Putting Grades in Context

Author(s): Talia Bar, Vrinda Kadiyali, and Asaf Zussman

Source: *Journal of Labor Economics*, Vol. 30, No. 2 (April 2012), pp. 445-478
Published by: The University of Chicago Press on behalf of the Society of Labor Economists and the NORC at the University of Chicago
Stable URL: http://www.jstor.org/stable/10.1086/663591
Accessed: 28/08/2014 11:36

Putting Grades in Context

Talia Bar, *Cornell University*

Vrinda Kadiyali, *Cornell University*

Asaf Zussman, *Hebrew University*

Concerns over grade inflation and disparities in grading practices have led institutions of higher education in the United States to adopt various grading reforms. An element common to several reforms is providing information on the distribution of grades in different courses. The main aims of such "grades in context" policies are to make grades more informative to transcript readers and to curb grade inflation. We provide a simple model to demonstrate that such policies can have complex effects on patterns of student course enrollment. These effects may lower the informativeness of some transcripts, increase the average grade, and lower welfare.

I. Introduction

A. Motivation

Grade inflation and disparities in grading practices are widespread phenomena in higher education and have attracted significant attention inside and outside academia.[1] A primary concern with these phenomena is that they make grades less informative. Transcript readers (e.g., potential em-

We are grateful to Peter Arcidiacono, Haim Bar, Steve Coate, Ani Guerdjikova, and Mukul Majumdar for useful comments and to Vidya Atal and Yanlei Ma for excellent research assistance. Contact the corresponding author, Talia Bar, at tb97@cornell.edu.

[1] For examples of press coverage, see Bruno (2007), Bartlett and Wasley (2008), Primack (2008), Wasley (2008), Harford (2009), Lewin (2010), and Foderaro (2011). We discuss the relevant academic literature below.

[*Journal of Labor Economics*, 2012, vol. 30, no. 2]

0734-306X/2012/3002-0006$10.00

ployers and graduate schools) might have less information on which to base their decisions (e.g., hiring and admission), and students might not obtain reliable signals regarding their own strengths. Thus, grade inflation and disparities in grading practices are associated with potential distortions.

Concerns over these problems have led institutions of higher education in the United States to reform their grading practices. Two approaches have been taken: "grade rationing" (or "forced curves") and "putting grades in context." Grade rationing restricts instructors' choice of grading policies. For example, Princeton University limits A grades to an average of 35% across departments, and New York University's Law School has a detailed mandatory grading curve. The second approach, putting grades in context, is to provide information on the distribution of grades in different courses with the aim of improving the reliability of grades as signals of student quality. Proponents of the second approach suggest that putting grades in context may result in a weaker incentive for students to select leniently graded courses, which in turn may assist in curbing grade inflation. Some proponents of this approach also expect that faculty would be more reluctant to grade leniently if they knew that this information would be exposed to their colleagues and to readers of transcripts. At the first glance, it may appear that the two approaches to reform grading—imposing grade distributions and grades in context policies—would have the same consequences, but in fact their effects can be quite different. For example, forcing grading curves essentially eliminates grade inflation and controls grading leniency, but as our article demonstrates, putting grades in context might result in an increase in grades.

The goal of this article is to study the potential effects of grades in context policies. We offer a simple stylized model that illustrates that these policies can have unintended consequences. Importantly, we show that the provision of information about grading policies could result in an *increase* in average grades and in a *decrease* in the reliability of information on student ability conveyed to transcript readers. The key to these results is that the provision of information about grading policies affects the way students select courses.

To understand the impact of grades in context reforms, it is important to know who has access to information about grading policies. We say that students are informed if they have access to the information before selecting courses; we say that transcript readers (from now on simply "employers") are informed if they have access to the information before hiring. We consider four regimes: (1) neither students nor employers are informed, (2) only employers are informed, (3) only students are informed, and (4) students and employers are informed.

Consider some examples of grades in context policies. Indiana University provides online information on the number of students in each grade category in each course. Similarly, since 1998, Cornell University has posted online reports providing information on enrollment and median grades in different courses. Analysis conducted by Bar, Kadiyali, and Zussman (2009) suggests that Cornell students used these reports to select courses.[2] If such online information is not as easily accessed by potential employers (since it requires sifting through reports for many courses and candidates), the regime in which only students are informed might best describe this situation. At other universities, information about grading policies is included in students' transcripts. This makes it easier for employers to judge student quality. Columbia University transcripts report the percentage of A-range grades for certain classes. Median grades are reported both online and in the transcripts of Dartmouth College students.

We show that students' course selection and the information content of grades (from the perspective of employers) depend on the information regime. When information on grading policies is provided only to students, some of them become more attracted to leniently graded courses. When the information is provided to both students and employers, some students have an increased incentive to choose strictly graded courses, where the same grade provides a better signal. Other students have an increased incentive to choose leniently graded courses, where they can receive a higher grade and be pooled with higher-ability students. Changes in students' course selection patterns may lead to an increase in the average grade across courses, contributing to grade inflation, and can result in grades becoming a less reliable measure of student ability. Before turning to the model, we discuss related literature.

B. Related Literature

Our study is related to a growing literature on grade inflation and on disparities in grading practices. For surveys of this literature, see Rosovsky and Hartley (2002) and Johnson (2003). We first describe a few theoretical contributions and then discuss the empirical literature. Ostrovsky and Schwarz (2010) explore information disclosure by universities in matching markets. Students are ranked by their ability and jobs are ranked by their desirability to students. This results in a unique stable assortative match. Strategically introducing some noise to the transcript may allow schools to increase placement into moderately desirable positions. Yang and Yip

[2] We note that there are additional resources for students to obtain (likely less reliable) information on professors' grading policies. These include Internet sites such as Ratemyprofessor.com and CourseRank—where students report on grading in courses—and informal word-of-mouth networks among students in a university. Recently, CourseRank has attempted to use Freedom of Information Act measures to obtain grades from public universities; this is expected to vastly improve its coverage and reliability.

(2003) model grade inflation that results from adverse selection in the labor market. Their model involves matching between two types of students and two types of jobs. Only high-ability students are productive in the high-type job. Grade inflation is defined by the share of low-ability students who are assigned a high grade. In their model, wages depend on job type but cannot depend on the student's school. This results in a free-rider problem, where each school has a tendency to inflate grades, lowering the reputation of all students. The analysis demonstrates that grade inflation reduces productivity and welfare. Chan, Li, and Suen (2007) provide a signaling model of grade inflation. In their model, employers know the distribution of grades but not the distribution of student abilities. When a school gives a lot of high grades, employers cannot distinguish between lenient grading standards and a large proportion of high-ability students.

Our analysis shares some common features with these theoretical contributions. For example, in our model and in the previous studies, employers rely on grades as signals of students' abilities since the latter are not directly observed. Additionally, in all models the focus is on the information content of grades rather than on the phenomenon of a continuous rise in grades.

An important difference, however, is that all the previous studies have a single grading policy for the school, whereas we allow for within-school variation in grading policies. Students may take grading policies into account when choosing courses. In this respect our model is closer to that of Rosar and Schulte (2010). They present a model with an agent who is only imperfectly informed about his type and a designer who can construct a device (a test) that, if used, directly reveals information about the agent's type. A key feature in their model is that the agent can opt out of taking the test. Similarly, in our model a student chooses which course (test) to take. Rosar and Schulte find that, even if the designer's goal is to learn the agent's type, the optimal device might generate imperfect information. Similarly, in our analysis we show that more information about grading policies is not necessarily better.

Within the empirical literature on university grading, the most relevant study is that by Bar et al. (2009), who analyze the effects of a Cornell University reform that provided online grade distribution information. The authors find that this grades in context reform led students to shop for leniently graded courses, an effect that was weaker for high-ability students. They also demonstrate that the change in enrollment patterns contributed to grade inflation. These findings are consistent with predictions that our simple model makes for the regime in which only students are informed.

Two key assumptions in our model are that there is within-school variation in the leniency of grading policies and that students strategically choose courses on the basis of this information (when it is available to them). There is empirical support for both assumptions. Recent studies have analyzed disparities in grading practices within a school. Achen and Courant (2009) document differences in average grades across fields of study at the University of Michigan. The authors attribute such differences to the elasticity of enrollment demand and to the cost for professors of assigning low grades, a cost that they argue likely depends on the availability of objective assessment methods. Bar and Zussman (forthcoming) show that differences in student grading outcomes across courses at an elite university in the United States are associated with the political orientation of faculty members. Students' grade-driven course selection has also been documented in several empirical studies. Fournier and Sass (2000) found that instructors' grading policies influenced students' subsequent curriculum choices. Johnson (2003) showed that students tend to choose courses offered by leniently grading instructors.

While previous studies have explored the causes and consequences of grade inflation and disparities in grading practices, we are not aware of any study that analyzes policies aimed at curbing grade inflation and reducing disparities. This article contributes to filling the void.

The rest of the article proceeds as follows. Section II introduces the model. Section III presents our core analysis and findings. Section IV examines welfare effects of putting grades in context policies. Section V reconsiders the behavior of low-ability students and accounts for the possibility that faculty may respond to grade reforms. Section VI offers concluding remarks. All proofs are provided in the appendix.

II. Model

In our model students can choose between two courses $i = 1, 2$.[3] Students differ in their tastes for courses and in their academic abilities. The courses are horizontally differentiated and located at the endpoints of a line segment $[0, 1]$ as in Hotelling (1929). Each student has a taste $\tau \in [0, 1]$. The distribution of tastes has a continuously differentiable cumulative distribution function $H(\tau)$ with a density $h(\tau) > 0$. We denote by $d(\tau, i)$ the distance of a student with taste τ from course i, $d(\tau, 1) = \tau$ and $d(\tau, 2) = 1 - \tau$. Students incur a cost (or a disutility) associated with choosing a course. The disutility from taking course i is given by $c(d(\tau, i))$. The function $c(d)$ is increasing and differentiable, and $c(0) = 0$. Each

[3] Analyzing only two options is standard in spatial models of product differentiation on which we build. One way to interpret these two options is as two sections in a single course; each section has a different focus to it, e.g., because of the interests of the instructor. Alternatively, the options may represent two elective courses within the same department or two programs in the same school. We believe that our model, while simple, illustrates important issues that would also apply in more complex course selection situations.

student has an ability $\theta \in [0, 1]$. A higher θ represents a higher ability. The distribution of students' abilities is given by a continuously differentiable cumulative distribution function $F(\theta)$, with a density $f(\theta) > 0$. The distributions of tastes and abilities are exogenous and commonly known. Taste and ability are independent.[4] Each student enrolls in one course.

There are two instructors, one in each course. A "grading policy" is defined by a threshold ability $\theta_i \in (0, 1)$ so that in course i grades are given by

$$g(\theta, i) = \begin{cases} \text{A} & \text{if } \theta \geq \theta_i \\ \text{B} & \text{if } \theta < \theta_i. \end{cases} \tag{1}$$

Grading policies differ by course and are taken as given. These policies can depend on the instructor's personal preferences, organizational incentives, and the nature of the subject being taught. Assuming exogenous grading policies allows us to focus on students' course selection and to illustrate the complex effects of grades in context reforms. In Section V we discuss the possibility that instructors respond to such reforms.

Employers cannot observe students' abilities or tastes directly. Instead, they rely on observed grades (and course choice, depending on the regime) to convey this information. Employers form rational expectations about the abilities of graduating students based on the information contained in transcripts. Higher ability is associated with higher productivity and therefore also with a higher wage. We therefore assume that the utility of a student of ability θ and taste τ enrolled in course i and who is awarded a grade g is given by

$$u(\theta, \tau; g, i) = E(\tilde{\theta} | g \text{ in } i) - c(d(\tau, i)). \tag{2}$$

The term $c(d(\tau, i))$ is the student's disutility from taking course i. The term $E(\tilde{\theta} | g \text{ in } i)$ is the expected ability of students with a grade g in course i.[5] The conditional distribution of abilities used in this expectation depends on the information available, on the grading policies, and on the equilibrium course selection of all students. This will be made clearer when we discuss each of the information regimes below.

[4] This simplifying assumption seems plausible in the context of course selection within a major. For example, there is no reason to assume that students who are more interested in ancient Greece than in ancient Rome have higher (or lower) ability than those who have the opposite preference. One might expect some correlation between taste and ability in the choice of discipline. For example, it is possible that low-ability students are relatively more attracted to the humanities and high-ability students are more attracted to the sciences, which may explain in part observed differences in the distribution of Scholastic Aptitude Test scores between disciplines.

[5] Figures 1 and 2 below illustrate regions of student types (θ, τ) with each grade in each course in two of the information regimes that will be considered later.

III. Putting Grades in Context

We examine equilibrium course selection patterns under four distinct grading information provision regimes: (1) no information, (2) employers only, (3) students only, and (4) students and employers. In the first two regimes, students do not know professors' grading policies. In reality, students might have some limited information gleaned from peers who have previously taken the class. In the absence of official information on grading policies, however, students will be less informed than when official information is provided. The uninformed students assumption is an idealization designed to capture in a simple way the relative change from less information to more information.[6] The first regime, which we use as a benchmark for comparison, approximates a school that has no grades in context policy (and there are no other reliable sources of information on grading distributions). The second regime approximates a school that provides official grade distribution information to employers but not to students.[7]

In the third and fourth regimes, information about grading policies is provided to students, who may take it into account in selecting courses. In the third regime, the information is provided only to students. As argued above, this could represent a situation in which grading information is available online. In the fourth regime, information about grading policies is provided to both students and employers. For example, the information may be available online as well as in transcripts.

A student's utility from taking a course and receiving a certain grade depends on his expected ability, which in turn is determined by the information regime. When employers are uninformed about within-school differences in grading policies, a student's expected ability depends on his grade but not on the course he took. When employers are informed about within-school differences in grading policies, a student's expected ability depends both on his grade and on the course he took. An equilibrium in our model is defined as follows.

Definition 1 (Equilibrium). $\sigma : [0, 1]^2 \rightarrow \{1, 2\}$ is an equilibrium

[6] In an empirical study, Bar, Kadiyali, and Zussman (2005) examined how course enrollment was related to the course's lagged median grade before and after Cornell University started publishing course median grades online. Course enrollment was found to be independent of the lagged median grade before the policy change but positively associated with it after the change. This is consistent with our argument that the provision of official grade information can make a crucial difference.

[7] We include this information regime for completeness. In practice, this regime might be less plausible than the others. Universities that provide grading information to employers typically also provide it to students. The information either is available online (where students can access it) or is included in the transcript, in which case students might obtain the official information from peers.

course selection mapping if (i) for all (θ, τ), the course $\sigma(\theta, \tau)$ maximizes the student's utility (2) and (ii) given the course selection mapping σ and the information regime, employers' expectations about a student's ability are correct.

For uninformed employers, expectations about a student's ability are correct conditional on the student's grade. For informed employers, expectations about a student's ability are correct conditional on the student's grade and on the course chosen.

A student's grade depends on his ability and on the grading policy. All students of ability θ in the same course expect the same grade. If a student of type (θ', τ') strictly prefers course 1, then all other students with the same ability θ' and with a taste parameter $\tau < \tau'$ will also prefer course 1. An equilibrium course selection mapping is therefore characterized by a boundary curve $\boldsymbol{\tau}(\theta)$ so that $\sigma(\theta, \tau) = 1$ if $\tau < \boldsymbol{\tau}(\theta)$ and $\sigma(\theta, \tau) = 2$ if $\tau > \boldsymbol{\tau}(\theta)$. A student of ability θ and taste $\tau = \boldsymbol{\tau}(\theta) \in (0, 1)$ is indifferent between the courses, and if $\boldsymbol{\tau}(\theta) = 1$ or 0, the student chooses course 1 or course 2, respectively.

A. Uninformed Students

Our premise is that unless otherwise informed, students and employers treat the two courses symmetrically.

Definition 2. Students and employers "treat courses symmetrically" if (i) students and employers have identical and independent prior beliefs about the leniency of grading in each course, (ii) uninformed employers' expectations depend only on grades, and (iii) informed employers' expectations depend only on the observed policies and the student grade; that is, for any grade g and grading policies θ', θ'', employers' expectations are such that

$$E(\tilde{\theta} | g \text{ in } 1)|_{(\theta'', \theta')} = E(\tilde{\theta} | g \text{ in } 2)|_{(\theta', \theta'')}.$$

When the two courses are treated symmetrically and students are uninformed, students have the same expectations for how they would be perceived by employers in either course. As a result the following proposition holds.

Proposition 1 (Uninformed students). If courses are treated symmetrically, then in equilibrium in regimes 1 and 2, students choose courses solely according to their tastes, $\boldsymbol{\tau}(\theta) \equiv \frac{1}{2}$.

Corollary 1 immediately follows.

Corollary 1. Enrollment into the leniently graded course and the average grade across courses are the same in the employers-only regime as in the no-information regime.

In Section III.C, we explain how employers' information about student ability is different in regimes 1 and 2.

B. Informed Students

We now consider the two information regimes with informed students for situations in which the grading policies in the two courses are different, $\theta_1 \neq \theta_2$. Without loss of generality, assume that course 1 is the leniently graded course. From now on we refer to it as course L and to course 2, the strictly graded course, as course S. The grading policies partition students into three ranges of ability: high ($\theta \geq \theta_S$), intermediate ($\theta_L \leq \theta < \theta_S$), and low ($\theta < \theta_L$). Students in the same ability range and in the same course are perceived by employers to have the same expected ability. Therefore, the boundary $\tau(\theta)$ of equilibrium course selection is characterized by three threshold taste levels (τ_1^*, τ_2^*, τ_3^*):

$$\tau(\theta) = \begin{cases} \tau_1^* & \text{if } \theta < \theta_L \\ \tau_2^* & \text{if } \theta_L \leq \theta < \theta_S \\ \tau_3^* & \text{if } \theta \geq \theta_S. \end{cases}$$

Students of type (θ', τ') such that $\tau' < \tau(\theta')$ enroll in course L. The equilibrium enrollment into course L is

$$N_L(\tau_1^*, \tau_2^*, \tau_3^*) = H(\tau_1^*)F(\theta_L) + H(\tau_2^*)[F(\theta_S) - F(\theta_L)] + H(\tau_3^*)[1 - F(\theta_S)]. \tag{3}$$

In the following subsections we characterize the thresholds that define an equilibrium in regimes 3 and 4.

1. Informed Students, Uninformed Employers

Consider the students-only regime. Employers evaluate students on the basis of their grade but not their course selection. The behavior of students in this regime also fits a situation in which employers (or other outside readers of the transcript who consider grades for scholarships, grants, graduation requirements, and so on) have an "absolute" standard, for example, to hire students with a grade point average (GPA) of 3.5 or more.

According to the grading policy defined in (1), low- ($\theta < \theta_L$) and high- ($\theta \geq \theta_S$) ability students will have the same grade (B and A, respectively) in either course and, hence, the same expected ability. Therefore, these students choose a course solely on the basis of their tastes, $\tau_1^* = \tau_3^* = \frac{1}{2}$.[8]

Intermediate-ability students, $\theta_L \leq \theta < \theta_S$, receive a B in course S and an A in course L. Students with grade A are of high or intermediate ability and students with grade B are of intermediate or low ability. Therefore,

[8] Grading practices might sometimes be such that low-ability students expect a lower grade in the strictly graded course than in the leniently graded course. In this case, low-ability students will be more attracted to the leniently graded course. In Sec. V.A, we provide such an example.

$E(\tilde{\theta}|\text{A}) > E(\tilde{\theta}|\text{B})$. Thus, intermediate-ability students are associated with a higher expected ability if they enroll in the leniently graded course. If their taste parameter is $\tau < \frac{1}{2}$, they clearly prefer course L. If their taste parameter is $\tau > \frac{1}{2}$, they face a trade-off between their taste for course S and the higher expected ability in course L. In equilibrium, the threshold taste for intermediate-ability students is $\tau_2^* = \bar{\tau} > \frac{1}{2}$, where $\bar{\tau}$ satisfies

$$E(\tilde{\theta}|\text{A}) - c(\bar{\tau}) = E(\tilde{\theta}|\text{B}) - c(1 - \bar{\tau}) \quad \text{and} \quad \bar{\tau} < 1 \tag{4}$$

or

$$E(\tilde{\theta}|\text{A}) - c(1) \geq E(\tilde{\theta}|\text{B}) \quad \text{and} \quad \bar{\tau} = 1.$$

When $\bar{\tau} < 1$, a student of intermediate ability and taste $\bar{\tau}$ is indifferent between taking course L and course S. It is possible, however, that all intermediate-ability students prefer course L, in which case $\bar{\tau} = 1$. We summarize these findings in proposition 2.

Proposition 2 (Informed students). In equilibrium in regime 3, low- and high-ability students choose according to their tastes, $\tau_1^* = \tau_3^* = \frac{1}{2}$; intermediate-ability students are attracted to the leniently graded course, $\tau_2^* = \bar{\tau} > \frac{1}{2}$.

The results of proposition 2 are illustrated in figure 1. Corollary 2 follows from proposition 2.

Corollary 2. Enrollment into the leniently graded course and the average grade across courses are higher in the students-only regime than in the no-information regime.

These predictions are consistent with the empirical findings of the Bar et al. (2009) study of Cornell's grades in context reform.

2. *Informed Students and Employers*

In the students and employers regime, employers' expectations about student ability depend on the student's grade as well as on his course choice. In course L, only low-ability students obtain a B. In course S, low- and intermediate-ability students obtain a B. Therefore, the conditional expected abilities satisfy $E(\tilde{\theta}|\text{B in } S) \geq E(\tilde{\theta}|\text{B in } L)$. Any student with $\theta < \theta_L$ and $\tau > \frac{1}{2}$ will clearly prefer course S because, for this student, course S results in at least as high an expected ability and a lower taste disutility. In equilibrium, $\tau_1^* = \bar{\tau}_1 \leq \frac{1}{2}$. It is possible that in equilibrium all low-ability students take course S, so $\bar{\tau}_1 = 0$. When $\bar{\tau}_1 > 0$, a student with $\theta < \theta_L$ and taste $\bar{\tau}_1$ is indifferent between the two courses. In equilibrium,

$$E(\tilde{\theta}|\text{B in } L) - c(\bar{\tau}_1) = E(\tilde{\theta}|\text{B in } S) - c(1 - \bar{\tau}_1) \quad \text{and} \quad \bar{\tau}_1 > 0 \tag{5}$$

or

$$E(\tilde{\theta}|\text{B in } L) \leq E(\tilde{\theta}|\text{B in } S) - c(1) \quad \text{and} \quad \bar{\tau}_1 = 0.$$

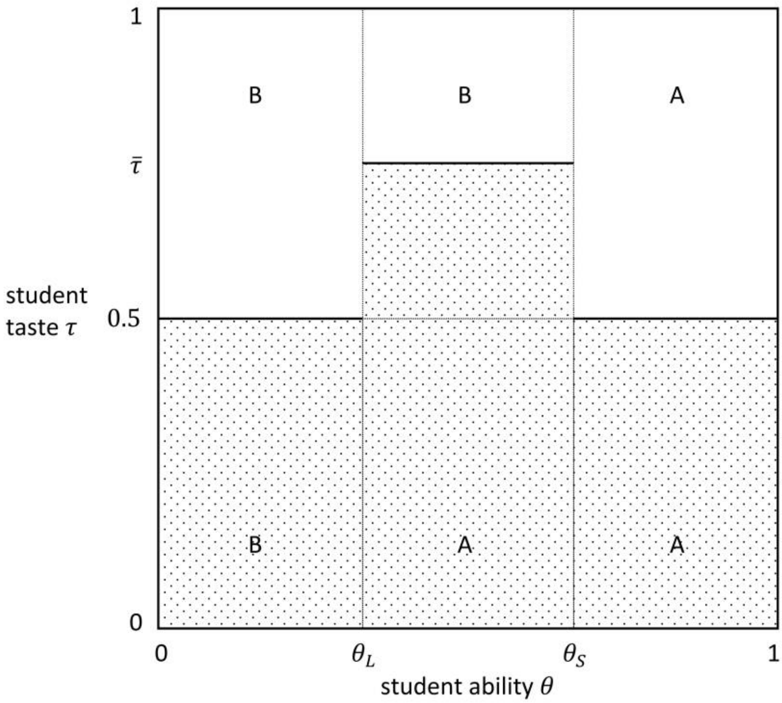

FIG. 1.—Course selection when only students are informed. The dotted areas represent students enrolled in the leniently graded course; all other students are enrolled in the strictly graded course; the letters A and B represent grades awarded to students in the given ability-taste ranges.

Similarly, high-ability students ($\theta \geq \theta_S$) can expect to be awarded a grade of A regardless of which course they choose. Because in the strictly graded course only high-ability students obtain an A but in the leniently graded course high- and intermediate-ability students obtain this grade, $E(\tilde{\theta}|\text{A in } S) > E(\tilde{\theta}|\text{A in } L)$.[9] This implies that $\tau_3^* = \bar{\tau}_3 < \frac{1}{2}$. When $\bar{\tau}_3 > 0$, a student with ability $\theta \geq \theta_S$ and taste $\bar{\tau}_3$ is indifferent between the two courses:

$$E(\tilde{\theta}|\text{A in } L) - c(\bar{\tau}_3) = E(\tilde{\theta}|\text{A in } S) - c(1 - \bar{\tau}_3) \quad \text{and} \quad \bar{\tau}_3 > 0 \tag{6}$$

or

[9] Here the inequality is strict because at least some intermediate-ability students will choose course L. This is argued more formally in the proof of proposition 3 in the appendix.

$$E(\tilde{\theta}|\text{A in } L) \leq E(\tilde{\theta}|\text{A in } S) - c(1) \quad \text{and} \quad \bar{\tau}_3 = 0.$$

An intermediate-ability student with $\theta_L \leq \theta < \theta_S$ is awarded an A in course L and a B in course S. The expected ability of an intermediate-ability student is higher in the leniently graded course, $E(\tilde{\theta}|\text{A in } L) > E(\tilde{\theta}|\text{B in } S)$, because $E(\tilde{\theta}|\text{A in } L)$ is a weighted average of high- and intermediate-ability students and $E(\tilde{\theta}|\text{B in } S)$ is a weighted average of intermediate- and low-ability students. Therefore, in equilibrium, $\tau_2^* = \bar{\tau}_2 > \frac{1}{2}$. When $\bar{\tau}_2 < 1$, a student with ability $\theta_L \leq \theta < \theta_S$ and taste $\bar{\tau}_2$ is indifferent between the two courses:

$$E(\tilde{\theta}|\text{A in } L) - c(\bar{\tau}_2) = E(\tilde{\theta}|\text{B in } S) - c(1 - \bar{\tau}_2) \quad \text{and} \quad \bar{\tau}_2 < 1 \tag{7}$$

or

$$E(\tilde{\theta}|\text{A in } L) - c(1) \geq E(\tilde{\theta}|\text{B in } S) \quad \text{and} \quad \bar{\tau}_2 = 1.$$

The expected abilities in (5)–(7) are functions of the equilibrium thresholds $\bar{\tau}_i$. We write them explicitly in the proof of proposition 3.

Proposition 3 (Informed students and employers). In equilibrium in regime 4, low- and high-ability students are attracted to the strictly graded course, $\bar{\tau}_1 \leq \frac{1}{2}$ and $\bar{\tau}_3 < \frac{1}{2}$. Intermediate-ability students are attracted to the leniently graded course, $\bar{\tau}_2 > \frac{1}{2}$.

Figure 2 illustrates our findings in this proposition. Intuitively, once information on grading policies is available, students who expect the same grade in both courses (low- and high-ability students) are more attracted to the strictly graded course, where the grade signals a higher expected ability. In contrast, students who obtain a higher grade in the leniently graded course (intermediate-ability students) are more attracted to it because they will be pooled with higher-ability students.

In selecting courses, students take into account not only grades but also their tastes. Students who have a strong taste for a certain course will remain in that course even if they will have a lower grade. When the courses offered are similar, however, course selection might be determined solely by grades. We think of courses as being similar (not too differentiated) when the maximal disutility that a student can get from choosing a nonideal course, $c(1)$, is small relative to what the student can gain in terms of expected ability.

Definition 3. We say that courses are "similar" if

$$\begin{aligned} 0 < c(1) < \min\{&E(\tilde{\theta}|\tilde{\theta} \geq \theta_S) - E(\tilde{\theta}|\tilde{\theta} \geq \theta_L), \\ &E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S)) - E(\tilde{\theta}|\tilde{\theta} < \theta_S)\}. \end{aligned} \tag{8}$$

For example, for a uniform distribution of abilities, the courses are similar when

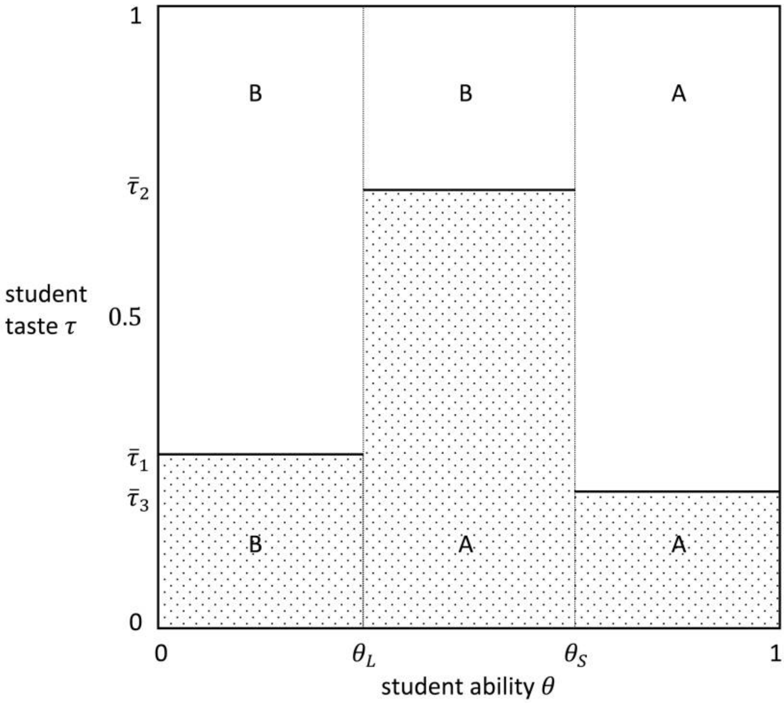

FIG. 2.—Course selection when both students and employers are informed. The dotted areas represent students enrolled in the leniently graded course; all other students are enrolled in the strictly graded course; the letters A and B represent grades awarded to students in the given ability-taste ranges.

$$0 < c(1) < \min\left\{\frac{\theta_S - \theta_L}{2}, \frac{\theta_L}{2}\right\}.$$

When the courses are similar, students select the course in which they would be perceived to have the highest ability because their taste preference for a particular course is not very strong.

LEMMA 1. If courses are similar, then all high-ability students choose course S ($\bar{\tau}_3 = 0$), all intermediate-ability students choose course L ($\bar{\tau}_2 = 1$), and low-ability students choose according to their tastes ($\bar{\tau}_1 = \frac{1}{2}$).

The effect on enrollment of providing information to both students and employers is ambiguous because enrollment into the leniently graded course decreases for some students and increases for others. For similar courses, we show that enrollment into course L is more likely to increase with the provision of information when the difference between grading policies is large (resulting in a large volume of intermediate-ability students who prefer course L) and when course S has a very strict policy (resulting in a lower volume of high-ability students who prefer course S).

Corollary 3. If courses are similar, then enrollment into the leniently graded course is larger in the students and employers regime than in the no-information regime if and only if

$$[1 - H(\tfrac{1}{2})][F(\theta_S) - F(\theta_L)] > H(\tfrac{1}{2})[1 - F(\theta_S)].$$

Whether or not enrollment into course L increases, the average grade across the two courses increases when putting grades in context. Increased enrollment of intermediate-ability students into the leniently graded course (where they receive a higher grade) contributes to an increase in the average grade. The decrease in enrollment of low- and high-ability students into the leniently graded course does not affect the average grade because these students receive the same grade in either course. Hence, overall, the changed pattern of course selection leads to a higher average grade.

Corollary 4. The average grade across courses is higher in the students and employers regime than in the no-information regime.

We also compare the students and employers regime to the students-only regime. Our findings in propositions 2 and 3 suggest that high- and low-ability students are more likely to enroll in the leniently graded course in the students-only regime than in the students and employers regime ($\bar{\tau}_1 \leq \frac{1}{2}$ and $\bar{\tau}_3 < \frac{1}{2}$). Therefore, enrollment into the leniently graded course is unambiguously higher in the students-only regime if $\bar{\tau}_2 \leq \bar{\tau}$. We can show that this holds if there is a unique equilibrium in the students-only regime. Lemma 2 in the appendix provides sufficient conditions for a unique equilibrium. The intuition why $\bar{\tau}_2 \leq \bar{\tau}$ is that intermediate-ability students have more to gain from enrolling in the leniently graded course (obtaining the grade A instead of B) when employers are uninformed than when employers are informed.

Proposition 4. Suppose that there is a unique equilibrium in the students-only regime. In this case, intermediate-ability students are more attracted to the leniently graded course in the students-only regime than in the students and employers regime, $\bar{\tau} \geq \bar{\tau}_2$. Thus, enrollment into the leniently graded course is higher and the average grade is at least as high in the students-only regime.

C. Information Quality

1. *The Variance of Student Abilities*

We now ask what effect putting grades in context has on the quality (or accuracy) of the information employers obtain about students' abilities. We use the variance of conditional distributions as a measure of information quality. For tractability, results in this subsection are derived for uniform distributions of students' tastes and abilities. This special case illustrates the ambiguity of the effects.

We first show that compared to the regimes in which employers are not informed about grading policies, when employers are informed, they have better information (lower variance of student abilities) on students with grade A in course S because course S separates the high-ability students from the rest.[10]

For students with grade A in course L, we show that the variance of students' abilities can increase with the provision of information on grading policies to employers. To show this effect, we note that when the difference between grading policies vanishes, $\theta_S \to \theta_L$, or when hardly any student gets an A in course S, $\theta_S \to 1$, the variance approaches the same value in all regimes. Around these points, a small increase in θ_S (i) lowers the variance of abilities of students with grade A in the no-information regime, (ii) has no effect on the variance of abilities of students with an A in course L in the employers-only regime (as it depends only on θ_L), and (iii) increases the variance of abilities of students with an A in course L in the students and employers regime. This allows us to compare the variances of student abilities in the different regimes for values of θ_S that are close to θ_L or close to one. We denote the variance by Var_r, with the subscript indicating the regime.

PROPOSITION 5. Assume a uniform distribution of student abilities and tastes. If employers are informed (regimes $r = 2$ and $r = 4$), then

i. information about high-ability students in course S is more accurate: $\mathrm{Var}_r(\tilde{\theta}|\mathrm{A\ in\ }S) \leq \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$;
ii. information about high-ability students in course L can be either less or more accurate: $\mathrm{Var}_r(\tilde{\theta}|\mathrm{A\ in\ }L) \geq \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$ when $\theta_S \to \theta_L$ but $\mathrm{Var}_r(\tilde{\theta}|\mathrm{A\ in\ }L) \leq \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$ when $\theta_S \to 1$;
iii. if the courses are similar (see definition 3), then in regime 4 information about high-ability students is more accurate: $\mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ }L) \leq \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$.

We also compare the quality of information between the students-only regime, the students and employers regime, and the no-information regime. We find that more information on grading policies can increase the variance of students' abilities when the grading policies are close.

PROPOSITION 6. Assume a uniform distribution of student abilities and tastes. If employers are informed, then

[10] This result also holds for other distributions, but it relies on the property that the variance of abilities conditional on $\theta \geq \theta_0$ is decreasing in θ_0, which does not hold for all distributions.

i. when $\theta_S \to 1$, we have $\mathrm{Var}_4(\tilde{\theta}|\mathrm{A} \text{ in } L) \leq \mathrm{Var}_3(\tilde{\theta}|\mathrm{A}) \leq \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$;
ii. when $\theta_S \to \theta_L$, we have $\mathrm{Var}_4(\tilde{\theta}|\mathrm{A} \text{ in } L) \geq \mathrm{Var}_3(\tilde{\theta}|\mathrm{A}) \geq \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$.

2. *Ranking Bias*

Within each course, grades are weakly monotone in student ability. When grading policies are not the same across courses, however, the ranking of students according to their expected ability might not be monotone in ability.

Definition 4. We say that there is a "ranking bias" when there are students with abilities $\theta_h > \theta_l$, so that student h is perceived to have a strictly lower expected ability (given his grade and course choice) than student l.

In all information regimes, for students in the intermediate-ability range $\theta_h, \theta_l \in [\theta_L, \theta_S)$, if $\tau_h > \tau_2^* > \tau_l$ so that h chooses course S, getting a B, and l chooses course L, getting an A, there is a ranking bias. When employers are uninformed, a ranking bias may occur only for students in the intermediate-ability range. In contrast, when employers are informed (regimes 2 and 4), a ranking bias may also occur in the high- and low-ability ranges. In these ranges the two students will have the same grade, but if $\tau_h < \tau_j^* < \tau_l$ ($j = 1, 3$), then employers perceive the ability of student h, who selected course L, as lower.

When employers are informed, providing students with information about grading policies allows some students to trade off their tastes for courses with improvements in expected ability. If the courses are similar, then by lemma 1, in the students and employers regime, all high-ability students select the strictly graded course and all intermediate-ability students select the leniently graded course. This eliminates ranking bias.

When courses are similar, we also showed in proposition 5 that in the students and employers regime, information about high-ability students is more accurate. We note that Cornell University's grades in context policy reports a combined median grade for different offerings of the same course (e.g., Intermediate Microeconomics taught in the same semester by different professors). Hence, the policy provides less grade distribution information for these courses that are likely similar. Our analysis suggests that perhaps for such courses putting grades in context can have more information benefits.

D. Grade Rationing

As mentioned in the introduction, two approaches have been adopted by different universities to curb grade inflation: grade rationing and grades in context policies. So far we have discussed grades in context policies; we now briefly turn our attention to grade rationing policies and compare the two approaches. Grade rationing policies directly impose grade distributions instead of only exposing this information. In the context of our model, we interpret grade rationing as imposing an equal policy $\theta_1 = \theta_2$ in both courses (or a constant and identical share of A grades per course). Under the assumptions of our model, the following proposition holds.

PROPOSITION 7. If policies $\theta_1 = \theta_2$ are imposed, then there is an equilibrium in which (i) students choose courses according to their tastes, (ii) the school can implement any average grade between B and A, and (iii) there is no ranking bias.

In our simple model, grade rationing seems to have some advantages over putting grades in context. Students select courses according to their tastes, grade inflation is curbed, and there is no ranking bias. However, grade rationing can also have complex effects on course selection and on the information content of grades. For example, if low-ability students are attracted to one course and high-ability students are attracted to the other, when an identical share of A's is imposed in each course, students would not necessarily choose according to tastes. The reason is that being identified as a top student in a class that attracts high-ability students would signal a higher expected ability than being identified as a top student in a class that attracts low-ability students.[11] A detailed analysis of grade rationing and its consequences is beyond the scope of our article.

IV. School Goals and Welfare

In this section we explore the welfare effects of grades in context reforms. These effects naturally depend on the form of the welfare criterion. Following Chan et al. (2007), we assume that in the labor market all workers are paid their productivities, so that (on average) employers are not affected by grade reporting policies. Consider first a utilitarian welfare function,

$$W_0 = \int_0^1 \int_0^1 u(\theta,\ \tau) dH(\tau) dF(\theta).$$

Each student's utility as defined in (2) is his expected ability as perceived by employers minus the disutility from course selection. Employers have rational expectations about student abilities. Therefore, the aggregate expected ability of students is the same in all four information regimes. The utility loss from course selection does, however, vary by regime. It is minimized when students are uninformed.

PROPOSITION 8. For the utilitarian welfare function W_0:

[11] Opponents of grade rationing sometimes also worry that the policy would discourage cooperation between students and that it does not allow grades to reflect a favorable draw of student abilities in a particular period.

i. Welfare is the same in the employers-only regime as in the no-information regime but is lower in the two regimes with informed students.
ii. If courses are similar, welfare is higher in the students-only regime than in the students and employers regime.

Chan et al. (2007) assume that schools care more about helping their good students (who, as alumni, may be more likely to donate money or bring fame to the school). Letting $R > 1$ be the relative weight on the utility of high-ability students, define an objective function

$$W_1 = W_0 + (R - 1)\int_{\theta_S}^{1}\int_{0}^{1} u(\theta, \tau)dH(\tau)dF(\theta).$$

PROPOSITION 9. For the objective function W_1, when compared to the no-information regime,

i. W_1 is lower in the students-only regime;
ii. W_1 is lower in the students and employers regime for R close to one, but it can be higher when R is large enough and $c(1)$ is sufficiently low.

The school's objective function has a lower value in the students-only regime because the expected ability of students with grade A is lower because of the increased selection of intermediate-ability students into the leniently graded course. In the students and employers regime, expected ability increases for high-ability students in course S but decreases for high-ability students in course L. When the disutility from taking nonideal courses is low enough, all high-ability students take the strictly graded course. They have a higher expected ability than they would have had in the no-information regime. For low enough disutility, high-ability students are better off in the students and employers regime. Hence, W_1 increases with information when the relative weight on the utility of high-ability students is high enough.

The objective functions considered above abstract from the possibility that matching of student ability to jobs is important (see Ostrovsky and Schwarz 2010). Depending on technology and market structure, it is possible that putting grades in context increases or decreases market efficiency. For example, suppose that the productivity of high-ability students is higher when matched to certain highly desirable jobs. Putting grades in context has a positive effect in that it allows identifying a subset of high-ability students (i.e., those with grade A in course S). On the other hand, high-ability students in course L will be pooled with a larger set of intermediate-ability students, reducing match quality for the former group.

V. Variations of the Model

A. Behavior of Low-Ability Students

In the model analyzed in the previous sections, low-ability students choose courses solely according to their tastes in the students-only regime. In the students and employers regime, in contrast, low-ability students are attracted to the strictly graded course, where the low grade is a better signal. However, if low-ability students expect a lower grade in the strictly graded course than in the leniently graded course, they may become more attracted to the leniently graded course. This can happen when grading policies include more than two grade categories (say A, B, and C) and students face some uncertainty about grades. We illustrate this with an example.

1. Example

Suppose that the two courses have the following grading policies: course S assigns grade A to high-ability students ($\theta \geq \theta_S$), grade B to intermediate-ability students ($\theta_L \leq \theta < \theta_S$), and grade C to low-ability students ($\theta < \theta_L$); course L assigns grade A to high-ability students, grade A or B with probabilities p_A and $1 - p_A$ to intermediate-ability students, and grade B or C with probabilities p_B and $1 - p_B$ to low-ability students.

In the students-only regime, high-ability students choose courses solely according to their tastes, $\tau_3^* = \frac{1}{2}$. Low- and intermediate-ability students are attracted to course L, $\tau_1^* \geq \frac{1}{2}$ and $\tau_2^* \geq \frac{1}{2}$, where they expect a higher grade. Depending on parameter values, it is possible that low-ability students are more attracted than intermediate-ability students to the leniently graded course ($\tau_1^* > \tau_2^*$).

In equilibrium, in the students and employers regime, low-ability students are attracted to course L, $\bar{\tau}_1 \geq \frac{1}{2}$, and high-ability students are attracted to course S, $\bar{\tau}_3 < \frac{1}{2}$. The effect on intermediate-ability students is ambiguous: $\bar{\tau}_2 \geq \frac{1}{2}$ when p_A is sufficiently high and $\bar{\tau}_2 \leq \frac{1}{2}$ when p_A is sufficiently low. These findings are shown in the appendix.

2. Discipline Choice—Discussion

In universities across the United States, grades tend to be lower in the natural sciences than in the humanities (Johnson 2003). Assuming that most students and employers are aware of this fact, the availability of grade distribution information most closely resembles the students and employers regime. A policy that provides official information about cross-discipline differences in grading distributions is therefore not likely to have a significant effect on discipline choice.

Our findings regarding the students and employers regime suggest that high-ability students will be attracted to the natural sciences. If grades of low-ability students tend to be higher in the humanities than in the natural sciences, as in the example above, then we also expect that low-ability students would be attracted to the humanities.

A caveat to the above discussion is that employers might be interested in specific majors. For example, investment banks might look for economics or business majors, ratings agencies might want political science majors for political risk assessment, and so forth. In such situations, within-major differences in grade distributions matter, offering room for students to shop for leniently graded courses.

Our analysis abstracted from the possibility that tastes and abilities are correlated. We note that the patterns of selection described for disciplines might alternatively (or additionally) be explained by a difference in tastes between students of different abilities. For example, if high-ability students enjoy natural science courses more than they enjoy humanities courses and vice versa for low-ability students, then we might see more low-ability students in the humanities even if these students are not attracted to leniently graded courses.

B. Faculty Response

Our analysis has so far assumed that grading policies are constant. It is possible, however, that faculty grading policies will respond to grade information reforms. This interaction can be modeled as a two-stage game. In stage 1, faculty choose their grading policies, and in stage 2, students select courses as described in the previous sections. We compare equilibrium choices of grading policies under the no-information regime and under the students-only and students and employers regimes.

Assume that instructors have ideal grading policies denoted by θ_i^*. All else being equal, the instructor in course L prefers a more lenient policy, $\theta_L^* < \theta_S^*$. The utility of each instructor may, however, also depend on other factors that are affected by the grading policies of both instructors. Specifically, suppose that instructors prefer higher enrollment.[12] This can be the case when instructors receive more resources or enjoy higher job security when teaching larger classes. Enrollment $N_i(\theta_L, \theta_S)$ depends on the grading policies of both instructors and on the information regime. Let $V(N_i)$ be the instructor's benefit from enrollment, $V'(N_i) > 0$. Define the instructor's utility from grading policies by

$$U_i(\theta_L, \theta_S) = V(N_i(\theta_L, \theta_S)) - \tfrac{1}{2}(\theta_i - \theta_i^*)^2 \quad \text{for } i = L, S.$$

The first term captures the strategic component of payoffs. The second term is the disutility incurred by an instructor who has an ideal policy, θ_i^*, but is induced to choose a nonideal one, θ_i.

[12] Achen and Courant (2009) provide empirical evidence consistent with this assumption.

In the no-information regime, there are no strategic effects. Instructors therefore choose their ideal policies, θ_L^* and θ_S^*. In the students-only and students and employers information regimes, in contrast, an interior equilibrium is a solution to the system of best-response functions, $dU_i(\theta_L, \theta_S)/d\theta_i = 0$, which can be written as

$$\theta_i = \theta_i^* + V'(N_i(\theta_L, \theta_S))\frac{\partial N_i(\theta_L, \theta_S)}{\partial \theta_i} \quad \text{for } i = L, S. \tag{9}$$

If enrollment decreases with the strictness of grading, $\partial N_i(\theta_L, \theta_S)/\partial \theta_i < 0$, then the equilibrium with information has more lenient grading policies, $\theta_i < \theta_i^*$.[13] In this case, providing grade distribution information results in an increase in the average grade for two reasons: (1) student selection and (2) more lenient grading standards.

VI. Concluding Remarks

Several institutions of higher education in the United States have implemented reforms that provide information on grading policies. Other institutions are considering the adoption of such reforms. Additionally, advances in information technology and online social networks facilitate the dissemination of information on grading policies, at least among students. For example, Phillips (2011) describes a new website in which "students who register for CourseRank will be able to take into account a professor's grade distribution, along with peer reviews and ratings, when deciding whether to take a class." Our model illustrates that providing information about grade distributions to students, or to both students and employers, may have adverse unintended consequences. The provision of information affects course selection patterns and ability distributions across courses. It may result in higher enrollment into leniently graded courses and in an increased average grade across courses. These effects can be exacerbated by instructors' responses to the policy reform.

How students are affected by the provision of information depends on their abilities and tastes. If information is provided to both students and employers, some students (those with a strong preference for the leniently graded course) become worse off because employers will associate them with a lower expected ability. Other students (those with a strong preference for the strictly graded course) become better off because employers

[13] Enrollment decreases with grading strictness for at least some range of parameters. For example, in the students-only regime, when $\bar{\tau} = 1$,

$$\frac{\partial N_L(\theta_L, \theta_S)}{\partial \theta_L} = -[1 - H(\tfrac{1}{2})]f(\theta_L) < 0$$

and

$$\frac{\partial N_S(\theta_L, \theta_S)}{\partial \theta_S} = -[1 - H(\tfrac{1}{2})]f(\theta_S) < 0.$$

will associate them with a higher expected ability. If students are paid their expected ability, utilitarian welfare is lower when information is provided.

An additional potential effect of putting grades in context is that instructors might try to circumvent the policy. Once grading policies become observed by employers, professors might resort to alternatives for lenient grading, such as reducing the amount of effort required in the class. Students might be able to obtain information about how demanding courses are from online sources that are less easily accessible to employers. If watered-down courses attract more students, the quality of education will decline. Here too we would expect the effects on selection to vary by student ability. Low-ability students might be more attracted to the less demanding courses.

Putting grades in context reforms may lead to a decline in the quality of information employers obtain about student abilities. This could in turn adversely affect the efficiency of job market matching. The provision of information about grade distributions can thus have important consequences. We hope that our analysis will stimulate further research on the question of reforms in grading policies.

Appendix

Proof of Proposition 1

Assume first uninformed students and employers. In this case the equilibrium expected ability depends only on the grade: $E(\tilde{\theta}|\mathrm{B})$ and $E(\tilde{\theta}|\mathrm{A})$. With the equilibrium value for each letter grade fixed, a student of ability θ will be perceived by employers to have an expected ability

$$\mathrm{prob}(\theta_i \leq \theta)E(\tilde{\theta}|\mathrm{A}) + \mathrm{prob}(\theta_i > \theta)E(\tilde{\theta}|\mathrm{B})$$

from course i. Because students have the same prior on the grading policies θ_i, given θ, a student expects to have a grade A in either course with the same probability: $\mathrm{prob}(\theta_1 \leq \theta) = \mathrm{prob}(\theta_2 \leq \theta)$. Hence, the expected ability of a student is the same in either course. Therefore, students with a taste parameter $\tau < \frac{1}{2}$ prefer course $i = 1$, and students with a taste parameter $\tau > \frac{1}{2}$ prefer course $i = 2$; that is, students select according to their tastes.

Consider now uninformed students and informed employers. For any given pair of grading policies (θ_1, θ_2), employers' expectations about students' abilities will depend on the student's grade and on grading policies, (θ_1, θ_2). We make three observations: (1) Because students' expectations about grading policies are independently and identically distributed, students assign the same probability (or density) to the events (θ', θ'') and (θ'', θ'). (2) A student would have the same grade g in course i when the policies are (θ', θ'') as in course j when the policies are (θ'', θ'). (3) In this case, employers perceive the same expected ability:

$$E(\tilde{\theta}|g \text{ in } 1)|_{(\theta'',\theta')} = E(\tilde{\theta}|g \text{ in } 2)|_{(\theta',\theta'')}.$$

Put together, these observations imply that, for every θ,

$$E_{(\theta_1,\theta_2)}(E(\tilde{\theta}|g(\theta, 1) \text{ in } 1)|_{(\theta_1,\theta_2)}) = E_{(\theta_1,\theta_2)}(E(\tilde{\theta}|g(\theta, 2) \text{ in } 2)|_{(\theta_1,\theta_2)}).$$

The student has the same expected ability (as perceived by employers) in either course and thus chooses according to tastes. QED

Proof of Corollary 1

Follows immediately from the fact that in both regimes students choose according to their tastes. QED

Proof of Proposition 2

When only students are informed, employers' expectations depend only on the grade: $E(\tilde{\theta}|\mathrm{B})$ and $E(\tilde{\theta}|\mathrm{A})$. Low-ability students will have a B in either course and high-ability students will have an A in either course. Hence, these students choose according to their tastes: course L if $c(\tau) < c(1 - \tau)$, that is, $\tau < \frac{1}{2}$, and course S if $\tau > \frac{1}{2}$.

Intermediate-ability students will have an A in course L and a B in course S. We argue that $E(\tilde{\theta}|\mathrm{A}) > E(\tilde{\theta}|\mathrm{B})$. Note that

$$E(\tilde{\theta}|\tilde{\theta} \geq \theta_S) > E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S)) > E(\tilde{\theta}|\tilde{\theta} < \theta_L). \quad \text{(A1)}$$

The expected ability conditional on a grade A is

$$E(\tilde{\theta}|\mathrm{A}) = \frac{H(\bar{\tau})\int_{\theta_L}^{\theta_S} \theta f(\theta)d\theta + \int_{\theta_S}^{1} \theta f(\theta)d\theta}{H(\bar{\tau})[F(\theta_S) - F(\theta_L)] + [1 - F(\theta_S)]},$$

which is a combination of expected abilities of students in the high and intermediate ranges,

$$E(\tilde{\theta}|\mathrm{A}) = \alpha E(\tilde{\theta}|\tilde{\theta} \geq \theta_S) + (1 - \alpha)E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S)), \quad \text{(A2)}$$

where

$$\alpha = \frac{1 - F(\theta_S)}{H(\bar{\tau})[F(\theta_S) - F(\theta_L)] + [1 - F(\theta_S)]}.$$

For $\theta_S < 1$, $\alpha > 0$. By (A1) and (A2), $E(\tilde{\theta}|\mathrm{A}) > E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S))$.

Similarly, conditional on the grade B,

$$E(\tilde{\theta}|\mathrm{B}) = \beta E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S)) + (1 - \beta)E(\tilde{\theta}|\tilde{\theta} < \theta_L), \quad \text{(A3)}$$

where

$$\beta = \frac{[1 - H(\bar{\tau})][F(\theta_S) - F(\theta_L)]}{F(\theta_L) + [1 - H(\bar{\tau})][F(\theta_S) - F(\theta_L)]}.$$

Hence, $E(\tilde{\theta}|\mathrm{B}) \leq E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S))$. Combining these findings, we have

$$E(\tilde{\theta}|\text{A}) > E(\tilde{\theta}|\text{B}).$$

Thus, intermediate-ability students with taste parameter $\tau \leq \frac{1}{2}$ strictly prefer course L. An equilibrium threshold must satisfy $\bar{\tau} > \frac{1}{2}$.

Finally, we argue that an equilibrium $\bar{\tau}$ exists. Let

$$\Delta u(\tau) = E(\tilde{\theta}|\text{A}) - E(\tilde{\theta}|\text{B}) + c(1-\tau) - c(\tau), \tag{A4}$$

where the expectations are defined in (A2) and (A3) using the variable τ wherever $\bar{\tau}$ appears. If $\Delta u(1) \geq 0$, then $\bar{\tau} = 1$ is an equilibrium. If $\Delta u(1) < 0$, then because $\Delta u(\frac{1}{2}) > 0$ and $\Delta u(\tau)$ is continuous, there exists $\bar{\tau} \in (\frac{1}{2}, 1)$ so that $\Delta u(\bar{\tau}) = 0$. QED

Proof of Corollary 2

i. In the no-information regime, enrollment into course L is $H(\frac{1}{2})$ because students choose according to their tastes. By (3) and because $\bar{\tau} > \frac{1}{2}$, in the students-only regime, enrollment into course L satisfies

$$H(\tfrac{1}{2})F(\theta_L) + H(\bar{\tau})[F(\theta_S) - F(\theta_L)] + H(\tfrac{1}{2})[1 - F(\theta_S)] > H(\tfrac{1}{2}).$$

ii. In the student-only regime, students' grades are either the same as or higher than in the no-information regime. Hence, the average grade is higher. QED

Proof of Proposition 3

By definition of the grading policies, in this information regime we have

$$E(\tilde{\theta}|\text{B in } L) = E(\tilde{\theta}|\tilde{\theta} < \theta_L), \tag{A5}$$

$$E(\tilde{\theta}|\text{B in } S) = \alpha E(\tilde{\theta}|\tilde{\theta} < \theta_L) + (1-\alpha)E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S)), \tag{A6}$$

where

$$\alpha = \frac{[1 - H(\bar{\tau}_1)]F(\theta_L)}{[1 - H(\bar{\tau}_1)]F(\theta_L) + [1 - H(\bar{\tau}_2)][F(\theta_S) - F(\theta_L)]},$$

$$E(\tilde{\theta}|\text{A in } S) = E(\tilde{\theta}|\tilde{\theta} \geq \theta_S), \tag{A7}$$

$$E(\tilde{\theta}|\text{A in } L) = \beta E(\tilde{\theta}|\tilde{\theta} \geq \theta_S) + (1-\beta)E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S)), \tag{A8}$$

where

$$\beta = \frac{H(\bar{\tau}_3)[1 - F(\theta_S)]}{H(\bar{\tau}_2)[F(\theta_S) - F(\theta_L)] + H(\bar{\tau}_3)[1 - F(\theta_S)]}.$$

By (A5) and (A6) we have $E(\tilde{\theta}|\text{B in } S) - E(\tilde{\theta}|\text{B in } L) \geq 0$. Hence, in equilibrium all low-ability students with taste parameters $\tau > \frac{1}{2}$ enroll in course S, which implies that $\bar{\tau}_1 \leq \frac{1}{2}$. Compared to the no-information regime, enrollment of low-ability students to the strictly graded course S is higher.

Next, we show that $E(\tilde{\theta}|\text{A in } L) - E(\tilde{\theta}|\text{B in } S) > 0$. This implies that all intermediate-ability students with taste parameter $\tau \leq \frac{1}{2}$ enroll in course L, and hence $\bar{\tau}_2 > \frac{1}{2}$. The inequality holds true because by (A6), (A8), and $\alpha > 0$, we have

$$E(\tilde{\theta}|\text{B in } S) < E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S)) \leq E(\tilde{\theta}|\text{A in } L).$$

Hence, compared to the no-information regime, enrollment of intermediate-ability students in the strictly graded course S is lower.

Finally, $\bar{\tau}_2 > \frac{1}{2}$ implies $\beta < 1$. Hence, by (A7) and (A8) we have $E(\tilde{\theta}|\text{A in } S) > E(\tilde{\theta}|\text{A in } L)$ and $\bar{\tau}_3 < \frac{1}{2}$. QED

Proof of Lemma 1

By (A7), $E(\tilde{\theta}|\text{A in } S) = E(\tilde{\theta}|\tilde{\theta} \geq \theta_S)$. Additionally, because $\bar{\tau}_2 > \frac{1}{2} > \bar{\tau}_3$, we know that $E(\tilde{\theta}|\text{A in } L) \leq E(\tilde{\theta}|\tilde{\theta} \geq \theta_L)$. Hence,

$$E(\tilde{\theta}|\text{A in } S) - E(\tilde{\theta}|\text{A in } L) \geq E(\tilde{\theta}|\tilde{\theta} \geq \theta_S) - E(\tilde{\theta}|\tilde{\theta} \geq \theta_L) > c(1).$$

The second inequality holds because the courses are similar (see [8]). By the equilibrium condition (6), this implies that all high-ability students prefer course S, $\bar{\tau}_3 = 0$.

By (A8), $E(\tilde{\theta}|\text{A in } L) \geq E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S))$ and $E(\tilde{\theta}|\text{B in } S) = E(\tilde{\theta}|\tilde{\theta} < \theta_S)$. Hence,

$$E(\tilde{\theta}|\text{A in } L) - E(\tilde{\theta}|\text{B in } S) \geq E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S)) - E(\tilde{\theta}|\tilde{\theta} < \theta_S) > c(1).$$

The second inequality holds because the courses are similar. By the equilibrium condition (7), this implies $\bar{\tau}_2 = 1$.

Finally, because $\bar{\tau}_2 = 1$, $E(\tilde{\theta}|\text{B in } S) = E(\tilde{\theta}|\text{B in } L)$. Therefore, for low-ability students, $\bar{\tau}_1 = \frac{1}{2}$. QED

Proof of Corollary 3

In the students and employers regime, when the courses are similar, by lemma 1, $(\tau_1^*, \tau_2^*, \tau_3^*) = (\frac{1}{2}, 1, 0)$. Substituting enrollment defined by (3),

$$N_L(\tfrac{1}{2}, 1, 0) = H(\tfrac{1}{2})F(\theta_L) + [F(\theta_S) - F(\theta_L)] > H(\tfrac{1}{2})$$

$$\Leftrightarrow [1 - H(\tfrac{1}{2})][F(\theta_S) - F(\theta_L)] > H(\tfrac{1}{2})[1 - F(\theta_S)].$$

QED

Proof of Corollary 4

In the students and employers regime, students' grades are either the same as or higher than in the no-information regime. Hence, the average grade is higher. QED

Lemma 2. Consider the students-only regime and the function

$\Delta u(\tau) = E(\tilde{\theta}|\mathrm{A}) - E(\tilde{\theta}|\mathrm{B}) + c(1-\tau) - c(\tau)$ (where the expectations are defined in [A2] and [A3] using the variable τ wherever there is $\bar{\tau}$).

i. There is a unique equilibrium with $\bar{\tau} = 1$ if $\Delta u(\tau) > 0$ for all τ; a sufficient condition for this to hold is $c(1) < E(\tilde{\theta}|\tilde{\theta} \geq \theta_L) - E(\tilde{\theta}|\tilde{\theta} < \theta_S)$.

ii. There is a unique interior equilibrium ($\bar{\tau} < 1$) if $\Delta u(\tau)$ is convex and $\Delta u(1) < 0$; sufficient conditions are that tastes are uniformly distributed, $c(\tau)$ is linear or quadratic, and $c(1) > E(\tilde{\theta}|\tilde{\theta} \geq \theta_L) - E(\tilde{\theta}|\tilde{\theta} < \theta_L)$.

Proof of Lemma 2

i. When $\Delta u(\tau) > 0$ for all τ, there can be an equilibrium only with $\bar{\tau} = 1$ (because when $\bar{\tau} < 1$ is an equilibrium, then $\Delta u(\tau) = 0$). Under the assumption on $c(1)$,

$$\Delta u(\tau) \geq E(\tilde{\theta}|\tilde{\theta} \geq \theta_L) - E(\tilde{\theta}|\tilde{\theta} < \theta_S) - c(1) > 0$$

for all τ. The first inequality follows from the definition of $\Delta u(\tau)$ and the second from the assumption on $c(1)$.

ii. We know that

$$\Delta u(\tfrac{1}{2}) = E(\tilde{\theta}|\mathrm{A})|_{\bar{\tau}=1/2} - E(\tilde{\theta}|\mathrm{B})|_{\bar{\tau}=1/2} > 0 > \Delta u(1).$$

Hence, $\Delta u(\tau) = 0$ has at least one solution, $\bar{\tau} \in (\frac{1}{2}, 1)$. Suppose by contradiction that there is more than one solution to $\Delta u(\tau) = 0$. Denote two solutions by $\hat{\tau}$ and $\bar{\tau}$, where $\bar{\tau} > \hat{\tau}$. If $\Delta u'(\bar{\tau}) \geq 0$, then by convexity the function is weakly increasing for all $\tau > \bar{\tau}$, and we know $\Delta u(\bar{\tau}) = 0$. This contradicts $\Delta u(1) < 0$. It must be that $\Delta u'(\bar{\tau}) < 0$. By convexity, the function is decreasing for all $\tau < \bar{\tau}$. Therefore, in this range, $\Delta u(\tau) > 0$. This contradicts there being another zero $\hat{\tau} < \bar{\tau}$.

To show that $\Delta u(\tau)$ is convex when tastes are uniformly distributed and $c(\tau)$ is linear or quadratic, we show that $d^2\Delta u(\tau)/d\tau^2 > 0$. Note that for the linear or quadratic cost functions, $c(1-\tau) - c(\tau)$ is linear. We focus, therefore, on the derivatives of $[E(\tilde{\theta}|\mathrm{A}) - E(\tilde{\theta}|\mathrm{B})]$:

$$\begin{aligned}
&\frac{d}{d\tau}[E(\tilde{\theta}|\mathrm{A}) - E(\tilde{\theta}|\mathrm{B})] \\
&\quad = [F(\theta_S) - F(\theta_L)]\left\{\frac{\int_{\theta_L}^{\theta_S}\theta f(\theta)d\theta}{F(\theta_S) - F(\theta_L)} - \frac{\tau\int_{\theta_L}^{\theta_S}\theta f(\theta)d\theta + \int_{\theta_S}^{1}\theta f(\theta)d\theta}{\tau[F(\theta_S) - F(\theta_L)] + 1 - F(\theta_S)}\right\} \\
&\quad\quad \div \{\tau[F(\theta_S) - F(\theta_L)] + 1 - F(\theta_S)\} \\
&\quad\quad + [F(\theta_S) - F(\theta_L)]\left\{\frac{\int_{\theta_L}^{\theta_S}\theta f(\theta)d\theta}{F(\theta_S) - F(\theta_L)} - \frac{\int_{0}^{\theta_L}\theta f(\theta)d\theta + (1-\tau)\int_{\theta_L}^{\theta_S}\theta f(\theta)d\theta}{F(\theta_L) + (1-\tau)[F(\theta_S) - F(\theta_L)]}\right\} \\
&\quad\quad \div \{F(\theta_L) + (1-\tau)[F(\theta_S) - F(\theta_L)]\};
\end{aligned}$$

$$\frac{d^2}{d\tau^2}[E(\tilde{\theta}|\text{A}) - E(\tilde{\theta}|\text{B})]$$

$$\propto \left\{\frac{\int_{\theta_L}^{\theta_S}\theta f(\theta)d\theta}{F(\theta_S) - F(\theta_L)} - \frac{\int_0^{\theta_L}\theta f(\theta)d\theta + (1-\tau)\int_{\theta_L}^{\theta_S}\theta f(\theta)d\theta}{F(\theta_L) + (1-\tau)[F(\theta_S) - F(\theta_L)]}\right\}$$

$$\div \{F(\theta_L) + (1-\tau)[F(\theta_S) - F(\theta_L)]\}^2$$

$$+ \left\{\frac{\tau\int_{\theta_L}^{\theta_S}\theta f(\theta)d\theta + \int_{\theta_S}^{1}\theta f(\theta)d\theta}{\tau[F(\theta_S) - F(\theta_L)] + 1 - F(\theta_S)} - \frac{\int_{\theta_L}^{\theta_S}\theta f(\theta)d\theta}{F(\theta_S) - F(\theta_L)}\right\}$$

$$\div \{\tau[F(\theta_S) - F(\theta_L)] + 1 - F(\theta_S)\}^2 > 0.$$

Finally,

$$\Delta u(1) = E(\tilde{\theta}|\tilde{\theta} \geq \theta_L) - E(\tilde{\theta}|\tilde{\theta} < \theta_L) - c(1) < 0$$

under the condition in part ii on $c(1)$. QED

Proof of Proposition 4

Consider $\Delta u(\tau)$, defined in (A4) in the proof of proposition 2. We know that $\Delta u(\frac{1}{2}) > 0$ because $E(\tilde{\theta}|\text{A}) > E(\tilde{\theta}|\text{B})$. If $\Delta u(1) \geq 0$, then $\bar{\tau} = 1$ is the equilibrium and $\bar{\tau} \geq \bar{\tau}_2$. If instead $\Delta u(1) < 0$, $\bar{\tau}$ solves $\Delta u(\bar{\tau}) = 0$. Because the equilibrium is unique, in this case $\Delta u(\tau) > 0$ for $\tau < \bar{\tau}$ and $\Delta u(\tau) < 0$ for $\tau > \bar{\tau}$.

By (7), we know that

$$E(\tilde{\theta}|\text{A in } L) - E(\tilde{\theta}|\text{B in } S) \geq c(\bar{\tau}_2) - c(1 - \bar{\tau}_2)$$

(holding with equality when $\bar{\tau}_2$ is interior). The expectations on the left-hand side are given by (A8) and (A6). We now evaluate the difference $E(\tilde{\theta}|\text{A}) - E(\tilde{\theta}|\text{B})$ at $\bar{\tau}_2$. We argue that

$$E(\tilde{\theta}|\text{A}) - E(\tilde{\theta}|\text{B}) > E(\tilde{\theta}|\text{A in } L) - E(\tilde{\theta}|\text{B in } S).$$

To see why this is true, note that $E(\tilde{\theta}|\text{A}) > E(\tilde{\theta}|\text{A in } L)$ because both are weighted averages of $E(\tilde{\theta} \geq \theta_S)$ and $E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S))$, but in $E(\tilde{\theta}|\text{A})$ there is a larger weight on $E(\tilde{\theta} \geq \theta_S)$ (all the high-ability students are included in the mean, not just those in course L). Additionally, $E(\tilde{\theta}|\text{B}) \leq E(\tilde{\theta}|\text{B in } S)$ because both are weighted averages of $E(\tilde{\theta}|\tilde{\theta} \in [\theta_L, \theta_S))$ and $E(\tilde{\theta}|\tilde{\theta} < \theta_L)$, but in $E(\tilde{\theta}|\text{B})$ there is a larger weight on $E(\tilde{\theta}|\tilde{\theta} < \theta_L)$ (all low-ability students are included in the average, not just those in course S). It follows that, when evaluating at $\bar{\tau}_2$,

$$E(\tilde{\theta}|\text{A})|_{\bar{\tau}_2} - E(\tilde{\theta}|\text{B})|_{\bar{\tau}_2} > E(\tilde{\theta}|\text{A in } L) - E(\tilde{\theta}|\text{B in } S) \geq c(\bar{\tau}_2) - c(1 - \bar{\tau}_2).$$

Hence, $\Delta u(\bar{\tau}_2) > 0$ and $\bar{\tau}_2 < \bar{\tau}$. QED

Proof of Proposition 5

i. The variance of students' abilities in the no-information regime (regime 1) is

$$\mathrm{Var}_1(\tilde{\theta}|\mathrm{A}) = \int_{\theta_L}^{1}\int_{0}^{1/2} \frac{[\theta - E(\tilde{\theta}|\mathrm{A})]^2}{1 - \frac{1}{2}\theta_L - \frac{1}{2}\theta_S} d\tau d\theta + \int_{\theta_S}^{1}\int_{1/2}^{1} \frac{[\theta - E(\tilde{\theta}|\mathrm{A})]^2}{1 - \frac{1}{2}\theta_L - \frac{1}{2}\theta_S} d\tau d\theta$$

$$= \gamma \int_{\theta_L}^{1} \frac{[\theta - E(\tilde{\theta}|\mathrm{A})]^2}{1 - \theta_L} d\theta + (1-\gamma) \int_{\theta_S}^{1} \frac{[\theta - E(\tilde{\theta}|\mathrm{A})]^2}{1 - \theta_S} d\theta,$$

where $\gamma = (1 - \theta_L)/(2 - \theta_L - \theta_S)$.

We replace $E(\tilde{\theta}|\mathrm{A})$ with $E(\tilde{\theta}|\tilde{\theta} \geq \theta_L)$ and $E(\tilde{\theta}|\tilde{\theta} \geq \theta_S)$ in the first and second terms, respectively. Because the mean of a distribution minimizes the expected value of squared errors, it follows that

$$\mathrm{Var}_1(\tilde{\theta}|\mathrm{A}) \geq \gamma \int_{\theta_L}^{1} \frac{[\theta - E(\tilde{\theta}|\tilde{\theta} \geq \theta_L)]^2}{1 - \theta_L} d\theta$$

$$+ (1-\gamma) \int_{\theta_S}^{1} \frac{[\theta - E(\tilde{\theta}|\tilde{\theta} \geq \theta_S)]^2}{1 - \theta_S} d\theta$$

$$\geq \gamma \int_{\theta_S}^{1} \frac{[\theta - E(\tilde{\theta}|\tilde{\theta} \geq \theta_S)]^2}{1 - \theta_S} d\theta$$

$$+ (1-\gamma) \int_{\theta_S}^{1} \frac{[\theta - E(\tilde{\theta}|\tilde{\theta} \geq \theta_S)]^2}{1 - \theta_S} d\theta$$

$$= \int_{\theta_S}^{1} \frac{[\theta - E(\tilde{\theta}|\tilde{\theta} \geq \theta_S)]^2}{1 - \theta_S} d\theta = \mathrm{Var}_2(\tilde{\theta}|\mathrm{A\ in\ } S)$$

$$= \mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ } S).$$

The second inequality follows because for the uniform distribution, the conditional variance in a range $[\theta, 1]$ decreases with θ.

ii. Consider the variance of abilities for students with grade A in course L for threshold taste parameters τ_2^* and τ_3^*:

$$\mathrm{Var}(\tilde{\theta}|\mathrm{A\ in\ } L) = \frac{1}{3} \frac{\tau_2^*(\theta_S^3 - \theta_L^3) + \tau_3^*(1 - \theta_S^3)}{\tau_2^*(\theta_S - \theta_L) + \tau_3^*(1 - \theta_S)}$$

$$- \frac{1}{4}\left[\frac{\tau_2^*(\theta_S^2 - \theta_L^2) + \tau_3^*(1 - \theta_S^2)}{\tau_2^*(\theta_S - \theta_L) + \tau_3^*(1 - \theta_S)}\right]^2.$$

Evaluating at $\theta_S = \theta_L$ and at $\theta_S = 1$, we find in both cases that

$$\mathrm{Var}(\tilde{\theta}|\mathrm{A\ in\ } L) = \tfrac{1}{3}(1 + \theta_L + \theta_L^2) - \tfrac{1}{4}(1 + \theta_L)^2 = \mathrm{Var}(\tilde{\theta}|\tilde{\theta} \geq \theta_L).$$

Note that this expression does not depend on τ_i^*. Hence, in all information regimes for $\theta_S \to \theta_L$ and for $\theta_S \to 1$ in the limit,

$$\mathrm{Var}_1(\tilde{\theta}|\mathrm{A}) = \mathrm{Var}_2(\tilde{\theta}|\mathrm{A\ in\ }L) = \mathrm{Var}_3(\tilde{\theta}|\mathrm{A}) = \mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ }L). \quad \text{(A9)}$$

Consider the effect of θ_S on the variance of student abilities:

$$\frac{d\,\mathrm{Var}(\tilde{\theta}|\mathrm{A\ in\ }L)}{d\theta_S} = \frac{\partial\,\mathrm{Var}(\tilde{\theta}|\mathrm{A\ in\ }L)}{\partial\tau_2^*}\frac{d\tau_2^*}{d\theta_S} + \frac{\partial\,\mathrm{Var}(\tilde{\theta}|\mathrm{A\ in\ }L)}{\partial\tau_3^*}\frac{d\tau_3^*}{d\theta_S} + \frac{\partial\,\mathrm{Var}(\tilde{\theta}|\mathrm{A\ in\ }L)}{\partial\theta_S}.$$

Evaluated at $\theta_S = \theta_L$ and $\theta_S = 1$, the partial derivatives with respect to τ_2^* and τ_3^* are equal to zero. Hence, $d\,\mathrm{Var}/d\theta_S = \partial\,\mathrm{Var}/\partial\theta_S$:

$$\frac{\partial\,\mathrm{Var}}{\partial\theta_S} = \tfrac{1}{3}(\tau_2^* - \tau_3^*)\frac{3\theta_S^2[\tau_2^*(\theta_S - \theta_L) + \tau_3^*(1-\theta_S)] - [\tau_2^*(\theta_S^3 - \theta_L^3) + \tau_3^*(1-\theta_S^3)]}{[\tau_2^*(\theta_S - \theta_L) + \tau_3^*(1-\theta_S)]^2}$$
$$- \tfrac{1}{2}(\tau_2^* - \tau_3^*)[\tau_2^*(\theta_S^2 - \theta_L^2) + \tau_3^*(1-\theta_S^2)]$$
$$\times\ \frac{2\theta_S[\tau_2^*(\theta_S - \theta_L) + \tau_3^*(1-\theta_S)] - [\tau_2^*(\theta_S^2 - \theta_L^2) + \tau_3^*(1-\theta_S^2)]}{[\tau_2^*(\theta_S - \theta_L) + \tau_3^*(1-\theta_S)]^3}.$$

Evaluating at $\theta_S = \theta_L$ and then at $\theta_S = 1$, we have

$$\frac{d\,\mathrm{Var}}{d\theta_S} = \begin{cases} \dfrac{\tau_2^* - \tau_3^*}{6\tau_3^*}(1-\theta_L) > 0 & \text{if } \theta_S = \theta_L \\ \dfrac{\tau_2^* - \tau_3^*}{6\tau_2^*}(1-\theta_L) > 0 & \text{if } \theta_S = 1. \end{cases} \quad \text{(A10)}$$

In regime 4 (students and employers), $\bar{\tau}_2 > \bar{\tau}_3$; therefore, $d\,\mathrm{Var}/d\theta_S > 0$ at $\theta_S = \theta_L$ and at $\theta_S = 1$.

To find the effect on the variance in the no-information regime, we evaluate (A10) at $\tau_2^* = \frac{1}{2}$ and $\tau_3^* = 1$ (because all high-ability students get an A) to find

$$\left.\frac{d\,\mathrm{Var}_1}{d\theta_S}\right|_{\theta_S=\theta_L} = -\frac{1}{12}(1-\theta_L) < 0,$$
$$\left.\frac{d\,\mathrm{Var}_1}{d\theta_S}\right|_{\theta_S=1} = -\frac{1}{6}(1-\theta_L) < 0. \quad \text{(A11)}$$

It follows that in the neighborhood of $\theta_S = \theta_L$ and $\theta_S = 1$, the difference $\mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ }L) - \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$ is increasing. Moreover, at these points, $\mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ }L) = \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$ by (A9). Therefore, for $\theta_S > \theta_L$ but close to

$\theta_S = \theta_L$, we have $\mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ }L) > \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$, and for $\theta_S < 1$ but close to $\theta_S = 1$, we have $\mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ }L) < \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$.

The proof for comparing the employers-only to the no-information regime follows in a similar way, noting that $d\,\mathrm{Var}_2(\tilde{\theta}|\mathrm{A\ in\ }L)/d\theta_S = 0$.

iii. If the courses are similar, then

$$\mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ }L) \le \mathrm{Var}_1(\tilde{\theta}|\mathrm{A}) \Leftrightarrow$$

$$\tfrac{1}{3}(\theta_L^2 + \theta_L\theta_S + \theta_S^2) - \tfrac{1}{4}(\theta_S + \theta_L)^2 \le \tfrac{1}{3}(\theta_L^2 + \theta_L + 1) - \tfrac{1}{4}(1 + \theta_L)^2.$$

The inequality holds true because its left-hand side is increasing in θ_S and equal to the right-hand side when $\theta_S = 1$. QED

Proof of Proposition 6

We follow the same steps as for part ii of proposition 5. To compare the variance in regimes 3 and 1, we evaluate (A10) at $\tau_2^* = \bar{\tau}$ and $\tau_3^* = 1$ (because all high-ability students get an A) to find

$$\frac{d\,\mathrm{Var}_3}{d\theta_S} - \frac{d\,\mathrm{Var}_1}{d\theta_S} = \begin{cases} \dfrac{1}{12}(1-\theta_L)(2\bar{\tau}-1) > 0 & \text{if } \theta_S = \theta_L \\ \dfrac{1}{6}(1-\theta_L)\dfrac{2\bar{\tau}-1}{\bar{\tau}} > 0 & \text{if } \theta_S = 1; \end{cases} \tag{A12}$$

the inequalities hold because $\bar{\tau} > \frac{1}{2}$. This implies that for $\theta_S > \theta_L$ but close to $\theta_S = \theta_L$, we have $\mathrm{Var}_3(\tilde{\theta}|\mathrm{A\ in\ }L) > \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$, and for $\theta_S < 1$ but close to $\theta_S = 1$, we have $\mathrm{Var}_3(\tilde{\theta}|\mathrm{A\ in\ }L) < \mathrm{Var}_1(\tilde{\theta}|\mathrm{A})$.

To compare the variance in regimes 4 and 3, we evaluate (A10) to find

$$\frac{d\,\mathrm{Var}_4}{d\theta_S} - \frac{d\,\mathrm{Var}_3}{d\theta_S} = \begin{cases} \dfrac{\bar{\tau}_2 - \bar{\tau}_3}{6\bar{\tau}_3}(1-\theta_L) + \dfrac{1-\bar{\tau}}{6}(1-\theta_L) > 0 & \text{if } \theta_S = \theta_L \\ \dfrac{\bar{\tau}_2 - \bar{\tau}_3}{6\bar{\tau}_2}(1-\theta_L) + \dfrac{1-\bar{\tau}}{6\bar{\tau}}(1-\theta_L) > 0 & \text{if } \theta_S = 1. \end{cases} \tag{A13}$$

This implies that for $\theta_S > \theta_L$ but close to $\theta_S = \theta_L$, we have $\mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ }L) > \mathrm{Var}_3(\tilde{\theta}|\mathrm{A})$, and for $\theta_S < 1$ but close to $\theta_S = 1$, we have $\mathrm{Var}_4(\tilde{\theta}|\mathrm{A\ in\ }L) < \mathrm{Var}_3(\tilde{\theta}|\mathrm{A})$. QED

Proof of Proposition 7

i. Because $\theta_1 = \theta_2$,

$$E(\tilde{\theta}|\mathrm{A\ in\ }L) = E(\tilde{\theta}|\mathrm{A\ in\ }S) = E(\tilde{\theta}|\tilde{\theta} \ge \theta_i)$$

and

$$E(\tilde{\theta}|\mathrm{B\ in\ }L) = E(\tilde{\theta}|\mathrm{B\ in\ }S) = E(\tilde{\theta}|\tilde{\theta} < \theta_i).$$

Students of any ability level will have the same grade and the same expected ability in either course. Hence, choosing according to taste is optimal.

ii. Because all students with type $\theta \geq \theta_i$ obtain the grade A, the average grade (GPA) is given by GPA $= [1 - F(\theta_i)]\text{A} + F(\theta_i)\text{B}$, where A and B are assigned some numerical values (e.g., A = 4 and B = 3). Take $\theta_i = F^{-1}((\text{A} - \text{GPA})/(\text{A} - \text{B}))$ to obtain the desired GPA. Note that we assumed a positive density so that F^{-1} is well defined for $(\text{A} - \text{GPA})/(\text{A} - \text{B}) \in [0, 1]$.

iii. The policy does not result in a ranking bias because for every $\theta_h > \theta_l$, either both are on the same side of θ_i and obtain the same grade or $\theta_h \geq \theta_i > \theta_l$ and θ_h has the higher grade. QED

Proof of Proposition 8

i. Recall that for a student (θ, τ), we denoted by $\sigma(\theta, \tau)$ the course selected and by $g(\theta, \sigma(\theta, \tau))$ the grade in this course. In all information regimes,

$$\int_0^1 \int_0^1 [E(\tilde{\theta} | g(\theta, \sigma(\theta, \tau)) \text{ in } \sigma(\theta, \tau))] dH(\tau) dF(\theta) = \int_0^1 \int_0^1 \theta dH(\tau) dF(\theta).$$

Therefore, the utilitarian welfare comparison depends on

$$\int_0^1 \int_0^1 [c(d(\tau, \sigma(\theta, \tau)))] dH(\tau) dF(\theta).$$

This cost is the same in regimes 1 and 2 (with uninformed students) because in both regimes students select courses according to tastes. It is higher in regimes 3 and 4 (with informed students) because in these regimes, for every student (θ, τ), the disutility from course selection is at least as high as it is in the no-information regime.

ii. When the courses are similar, then in regimes 3 and 4, all low-ability students choose according to tastes, and all intermediate-ability students select course L. These students have the same disutility from course selection in either regime. In the students-only regime (regime 3), in contrast, students of high ability choose according to their tastes, resulting in a lower disutility than in the students and employers regime, where all high-ability students choose course S. Hence, overall, the aggregate disutility is lower and welfare is higher in the students-only regime. QED

Proof of Proposition 9

i.

$$W_1 = W_0 + (R - 1) \int_{\theta_S}^1 \int_0^1 [u(\theta, \tau)] dH(\tau) dF(\theta).$$

We know by proposition 8 that W_0 is lower in the students-only regime. The second term is also lower because for high-ability students, expected ability is lower and the disutility from course selection is at least as high in the students-only regime.

ii. For $R \to 1$ we have $W_1 \to W_0$. By proposition 8, W_0 is lower in the students and employers regime than in the no-information regime. Hence, for $R \to 1$, welfare is lower in the students and employers regime. When $c(1)$ is sufficiently low, courses are similar by lemma 1, all high-ability students select course S, and they have a higher expected ability. Additionally, the disutility from course selection is low. Hence, for sufficiently low $c(1)$, $\int_{\theta_S}^1 \int_0^1 u(\theta, \tau) dH(\tau) dF(\theta)$ is higher in the students and employers regime than in the no-information regime. As a result, for a large enough R, W_1 is larger in the students and employers regime. QED

Proof of Example in Section V.A.1

i. Students only: High-ability students expect the same grade A in either course, and thus, with uninformed employers, they will have the same expected ability in either course. They therefore choose according to tastes, $\tau_3^* = \frac{1}{2}$. Intermediate-ability students prefer course L if

$$p_A E(\tilde{\theta}|\text{A}) + (1 - p_A) E(\tilde{\theta}|\text{B}) - c(\tau) > E(\tilde{\theta}|\text{B}) - c(1 - \tau)$$

or

$$p_A [E(\tilde{\theta}|\text{A}) - E(\tilde{\theta}|\text{B})] > c(\tau) - c(1 - \tau).$$

The left-hand side is nonnegative, and therefore $\bar{\tau}_2 \geq \frac{1}{2}$. Similarly, low-ability students select course L if

$$p_B [E(\tilde{\theta}|\text{B}) - E(\tilde{\theta}|\text{C})] > c(\tau) - c(1 - \tau).$$

The left-hand side is nonnegative, and therefore $\bar{\tau}_1 \geq \frac{1}{2}$. In equilibrium we will have $\bar{\tau}_1 > \bar{\tau}_2$ if and only if

$$p_B [E(\tilde{\theta}|\text{B}) - E(\tilde{\theta}|\text{C})] > p_A [E(\tilde{\theta}|\text{A}) - E(\tilde{\theta}|\text{B})].$$

For example, if $p_A \to 0$ (almost all intermediate-ability students will have the grade B in course L) and $p_B > 0$, then $\bar{\tau}_1 > \bar{\tau}_2 \to \frac{1}{2}$.

ii. Students and employers: For high-ability students, $E(\tilde{\theta}|\text{A in } S) \geq E(\tilde{\theta}|\text{A in } L)$ because students with an A in course S are all high-ability students, whereas in course L some are intermediate-ability students. Hence, $\bar{\tau}_3 \leq \frac{1}{2}$.

Intermediate-ability students prefer course L if

$$p_A [E(\tilde{\theta}|\text{A in } L) - E(\tilde{\theta}|\text{B in } S)] + (1 - p_A)[E(\tilde{\theta}|\text{B in } L) - E(\tilde{\theta}|\text{B in } S)] > c(\tau) - c(1 - \tau). \tag{A14}$$

Because $E(\tilde{\theta}|\text{A in } L) - E(\tilde{\theta}|\text{B in } S) > 0$ and $E(\tilde{\theta}|\text{B in } L) - E(\tilde{\theta}|\text{B in } S) \leq 0$, for $p_A \to 1$, the left-hand side of (A14) is positive and we have $\bar{\tau}_2 \leq \frac{1}{2}$.

But when $p_A \to 0$, the left-hand side of (A14) is negative and we have $\bar{\tau}_2 \geq \frac{1}{2}$.

Low-ability students prefer course L if

$$p_B E(\tilde{\theta}|\text{B in } L) + (1 - p_B)E(\tilde{\theta}|\text{C in } L) - E(\tilde{\theta}|\text{C in } S) > c(\tau) - c(1 - \tau).$$

Note that since only low-ability students can get the grade C, $E(\tilde{\theta}|\text{C in } L) = E(\tilde{\theta}|\text{C in } S)$. The condition becomes

$$p_B[E(\tilde{\theta}|\text{B in } L) - E(\tilde{\theta}|\text{C in } L)] > c(\tau) - c(1 - \tau).$$

The right-hand side is nonnegative, and therefore $\bar{\tau}_1 \geq \frac{1}{2}$. QED

References

Achen, Alexandra C., and Paul N. Courant. 2009. What are grades made of? *Journal of Economic Perspectives* 23, no. 3:77–92.

Bar, Talia, Vrinda Kadiyali, and Asaf Zussman. 2005. Quest for knowledge and pursuit of grades. Unpublished manuscript, Department of Economics, Cornell University.

———. 2009. Grade information and grade inflation: The Cornell experiment. *Journal of Economic Perspectives* 23, no. 3:93–108.

Bar, Talia, and Asaf Zussman. Forthcoming. Partisan grading. *American Economic Journal: Applied Economics*.

Bartlett, Thomas, and Paula Wasley. 2008. Just say "A": Grade inflation undergoes reality check. *Chronicle of Higher Education*, September 5.

Bruno, Laura. 2007. Princeton leads in grade deflation. *USA Today*, March 28.

Chan, William, Hao Li, and Wing Suen. 2007. A signaling theory of grade inflation. *International Economic Review* 48, no. 3:1065–90.

Foderaro, Lisa W. 2010. Type-A-plus students chafe at grade deflation. *New York Times*, January 31.

Fournier, Gary M., and Tim R. Sass. 2000. Take my course, please: The effects of the principles experience on student curriculum choice. *Journal of Economic Education* 31, no. 4:323–39.

Harford, Tim. 2009. Outside edge: An easy answer to grade inflation. *Financial Times*, March 20.

Hotelling, Harold. 1929. Stability in competition. *Economic Journal* 39, no. 1:1–57.

Johnson, Valen E. 2003. *Grade inflation: A crisis in college education*. New York: Springer-Verlag.

Lewin, Tamar. 2010. A quest to explain what grades really mean. *New York Times*, December 26.

Ostrovsky, Michael, and Michael Schwarz. 2010. Equilibrium information disclosure: Grade inflation and unraveling. *American Economic Journal: Microeconomics* 2, no. 2:34–63.

Phillips, Lisa. 2011. Too much information. *New York Times*, January 9.
Primack, Phil. 2008. Doesn't anybody get a C anymore? *Boston Globe*, October 5.
Rosar, Frank, and Elisabeth Schulte. 2010. Imperfect private information and the design of information-generating mechanisms. Unpublished manuscript, Department of Economics, University of Bonn.
Rosovsky, Henry, and Matthew Hartley. 2002. *Evaluation and the academy: Are we doing the right thing?* Cambridge, MA: American Academy of Arts and Sciences.
Wasley, Paula. 2008. A statistics professor finds grade inflation a difficult problem. *Chronicle of Higher Education*, September 5.
Yang, Huanxing, and Chun Seng Yip. 2003. An economic theory of grade inflation. Unpublished manuscript, Department of Economics, University of Pennsylvania.