (ISSN: 2332-2039) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

