University Competition, Grading Standards and Grade Inflation

Popov, Sergey V. and Bernhardt, Dan

University of Illinois

4 November 2010

University Competition, Grading Standards and Grade Inflation*

Sergey V. Popov
Department of Economics
University of Illinois
popov2@illinois.edu

Dan Bernhardt
Department of Economics
University of Illinois
danber@illinois.edu

Draft: November 4, 2010

Abstract

We develop a model of strategic grade determination by universities distinguished by their distributions of student academic abilities. Universities choose grading standards to maximize total wages of graduates. Job placement and wages hinge on a firm's productivity assessment given a student's university, grade and productivity signal. We identify conditions under which better universities set lower grading standards, exploiting the fact that firms cannot distinguish between "good" and "bad" "A"s. In contrast, a social planner sets stricter standards at better universities. We show how increases in skilled jobs drive grade inflation, and determine when grading standards fall faster at better schools.

Keywords: grading standards, grading inflation, information.
JEL: I21.

*We thank Angelo Mele, Sara LaLumia, Bart Taub, Grigory Kosenok and Alexei Savvateev and participants at the 2010 Midwest Economic Association Meetings, the 2010 Missouri Economics Conference, Washington University, St. Louis, Novosibirsk State University (Russia), and the Higher School of Economics (Moscow) for helpful comments. All errors are ours.

1 Introduction

Universities award grades to measure the performance of students in courses. In turn, important decisions by third parties are based in part on GPAs — firms tend to offer higher wages to students with high GPAs, and graduate schools tend to admit high GPA students[1]. In this paper, we study how universities choose grading standards when they care about the decisions made by third parties based on GPAs. We characterize how student body qualities at different schools interact with the depth of the job market to affect equilibrium grading standards.

Our model reconciles three central empirical regularities describing grading over the past fifty years: (1) GPAs are higher at better schools, (2) GPAs have risen over time at all schools, and (3) grading standards have fallen faster over time at better schools. It is manifest that better universities award more high grades. For example, Rojstaczer (2003) finds that GPAs at private universities in 2006-2007 are 0.3 higher than at public universities. Table 1 reinforces these findings, presenting the evolution of grades at selected universities between 1960 and 2000. This table reveals that grades at better universities are uniformly higher. The table also highlights a uniform secular rise in GPAs over time. In addition, over the entire sample period, GPAs at better universities increased significantly faster, although there is no significant difference in grade inflation in different universities between 1980-2000.

We develop a model in which universities are distinguished by the distributions of "academic abilities" of their students: the distribution of student academic abilities at top schools conditionally stochastically dominates that at lesser schools. Firms value both "academic ability" and social skills, which are complements in production. There are two types of jobs, good and bad, which are distinguished by the higher marginal product of skills in good jobs. Good jobs are in limited supply. Universities determine which students receive "A" grades by setting endogenously-chosen

[1] There is a large empirical grading standard literature; see Bar and Zussman (2010), Rose and Betts (2004) and Bagues et al. (2008) for both the questions they study and their literature reviews.

	1960	1980	2000	1960 - 2000	1980-2000
Harvard U	2.7	3.05	3.41	0.71	0.36
Princeton U	2.83	3.13	3.36	0.53	0.23
Yale U	2.56	3.27	3.48	0.92	0.21
Columbia U		3.2	3.36		0.16
Stanford U	2.98	3.27	3.55	0.57	0.28
Northwestern U		3.02	3.35		0.33
U Chicago	2.5		3.26	0.76	
MIT	2.47	3.27	3.26	0.79	-0.01
Dartmouth C	2.47	3.06	3.33	0.86	0.27
Duke U	2.41	3.02	3.36	0.95	0.34
Average grade inflation				0.7613	0.2411
Standard deviation				0.1531	0.1145
U Illinois	2.77		3.12	0.35	
U Miami		2.7	3.05		0.35
Penn State U		2.86	2.99		13
U Wisconsin	2.51	2.89	3.13	0.62	0.24
U Texas		2.6	3		0.4
U Washington	2.31	2.97	3.12	0.81	0.15
UC Irvine		2.9	2.95		0.05
Lehigh U		2.6	2.97		0.37
Georgia Tech		2.78	2.97		0.19
George Washington U		3.03	3.25		0.22
Average grade inflation				0.5933	0.2563

GPA Data from GradeInflation.com. When year X was not available, the closest year was used; if there was no observation in ± 5 years, the observation was treated as missing. The top list of schools is from the top list of US News National Universities rankings, and the bottom list is from the bottom 50 and up of the same list for the sub-samples that had grade points in at least two time periods. Grade inflation is calculated as difference in GPAs.

Table 1: Evolution of GPAs at selected universities.

cutoffs on academic ability. Firms learn student social skills via job interviews, and forecast academic abilities using the information contained in the ability distribution at a student's university, the university's grading standard and the student's grade. Firms then make job assignments, and wages are competitively determined.

Universities understand how firms determine job placement and wages, and set grade standards to maximize the total expected wages of their graduates. Top universities would argue that their higher proportions of high grades simply reflect their better student bodies; a common grading standard would inevitably lead to more good grades at better schools. A central result of our analysis is that under weak conditions, top universities actually set *softer* grading standards: the marginal "A" student at a top university is *less* able than the marginal "A" student at a lesser university. The intuition for this result devolves from the basic observation that a marginal student at a top school can free ride on the better upper tail of students because firms cannot distinguish "good A" students from "bad A" students. In contrast, lesser schools must compete for better job assignments by raising the average ability of students who receive "A" grades, setting excessively high grading standards. It is the competition for good job assignments for graduates that generates this result—while giving more "A" grades lowers the expected productivity and hence wages of students who receive good job assignments, additional well-placed students more than offset this wage effect in the eyes of university.

Importantly, a social planner who seeks to maximize total output in society would choose the *opposite* ordering, setting more demanding grading standards at top schools whenever there is heterogeneity among students in social skills. Were academic skill the only source of heterogeneity, the social planner would set the same grading standard at each school. However, then the expected academic productivity of students with "A"s at top schools is higher. Firms, which do not see the abilities of students, would then "over-rate" marginal "A" students from top schools, assigning too many with low social skills to good jobs. The social planner wants to equalize the expected productivity of the marginal "A" students who receive good job assignments, and since the average social skill is less for students from top schools with good job assignments, the social planner sets a higher grading standard for "A"s at top schools.

Finally, we identify a plausible driving force underlying grade inflation at universities: the secular increase over time in the measure of good jobs relative to the measure of students, possibly reflecting the well-established shift toward skill-biased technologies. There is extensive evidence that skill demands at jobs have increased significantly, suggesting that there are now far more good jobs. We show that universities respond to an increase in good jobs by trying to place less able students at good jobs. In particular, universities reduce cutoffs for "A" grades.

The upward trend in grades is perhaps less interesting than the inference problems it creates for third parties. In and of itself, grade inflation might not increase inefficiency in the economy. However, we provide conditions under which grade inflation exacerbates differences in grading standards, further distorting hiring decisions. That is, we provide conditions under which the cutoff for an "A" falls by more at top schools. Intuitively, if there are very few good jobs, grade standards must be very high, so that with a common bounded support on ability, ability differences between the marginal "A" and average "A" student must be small at all schools. However, when more students get "A"s, this difference grows, and top schools exploit this via lowering their grading standards by more. We then provide conditions under which this greater reduction in grading standards at top schools is associated with a greater increase in the number of "A"s, i.e., for grade inflation at top schools to be higher.

We next review the literature. Section 2 presents the basic model of university competition. Section 3 derives equilibrium and social planner outcomes when all students have the same social skill. Section 4 extends the analysis to a setting in which students differ in social skills. Section 5 summarizes our findings and discusses the importance of our assumptions. An appendix contains all proofs.

The Literature. The paper closest to ours is probably a free-rider paper of Yang and Yip (2003), which also predicts more good grades in better schools. However, this paper predicts that all students receive the same wage, independently of grade: in their equilibrium, grades convey no information about students. Moreover, in their model, universities intentionally destroy value by explicitly lying about students' abilities.

Chen et al. (2007) consider a setting in which the measure of good students at a school is random, observed by schools, but not by firms. They model the intentional loss of academic reliability where the grading standard is not fixed, so that otherwise identical students who take identical actions might not receive the same grade. They argue that this is why grading standards have varied over time. However, they are silent as to why there should be significant unobservable variation in the quality of large populations of students at a university from year-to-year, especially given indirect, but broadly observable, measures of student quality such as mean SAT and ACT scores, and measures of typical high school class rank.

Dubey and Geanakoplos (2009) investigate how discreteness of grades influences student effort: they find that when students only care about relative rank, coarser grade structures can motivate students to study harder. MacLeod and Urquiola (2009) explore how the structure of the schooling market affects tradeoffs between studying effort, wealth and leisure.

A body of literature studies grading standards from the perspective of a central planner. Costrell (1994) studies how different policies toward standards affect student effort (he states that an egalitarian central planner is likely to pick lower grading standards than a total earnings maximizer), and provides a review of the grading literature; Betts (1998) provides an opposing argument.

2 The Competition Between Universities

The world contains two types of *universities*, $u \in \{H, I\}$. Universities are distinguished by the ability distributions of their student bodies. Abilities at a type H school are distributed according to a density $f_H(\theta)$, and the distribution at a type I school is $f_I(\theta)$. These densities are continuous and strictly positive on their common support, $[\underline{\theta}, \bar{\theta}]$. We capture the notion that the student body at a type H school is better with the concept of conditional first-order stochastic dominance: $f_H(x|x > t)$ first order stochastically dominates $f_I(\theta|\theta > t)$ for all $t \in [\underline{\theta}, \bar{\theta})$, written $f_H(\theta) \succeq_C f_I(\theta)$. In particular, the associated cumulative distribution functions satisfy $F_I(\theta|\theta > t) > F_H(\theta|\theta > t)$, for all $t \in (\underline{\theta}, \bar{\theta})$ and $\theta \in (t, \bar{\theta})$. To capture the fact that any single university admits a negligible portion of the entire pool of students, we assume there is a continuum of each type of university. The total measure of students is normalized to one, and measure $\alpha \in (0, 1)$ of students attend type H universities.

A student is distinguished by his (a) university type, (b) academic ability, θ, and (c) social skill, μ. Social skills, μ, are distributed according to the density $g(\cdot)$ and distribution $G(\cdot)$ with nonnegative full interval support, and are distributed independently of academic skills. We assume that the distribution of social skills is the same at all schools. This assumption reflects the observation that universities largely filter students via high school academic performance and academic tests such as the SAT or ACT. We make the standard increasing hazard rate assumption on g and f_u.

Both academic ability and social skills contribute to the work productivity of a student. Via job interviews, firms can observe μ, but they do not directly observe θ. There are two types of jobs. There is a positive measure Γ of "good" jobs in which the product of a student with ability (θ, μ) is $S\theta\mu$, and many "bad" jobs in which the product is $s\theta\mu$, where $S > s > 0$. Our qualitative findings largely extend to the class of technologies in which a worker's output is proportional to $\theta^\alpha \mu$ for some $\alpha > 0$, reflecting that we do not impose strong structure on the distributions of θ and μ.

Universities know the academic abilities of their students, but not their social skills.[2] Their problem is to assign a grade $g \in \{A, B\}$ to each student.[3] Universities seek to maximize the expected sum of wages earned by graduates.

Firms make students competitive wage offers that earn firms zero expected profits given their information. Firms do not observe student academic abilities. However, firms know the university that each student attended and the distribution of academic abilities at each school, and hence can extract information about academic abilities from grades. We assume that universities adopt grading strategies that take the form of a cutoff, so university u gives a student an "A" if and only if his academic ability θ exceeds a cutoff $\hat{\theta}_u$ chosen by the university. The same equilibrium outcome would obtain were the labels "A" and "B" reversed: we adopt the convention that an "A" grade refers to the better subpopulation of students. In addition, since giving all students "A" grades is equivalent to giving all students "B" grades, without loss of generality, we assume that if it is optimal for a university to give all students the same grade, then it gives all students "A" grades, as at Doonesbury's fictional Walden University.

More generally, our model is sufficiently sparse that equilibria can exist in which grading strategies do not take a cutoff form. Non-cutoff strategies can emerge in equi-

[2]Equivalently, universities could treasure academic integrity so that only θ affects grades.

[3]Lizzeri (1999) argues why universities are not interested in revealing too much information. Dubey and Geanakoplos (2009) suggest a story for why a coarse signal structure might help student motivation. For instance, many graduate programs formulate admission requirements in the form of thresholds, and these thresholds are largely consistent among departments.

librium simply because once firms form beliefs (which determine job assignments), they are not affected by which set of student abilities receive "A"s. However, such equilibria are not robust to natural refinements that pin down what universities do: non-cutoff strategies cannot be part of an equilibrium if either (a) firms observe the true ability of a small measure of students (and schools do not know which ones), from which firms infer average productivities of "A" and "B" students, or (b) employment continues for two periods, and firms learn a worker's true ability after the first period, and there is either complementary learning-by-doing or workers cannot be reassigned. Under such scenarios, universities have strict incentives to ensure their most able students receive "A"s, and hence that equilibrium grading strategies take cutoff forms.

We denote a student from a type u university with grade g as a ug student. The zero profit condition for firms implies that a student who receives a good job receives wage $SE[\theta|g,u,\hat{\theta}_u]\mu$, while a student with a bad job earns wage $sE[\theta|g,u,\hat{\theta}_u]\mu$.[4] The expected ability of a ug student is $E_{ug}[\theta] = \frac{\int_{\underline{\theta}}^{\bar{\theta}} I(\text{grade is g})\theta f_u(\theta)d\theta}{\int_{\underline{\theta}}^{\bar{\theta}} I(\text{grade is g}) f_u(\theta)d\theta}$.[5] Notice that issuing fewer "A"s raises the expected academic ability of *both* "A" *and* "B" students: increasing the grading standard $\hat{\theta}_u$ increases both $E_{uA}[\theta]$ and $E_{uB}[\theta]$. Also, were universities to set a common grading standard, then a type H university would have more "A" students because $f_H(\theta) \succeq_C f_I(\theta)$.

Firms assign a student from university u with grade g and social skills μ to a good job if and only if the student' expected productivity $\mu E_{ug}[\theta]$ exceeds a critical endogenous equilibrium standard, K. That is, K denotes the lowest expected productivity among students employed on good jobs, and $K/E[\theta|u,g,\hat{\theta}_u] \equiv \hat{\mu}_{ug}$ is the minimum level of social skills required from a ug student for placement at a good job. Each university is too small to affect the productivity standard K, but each university internalizes how its grading standard $\hat{\theta}_u$ affects the hiring standard $\hat{\mu}_{ug}$ set by firms.

Given an equilibrium productivity standard K for a good job, a type u university

[4]Much of the literature (e.g., Yang and Yip (2003) and Coate and Loury (1993)), imposes the restriction that everyone at the same job earns the same wage, regardless of their expected productivity.

[5]If the denominator is 0, we set $E_{uB}[\theta] = \underline{\theta}$ and $E_{uA}[\theta] = \bar{\theta}$ to preserve continuity.

maximizes the total income of its student body by choosing $\hat{\theta}_u$ to maximize

$$\pi_u = \max_{\hat{\theta}\in[\underline{\theta},\bar{\theta}]} S\int_{\hat{\mu}_{uA}}^{+\infty}\int_{\hat{\theta}_u}^{\bar{\theta}}\mu\theta dF_u(\theta)dG(\mu) + s\int_0^{\hat{\mu}_{uA}}\int_{\hat{\theta}_u}^{\bar{\theta}}\mu\theta dF_u(\theta)dG(\mu)$$
$$+S\int_{\hat{\mu}_{uB}}^{+\infty}\int_{\underline{\theta}}^{\hat{\theta}_u}\mu\theta dF_u(\theta)dG(\mu) + s\int_0^{\hat{\mu}_{uB}}\int_{\underline{\theta}}^{\hat{\theta}_u}\mu\theta dF_u(\theta)dG(\mu)$$
$$\text{s.t.}\quad \hat{\mu}_{uA}\frac{\int_{\hat{\theta}_u}^{\bar{\theta}}\theta dF_u(\theta)}{\int_{\hat{\theta}_u}^{\bar{\theta}}dF_u(\theta)} = \hat{\mu}_{uB}\frac{\int_{\underline{\theta}}^{\hat{\theta}_u}\theta dF_u(\theta)}{\int_{\underline{\theta}}^{\hat{\theta}_u}dF_u(\theta)} = K.$$

Subtracting the total productivity of all students of university u were they all employed on bad jobs (a constant that does not depend on the grading standard) from π_u, we can rewrite the university's objective as

$$(S-s)\int_{\hat{\mu}_{uA}}^{+\infty}\int_{\hat{\theta}_u}^{\bar{\theta}}\mu\theta dF_u(\theta)dG(\mu) + (S-s)\int_{\hat{\mu}_{uB}}^{+\infty}\int_{\underline{\theta}}^{\hat{\theta}_u}\mu\theta dF_u(\theta)dG(\mu).$$

Dividing this result by $S-s>0$ yields:

$$\Pi_u = \max_{\hat{\theta}\in[\underline{\theta},\bar{\theta}]}\int_{\hat{\mu}_{uA}}^{+\infty}\int_{\hat{\theta}_u}^{\bar{\theta}}\mu\theta dF_u(\theta)dG(\mu) + \int_{\hat{\mu}_{uB}}^{+\infty}\int_{\underline{\theta}}^{\hat{\theta}_u}\mu\theta dF_u(\theta)dG(\mu).$$

Thus, a university maximizes the total wage bill of its student body by maximizing the total product of all of its students who are employed at good jobs.

An immediate implication is that some "A" students from each university always receive good jobs. A student only receives a good job if his expected productivity exceeds the endogenous level K (associated with measure Γ of students receiving good jobs). The common support assumptions on academic abilities and social skills ensures that both universities have some of the most able students. By setting a sufficiently high grading standard, a university can ensure that its "A" students have productivities arbitrarily close to $\bar{\theta}$, and some of these students will also have high social skills and hence receive good jobs.

2.1 Equilibrium

A symmetric pure strategy equilibrium is a collection of grading standards $(\hat{\theta}^*_H, \hat{\theta}^*_I)$, social skill cutoffs, $\hat{\mu}^*_{ug}$, $u \in \{H, I\}$, $g \in \{A, B\}$, and minimal productivity standard, K^*, such that:

- $\hat{\theta}^*_u$ maximizes the total productivity of students at a type $u \in \{H, I\}$ university who receive good jobs given productivity standard K^*;

- Firms assign good jobs to maximize profit, $\hat{\mu}^*_{ug} E_{ug}[\theta|\hat{\theta}^*_u] = K^*$ if some, but not all, type ug students receive good jobs, $\underline{\mu} E_{ug}[\theta|\hat{\theta}^*_u] \geq K^*$ if all receive good jobs and $\bar{\mu} E_{ug}[\theta|\hat{\theta}^*_u] < K^*$ if none do.

- K^* "clears" the market: given $\hat{\theta}^*_H$, $\hat{\theta}^*_I$ and $\hat{\mu}^*_{ug}$, the measure of students with expected productivity of at least K^* is Γ:

$$\alpha \left[(1 - G(\hat{\mu}^*_{HA})) \left(1 - F_H(\hat{\theta}^*_H)\right) + (1 - G(\hat{\mu}^*_{HB}))\, F_H(\hat{\theta}^*_H) \right] +$$
$$+ (1-\alpha) \left[(1 - G(\hat{\mu}^*_{IA})) \left(1 - F_I(\hat{\theta}^*_I)\right) + (1 - G(\hat{\mu}^*_{IB}))\, F_I(\hat{\theta}^*_I) \right] = \Gamma. \quad (2.1)$$

The next proposition establishes the existence of a symmetric equilibrium (i.e., each type u school sets the same grading standard). Optimization by schools further pins down the product of students at good jobs—each university will choose a grading standard that maximizes the productivity of students who receive good jobs.

Proposition 1 *A pure strategy symmetric equilibrium exists. There is a unique equilibrium productivity standard K for a good job, and the expected product of students from school type u who receive good jobs in equilibrium is unique.*

Uniqueness of the grading standard is not guaranteed absent assumptions on the measure of good jobs. For example, if there were so many good jobs that every student from a type H university receives one, then a type H university could achieve this either by giving all students "A"s, or by giving a very few students "A"s, so that the expected productivity of its "B" students was high enough that they receive good jobs.

3 No Heterogeneity in Social Skills

We begin by analyzing the special case in which all students have the same social skills, $\mu = 1$. To solve for equilibrium, first notice that it is never an equilibrium for some, but not all, students with grade g from university u to get good jobs. Were this so, a university could marginally increase its grading standard, raising the average ability of both its "A" and "B" students. In turn, the expected productivity of all students with grade g is raised, so that all now receive good jobs, increasing total productivity of u alumni at good jobs. A direct implication is that in equilibrium with $\Gamma \in (0, 1)$, either only "A" students receive good jobs, or type H universities place all students on good jobs, whereas type I universities only place "A" students. Moreover, for an equilibrium to exist in which "B" students receive good jobs, Γ must exceed the measure α of students at H universities. Therefore, as long as $\Gamma \leq \alpha$, only "A" students receive good jobs in equilibrium when there is no variation in social skills.

So consider an equilibrium in which only "A" students receive good jobs. Were, say, HA students to have higher expected academic productivities than IA students, then a type H university could lower its standard for an "A" slightly, and more of its students would then receive good jobs. Therefore, in equilibrium, the expected productivity of "A" students at the two types of schools must be equal. Combined with the capacity constraint, this implies that equilibrium is fully described by:

$$\frac{\int_{\hat{\theta}_H^*}^{\bar{\theta}} \theta f_H(\theta) d\theta}{\int_{\hat{\theta}_H^*}^{\bar{\theta}} f_H(\theta) d\theta} = \frac{\int_{\hat{\theta}_I^*}^{\bar{\theta}} \theta f_I(\theta) d\theta}{\int_{\hat{\theta}_I^*}^{\bar{\theta}} f_I(\theta) d\theta} \tag{3.1}$$

$$\alpha \int_{\hat{\theta}_H^*}^{\bar{\theta}} f_H(\theta) d\theta + (1 - \alpha) \int_{\hat{\theta}_I^*}^{\bar{\theta}} f_I(\theta) d\theta = \Gamma. \tag{3.2}$$

The next result establishes that under a weak sufficient condition, type H universities set slacker grading standards than type I universities.

Proposition 2 *Suppose that Γ is small enough that in equilibrium some H students do not receive good jobs. Then, in equilibrium, $\hat{\theta}_H^* < \hat{\theta}_I^*$.*

Proposition 2 implies that there is a positive mass of students at type H universities who receive "A"s, but whose ability is low enough that they would receive "B"s at a type I university. This asymmetry simply says that better universities have incentives to dilute the mass of good "A" students with students who would receive "B"s elsewhere: even after dilution, the better upper tail of good students at type H

universities makes an average "A" student at a type H university as good as a typical "A" student from a type I university.

The result that $\hat{\theta}^*_H < \hat{\theta}^*_I$ holds *unless every* student at a type H university is assigned to a good job. In particular, $\Gamma \leq \alpha$ is far from necessary for this result, as in equilibrium a positive measure of students from type I universities receive "A" grades.

Social Planner. A social planner can set and enforce grading standards at schools. She seeks to maximize total output in the economy, setting grading standards to solve

$$\max_{\hat{\theta}_H, \hat{\theta}_I} \alpha \int_{\hat{\theta}_H}^{\bar{\theta}} \theta dF_H(\theta) + (1-\alpha) \int_{\hat{\theta}_I}^{\bar{\theta}} \theta dF_I(\theta) \tag{3.3}$$

$$\text{s.t. } \alpha \int_{\hat{\theta}_H}^{\bar{\theta}} dF_H(\theta) + (1-\alpha) \int_{\hat{\theta}_I}^{\bar{\theta}} dF_I(\theta) = \Gamma. \tag{3.4}$$

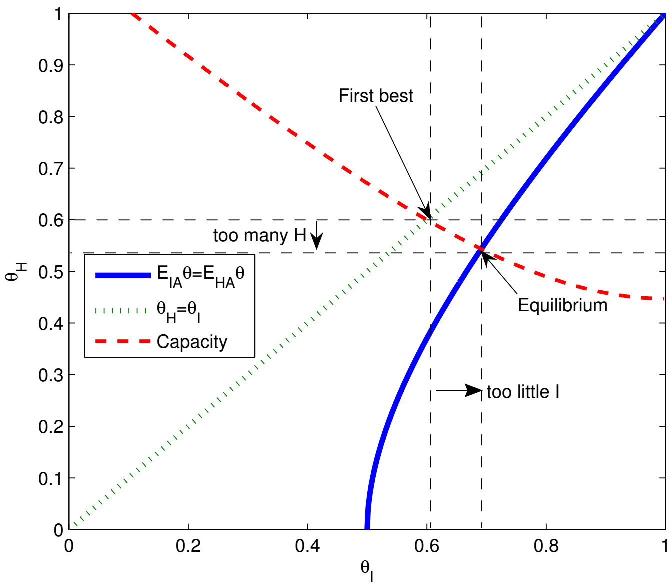
Proposition 3 *The social planner sets a common grading standard,* $\hat{\theta}^{\mathcal{P}}_H = \hat{\theta}^{\mathcal{P}}_I = \hat{\theta}^{\mathcal{P}}$.

When there is no heterogeneity in social skills, no inefficiencies are created by the fact that with the same grading standard, the expected productivity of "A" students is higher at a type H university: the "right" students receive good job assignments.

Corollary 1 *Suppose* Γ *is small enough that not all* H *students receive good jobs. Then* $\hat{\theta}^*_H < \hat{\theta}^{\mathcal{P}} < \hat{\theta}^*_I$*: in equilibrium, too many* H *students and too few* I *students receive good jobs.*

The corollary follows since the good jobs capacity is exhausted in both the social optimum and equilibrium, so both $\hat{\theta}^*_H < \hat{\theta}^*_I \leq \hat{\theta}^{\mathcal{P}}$ and $\hat{\theta}^{\mathcal{P}} \leq \hat{\theta}^*_H < \hat{\theta}^*_I$ cannot occur.

Figure 1 conveys the intuition, depicting equilibrium and social planner outcomes. CFOSD implies that the solid line of equal expected productivities lies below the 45° line associated with equal grading standards; Proposition 2 says that the intersection of the dashed capacity constraint line and the solid equal productivity line is below the equal grading standards line; and the capacity constraint must be negatively sloped. Strategic considerations not only induce better universities to set slacker grading standards than is socially optimal, but they also force lesser universities to set stricter



Note: $\Gamma = 0.25$, $\alpha = 0.5$. $f_H(x) = 2x$, $f_I(x) = 2 - 2x$, $[\underline{\theta}, \bar{\theta}] = [0, 1]$.

Figure 1: Comparison of Equilibrium and Social Planner outcomes.

standards in order to compete. Still, type I universities prefer this outcome to one in which no grades are disclosed, in which case no I alumni would receive good jobs.

We next characterize how the primitives of our economy affect outcomes.

More type H universities. An increase in the fraction α of top schools flattens the dashed capacity constraint line because changes in grading standards at type H universities now have bigger impacts on the measure of students with As. There are conflicting effects: increasing α shifts the composition of schools from those with high standards to those with low standards; but increasing α increases the supply of able students. To determine whether grading standards improve, observe that a marginal increase in α is a (not necessarily shape preserving) counterclockwise rotation of the capacity line, and one point of the line remains the same. Whenever this point is above the 45° line, the intersection with the (blue) equal productivity line shifts up and grading standards rise. Moreover, when $\alpha = \frac{1}{2}$, the rotation point is above 45° line, since $1 - F_H(t) > 1 - F_I(t)$. Thus, there exists a $\bar{\alpha} < \frac{1}{2}$ such that for $\alpha > \bar{\alpha}$, increasing the fraction of type H schools raises grading standards. That is, the supply effect dominates the composition effect when there are enough top universities.

Improvements in student body composition. Improving the distribution of student abilities at a type I university causes type H universities to set higher standards for an "A", but has ambiguous effects on I's grading standard. To see this, let $E_u[x] = \frac{\int_x^{\bar{\theta}} \theta dF_u}{\int_x^{\bar{\theta}} dF_u}$, and $m_u(x) = \int_x^{\bar{\theta}} dF_u$. Consider an improvement in I's distribution that corresponds to a weighted average of the ability distributions,

$$F_\lambda(x) = P(\theta < x|\lambda, I) = \lambda F_H(x) + (1-\lambda)F_I(x). \tag{3.5}$$

Then $E[\theta|\lambda, I, \theta > x] = \lambda E_H[x] + (1-\lambda)E_I[x]$, and $\int_x^{\bar{\theta}} dF_\lambda = \lambda m_H(x) + (1-\lambda)m_I(x)$. Let $(\theta_H^*(\lambda), \theta_I^*(\lambda))$ denote the equilibrium pair of grading standards when the ability distribution at a type I university is F_λ for $\lambda \in [0,1]$, where $(\theta_H^*, \theta_I^*) = (\theta_H^*(0), \theta_I^*(0))$. Differentiating the equilibrium conditions with respect to λ at $\lambda = 0$ yields

$$\begin{cases} \overbrace{E_H'[\theta_H^*]}^{\triangleq A > 0}\ \theta_H^*(0)' - \overbrace{E_I'[\theta_I^*)]}^{\triangleq B > 0}\ \theta_I^*(0)' = \overbrace{E_H[\theta_I^*] - E_I[\theta_I^*]}^{\triangleq C > 0} \\ \underbrace{f_H(\theta_H)}_{\triangleq D > 0}\ \theta_H^*(0)' + \underbrace{f_I(\theta_I)}_{\triangleq E > 0}\ \theta_I^*(0)' = \underbrace{m_H(\theta_I^*) - m_I(\theta_I^*)}_{\triangleq F > 0} \end{cases} \Rightarrow$$

$$\theta_H^*(0)' = \frac{CE + BF}{AE + BD} > 0 \qquad \theta_I^*(0)' = \frac{AF - CD}{AE + BD} \lesseqgtr 0.$$

Because the distribution of abilities at type H universities dominates that at type I universities, the parameters C and F are positive. Intuitively, improving the distribution at type I schools creates added "competition" for type H schools forcing them to raise standards. The ambiguous impact on type I schools reflects that (a) type I schools can lower grading standards and still have a higher average quality of "A" students, but (b) the better distribution also increases competition for type I schools, raising the average ability required for a good job placement. Analogously, one can show that a deterioration in the distribution of abilities at type H universities eases grading standards at type I universities, but has ambiguous effects at type H schools.

Good Jobs, Grading Standards and Grade Inflation. We next characterize how increases in the number of good jobs affect equilibrium outcomes. It follows directly that increasing Γ causes all universities to lower grading standards: this reduction in the average quality of "A" students implies there is grade inflation. In particular, increasing Γ shifts the capacity line outward away from the top right corner, shifting equilibrium outcomes away along the "equal expectations about "A" students" locus line. The social planner's choice shifts away along the 45° line, resulting in an equal decrease in grading standards and an increase in "A" grades at both universities.

We are especially interested in identifying when equilibrium grading standards fall by more at top schools, and when this translates into higher grade inflation at top schools. Define $Q_u(t) = E(\theta|u, \theta > t)$, where $Q(\bar{\theta}) = \bar{\theta}$ preserves continuity. We characterize the relative impact of Γ on grading standards via the implicit function $\hat{\theta}_H(\hat{\theta}_I)$ defined by $Q_H(\hat{\theta}_H) = Q_I(\hat{\theta}_I) = K$, by varying K (K falls with Γ): grading standards fall faster at type H schools than type I schools if and only if $\hat{\theta}'_H(\hat{\theta}_I) > 1$.

Proposition 4 *Suppose there are sufficiently few good jobs, Γ. Then a slight increase in Γ causes grading standards to fall faster at type H schools than type I schools.*

One would like to extend this result to settings where the number of good jobs is larger, maintaining only the premise that Γ is not so large that all H students receive good jobs. To do this, we consider a family of ability densities with linear right tails, where the linear right tail is "long" enough that it describes the abilities of A students:

$$f_u(\theta|a_u, b_u) = a_u + b_u\theta, \ \theta \in [t, 1] \text{ for some } t, \tag{3.6}$$

where, to ease presentation, we assume a $[0, 1]$ support (extensions to a support $[\underline{\theta}, \bar{\theta}]$ are routine). Positivity of the density implies that $a_u + b_u > 0$ and $a_u + b_u t > 0$, and $f_H(\cdot) \succeq_C f_I(\cdot)$ implies that $a_H b_I < a_I b_H$. We also need that b_H *sufficiently exceeds* b_I. When densities are linear on their full support, $b_H > b_I$ suffices.

Proposition 5 *Consider densities with linear right tails, where b_H sufficiently exceeds b_I, so that $a_H b_I(1+\hat{\theta}_I) < a_I b_H(1+\hat{\theta}_H)$. Then if $\hat{\theta}_u > t$, an increase in Γ causes grading standards to fall faster at type H universities than type I universities.*

Corollary 2 *If $f_H(\hat{\theta}_H) \geq f_I(\hat{\theta}_I)$ and grading standards fall faster in H, then the number of "A"s increases faster at type H universities than type I universities.*

For example, if $f_H(t)/f_I(t)$ increases in t, then there exists a $\tilde{t}$ such that $f_H(t') > f_I(t'), \forall t' > \tilde{t}$, in which case Corollary 2 follows if and only if there are sufficiently few good jobs, Γ.

In sum, when students only differ in academic ability, universities with better student bodies press some students from lesser universities out of good job assignments by setting slacker grading standards than is socially optimal. Further, under plausible scenarios, more good jobs causes greater grade inflation at top universities. We now explore how heterogeneity in social skills affects these conclusions.

4 Heterogeneous Social Skills

Suppose now that students differ in their social skills. For heterogeneity in social skills to alter outcomes, there must be a sufficient difference between the highest and lowest social skill that not all "A" students receive good job assignments. If not, then our previous analysis characterizes outcomes. We maintain the assumption that the dispersion is not sufficient for "B" students to receive good job assignments in equilibrium. Thus, the support $[\underline{\mu}, \bar{\mu}]$ of the distribution of social skills $G(\cdot)$ is neither very small, nor very large; i.e., heterogeneity in social skills is "intermediate" so that $K^* > max\{\underline{\mu} E_u[\theta|\theta > \hat{\theta}_u^*],\ \bar{\mu} E[\theta|\theta < \hat{\theta}_u^*]\}$.

Proposition 6 *Suppose $\mu g(\mu)$ is increasing in μ and heterogeneity in social skills is intermediate. Then, in the unique equilibrium, type I schools set higher grading standards than type H schools. Further, $\hat{\mu}_H^* \leq \hat{\mu}_I^*$, where the inequality is strict as long as some A students do not receive good jobs.*

The result says that as long as the density over social skills does not fall quickly, $g'(\mu) > -\frac{g(\mu)}{\mu}$, then top universities set slacker grading standards. When Γ is small enough that in equilibrium not all students at a school receive "A"s, and the dispersion in social skills is sufficient that an "A" student with the lowest social skill is not offered a good job, then equilibrium is characterized by interior solutions, and a university's best response is characterized by the intersection of

$$\hat{\theta}_u = K \frac{\frac{g(\hat{\mu}_u)}{1-G(\hat{\mu}_u)}\hat{\mu}_u}{E[\mu|\mu > \hat{\mu}_u] + \frac{g(\hat{\mu}_u)}{1-G(\hat{\mu}_u)}\hat{\mu}_u^2} \equiv R(\hat{\mu}_u)$$

$$\hat{\mu}_u E_u[\theta|\theta > \hat{\theta}_u] = K \Leftrightarrow \hat{\mu}_u Q(\hat{\theta}_u) = K.$$

The bottom efficient job assignment equation is decreasing in $(\hat{\mu}, \hat{\theta})$ space (see Figure 2). The CFOSD assumption implies that $E_I[\theta|\theta > \hat{\theta}] < E_H[\theta|\theta > \hat{\theta}]$, for every $\hat{\theta}$. Therefore, $\hat{\mu}^*_{HA} < \hat{\mu}^*_{IA}$. Since $\mu g(\mu)$ is increasing in μ, the right-hand side of the top equation is increasing in μ, which combined with $\hat{\mu}^*_{HA} < \hat{\mu}^*_{IA}$, implies $\hat{\theta}^*_H < \hat{\theta}^*_I$.

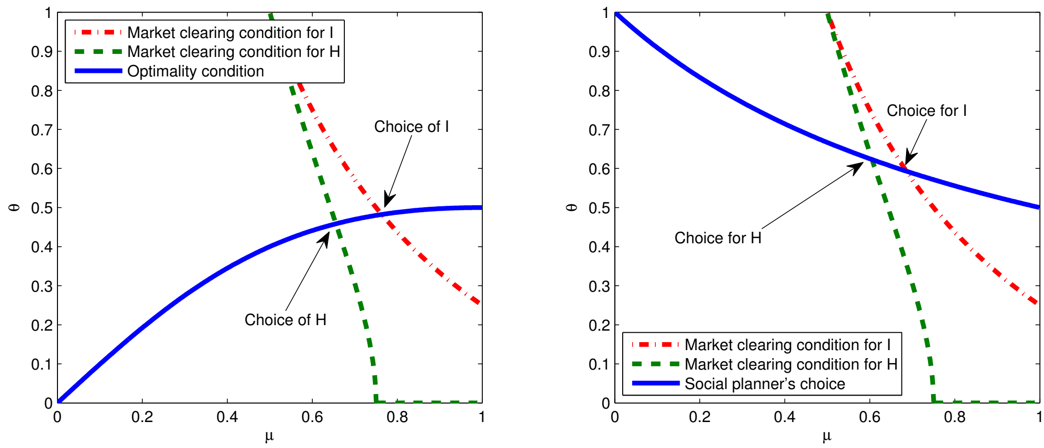


Figure 2: Equilibrium and Social Planner outcomes. μ is uniformly distributed on $[0, 1]$, $f_H(\theta) = 2\theta$, $f_I(\theta) = 2 - 2\theta$, $[\underline{\theta}, \bar{\theta}] = [0, 1]$.

Corollary 3 *In any interior equilibrium,* $E[\theta|H, \theta > \hat{\theta}^*_H] > E[\theta|I, \theta > \hat{\theta}^*_I]$.

The corollary follows directly from optimal job assignment, $\hat{\mu}^*_{HA} E[\theta|H, \theta > \hat{\theta}^*_H] = \hat{\mu}^*_{IA} E[\theta|I, \theta > \hat{\theta}^*_I]$; and Proposition 6, which states that $\hat{\mu}^*_{HA} < \hat{\mu}^*_{IA}$. Corollary 3 says that otherwise identical students from H and from I receive different wages at good jobs: in particular, students from lesser universities receive lower wages.

We now derive the qualitative impact of intermediate levels of heterogeneity in social skills on the social planner's grading standards.

Proposition 7 *Suppose heterogeneity in student social skills is intermediate, so that not all "A" students receive good jobs, and no "B" students receive good jobs. Then the social planner sets* $\hat{\theta}_H^{\mathcal{P}} > \hat{\theta}_I^{\mathcal{P}}$ *and* $\hat{\mu}_{HA}^{\mathcal{P}} < \hat{\mu}_{IA}^{\mathcal{P}}$.

The social planner sets $\hat{\mu}_{HA}E[\theta|\theta > \hat{\theta}_H] = \hat{\mu}_{IA}E[\theta|\theta > \hat{\theta}_I]$ and $E[\mu|\mu > \hat{\mu}_{HA}]\hat{\theta}_H = E[\mu|\mu > \hat{\mu}_{IA}]\hat{\theta}_I$. The first equality confirms that a social planner sets the same standards on the social skills of students assigned to good jobs as competitive firms: each marginal student has the same expected productivity given his social skills and the information contained in an "A" grade from his university. As a result, even though the distributions of social skills at universities are the same, the average social skill of "A" students from a type H university who receive good job assignments is less. The second equality says that the social planner sets grading standards to equate the expected productivities of the marginal "A" student at each university. Then, because the average social skill of students from a type H university with good jobs is less, a social planner sets a more demanding grading standard at type H universities.

Heterogeneity across universities in student body compositions leads to worse equilibrium outcomes than were there a common ability distribution at schools. One might conjecture that this heterogeneity also hinders the social planner because it causes firms to distort hiring decisions toward students from better universities. This conjecture is false: homogeneity harms a planner's ability to distinguish better populations of students. To see this, note that with heterogeneous distributions, the social planner *could* set common grading and social skill cutoffs; however, the planner chooses not to. It follows that homogenizing the university pool is suboptimal.

In sum, as long as there is not so much heterogeneity in social skills that some "B" students receive good jobs, then under mild conditions, better universities set lower grading standards, even though a social planner would make the opposite choice. These findings extend when some "B" students with exceptional skills also receive good jobs, as long as the distribution of social skills is such that "A" students dominate decision making. We believe that this is the relevant real world scenario. However, if, for example, $g(\mu)$ is sufficiently flat with sufficient dispersion, then type H schools may weigh the job prospects of "B" students by enough that they set higher grading standards, in order to raise the fraction of “B” students that receive good jobs.

5 Conclusion

The central message of this paper is that competition for good job assignments for graduates causes better universities to set lower standards for “A”s, because their marginal “A” students can ride on the coat tails of the better average qualities of “A” students. We show that a social planner sets the opposite ordering on grading standards. We also show that increases in the number of good jobs drives down standards for “A”s, and that under plausible scenarios, standards fall more at better schools.

Although our setting features just two types of universities, *none* of our analysis hinges on this modeling choice. To see that our results extend as long as university types can be CFOSD-ordered by their distributions over academic abilities, note that additional types enter via the equilibrium standard for a good job assignment, and our analysis simply establishes that better schools set lower grading standards. In particular, our analysis extends when there is a continuum of different university types. So, too, our analysis does not hinge on our common support assumption, as long as some students at a lesser schools get good jobs.

One feature that we do not integrate is to our model is effort, and how grading standards affect student effort, where effort both affects academic performance and on-the-job productivity. In such a setting, universities must account for how their grading standards affect effort choices, and equilibrium job assignment. Lower standards for “A” grades may induce students to exert less effort (especially if the density of ability for the marginal student is low), and universities will internalize this effect. One can clearly provide conditions under which endogenizing effort does not reverse the conclusion that better universities set lower grading standards. More generally, a thorough analysis of effort and grading is an interesting topic for future research.

6 Appendix

Proof of Proposition 1. Substitute the market-clearing conditions into the maximized objectives of the universities. Each university maximizes a continuous function on the convex set $[\underline{\theta}, \bar{\theta}]$. Therefore, by Berge's maximum theorem, there is an upper hemicontinuous best-response relation $\theta_u^*(K)$. Labor-market clearing implies that $K^*(\theta_H^*, \theta_I^*)$ is a continuous function of its arguments, since the densities are positive on their support. Substituting the best responses of the two types of universities into the market-clearing condition yields a upper hemicontinuous correspondence $K^*(K)$, defined on $[\underline{\theta}\underline{\mu}, \bar{\theta}\bar{\mu}]$ to itself, which has a fixed point by Kakutani's theorem.

Now suppose there were multiple equilibrium standards, K_1, K_2, for good job assignments, with $K_1 > K_2$. Then, choosing the same grading standard facing K_2, each university can place more students at good jobs. But then the labor markets cannot clear for both K_1 and K_2. Further, facing a single equilibrium standard K, the measure of students from type u university is pinned down by optimization—each university chooses a grading standard that maximizes the expected product of those receiving good jobs. $\square$

Proof of Proposition 2. Suppose only students with "A" grades receive good jobs. Equilibrium requires $E_{HA}[\theta] = E_{IA}[\theta]$. Define $Q_U(x) = \frac{\int_x^{\bar{\theta}} t f_U(t)dt}{\int_x^{\bar{\theta}} f_U(t)dt}$ for $x \in [\underline{\theta}, \bar{\theta})$, with $Q(\bar{\theta}) = \bar{\theta}$, to be the expected productivity of "A" students given any standard x; $Q_U(x)$ is trivially strictly increasing in x. By CFOSD, $Q_H(x) > Q_I(x)$ for all $x \in [\underline{\theta}, \bar{\theta})$, so the value y defined by $Q_H(y) = Q_I(x)$ is less than x. $\square$

Proof of Proposition 3. Because "A" students are hired before "B" students, the social planner gives "A" to all students who, in her opinion, should be employed on a good job, and the labor market assigns good jobs only to "A" students. The first-order conditions to the social planner's problem are:

$$\begin{aligned} -\alpha\hat{\theta}_H f_H(\hat{\theta}_H) + \alpha\lambda f_H(\hat{\theta}_H) &= 0 \\ -(1-\alpha)\hat{\theta}_I f_I(\hat{\theta}_I) + (1-\alpha)\lambda f_I(\hat{\theta}_I) &= 0. \end{aligned}$$

By the full support assumption, the densities are positive, so the first-order conditions simplify to $\hat{\theta}_H^{\mathcal{P}} = \lambda$ and $\hat{\theta}_I^{\mathcal{P}} = \lambda$. $\square$

Proof of Proposition 4. Observe that $Q_H^{-1}(\bar{\theta}) = Q_I^{-1}(\bar{\theta}) = \bar{\theta}$, and CFOSD implies that $Q_H^{-1}(\bar{\theta} - \varepsilon) > Q_I^{-1}(\bar{\theta} - \varepsilon)$ for any ε positive, but sufficiently small. Thus, $Q_H'(\bar{\theta}) < Q_I'(\bar{\theta})$. By the implicit function theorem, $\hat{\theta}_H'(\hat{\theta}_I \mid \hat{\theta}_I = \bar{\theta}) = \frac{Q_I'(\bar{\theta})}{Q_H'(\bar{\theta})} > 1$. Continuity of $\hat{\theta}_H'(\hat{\theta}_I)$ ensures that $\hat{\theta}_H'(\hat{\theta}_I) > 1$ over some non-degenerate interval, $[\tilde{\theta}_I, \bar{\theta})$. The result follows. $\square$

Proof of Proposition 5. By Proposition 4, there is an interval $[\tilde{\theta}_I, 1]$ where $\hat{\theta}_H'(\hat{\theta}_I)$ is at least 1, so that $\hat{\theta}_H'(\hat{\theta}_I) = 1$ when $\hat{\theta}_I = \tilde{\theta}_I$. Thus, for some $\hat{\theta}_H = \hat{\theta}_H(\tilde{\theta}_I)$ and $\hat{\theta}_I = \tilde{\theta}_I$, $Q_H(\hat{\theta}_H) = Q_I(\hat{\theta}_I)$ and $Q_H'(\hat{\theta}_H) = Q_I'(\hat{\theta}_I)$. The second derivative of $\hat{\theta}_H(\hat{\theta}_I)$ is

$$\frac{\partial^2 \hat{\theta}_H}{\partial \hat{\theta}_I^2} = \frac{Q_I''(\hat{\theta}_I)Q_H'(\hat{\theta}_H) - Q_I'(\hat{\theta}_I)Q_H''(\hat{\theta}_H)\frac{\partial \theta_H}{\partial \theta_I}}{(Q_H'(\hat{\theta}_H))^2} = \frac{Q_I''(\hat{\theta}_I)(Q_H'(\hat{\theta}_H))^2 - Q_H''(\hat{\theta}_H)(Q_I'(\hat{\theta}_I))^2}{(Q_H'(\hat{\theta}_H))^3}.$$

Thus, $Q_H'(\hat{\theta}_H) = Q_I'(\hat{\theta}_I)$, $\hat{\theta}_H(\hat{\theta}_I)$ is concave if $Q_I''(\hat{\theta}_I) < Q_H''(\hat{\theta}_H)$. We now solve for the shapes of derivatives of $Q(\cdot|\cdot)$ when densities have linear right tails:

$$Q(t|a, b) = \frac{2}{1-t} \frac{\frac{a}{2}(1-t^2) + \frac{b}{3}(1-t^3)}{b(1+t) + 2a},$$

$$Q'(t|a, b) = \frac{2(a+bt)(3a+2b+bt)}{3(2a+b+bt)^2} = \frac{2}{3}\left(1 - \frac{(a+b)^2}{(2a+b+bt)^2}\right),$$

$$Q''(t|a, b) = \frac{4b(a+b)^2}{3(2a+b+bt)^3}.$$

Observe that $Q'(\hat{\theta}_H|b_H) = Q'(\hat{\theta}_I|b_I)$ implies $\frac{a_H+b_H}{2a_H+b_H+b_H\theta_H} = \frac{a_I+b_I}{2a_I+b_I+b_I\theta_I}$, which in turn combined with $a_H b_I(1+\hat{\theta}_I) < a_I b_H(1+\hat{\theta}_H)$ implies $\frac{b_H}{b_H(1+\hat{\theta}_H)+2a_H} > \frac{b_I}{b_I(1+\hat{\theta}_I)+2a_H}$. Multiply both sides of this inequality of derivatives by $\frac{a_H+b_H}{b_H(1+\hat{\theta}_H)+2a_H}$ and $\frac{b_I+a_I}{b_I(1+\hat{\theta}_I)+2a_I}$, respectively, and multiply by $\frac{4}{3}$:

$$\underbrace{\frac{4b_H(a_H+b_H)^2}{3(2a_H+b_H+b_H\hat{\theta}_H)^3}}_{Q''(\hat{\theta}_H|a_H,b_H)} > \underbrace{\frac{4b_I(a_I+b_I)^2}{3(2a_I+b_I+b_I\hat{\theta}_I)^3}}_{Q''(\hat{\theta}_I|a_I,b_I)}.$$

Therefore, $\theta_H(\theta_I)$ is concave at $\tilde{\theta}$. Since $\frac{\partial \hat{\theta}_H}{\partial \theta_I} > 1$ for every $\hat{\theta}_I > \tilde{\theta}_I$, concavity of $\theta_H(\theta_I)$ at $\tilde{\theta}$ contradicts $\theta_H'(\tilde{\theta}_I) = 1$. $\square$

Proof of Corollary 2. $-f_u(\theta_u)d\hat{\theta}_u$ is the increase in "A"s at university u responding to $d\Gamma$. As increasing Γ causes grading standards to fall and $\frac{d\hat{\theta}_H}{d\hat{\theta}_I} > 1$, the result follows. $\square$

Proof of Proposition 6. University u solves:

$$\max_{\hat{\theta}_u,\hat{\mu}_u} \int_{\hat{\mu}_u}^{\bar{\mu}} \int_{\hat{\theta}_u}^{\bar{\theta}} \theta\mu dF_u(\theta)dG(\mu), \tag{6.1}$$

$$\text{s.t. } \hat{\mu}_u E_u[\theta|\theta > \hat{\theta}_u] = K. \tag{6.2}$$

The associated first-order conditions for interior solution are:

$$-g(\hat{\mu}_u)\hat{\mu}_u \int_{\hat{\theta}_u}^{\bar{\theta}} \theta dF_u(\theta) + \lambda E_u[\theta|\theta > \hat{\theta}_u] = 0,$$

$$-f_u(\hat{\theta}_u)\hat{\theta} \int_{\hat{\mu}_u}^{\bar{\mu}} \mu dG(\mu) + \lambda\hat{\mu}_u \frac{f_u(\hat{\theta}_u)}{1-F_u(\hat{\theta}_u)} \left(E_u[\theta|\theta > \hat{\theta}_u] - \hat{\theta}_u\right) = 0.$$

The densities are positive everywhere. Integrating and rearranging terms yields

$$E[\mu|\mu > \hat{\mu}_u]\hat{\theta}_u = \frac{g(\hat{\mu}_u)}{1-G(\hat{\mu}_u)}\hat{\mu}_u^2 \left(\frac{K}{\hat{\mu}_u} - \hat{\theta}_u\right).$$

Solve for $\hat{\theta}_u$:

$$\hat{\theta}_u = \frac{K\frac{g(\hat{\mu}_u)}{1-G(\hat{\mu}_u)}\hat{\mu}_u}{E[\mu|\mu > \hat{\mu}_u] + \frac{g(\hat{\mu}_u)}{1-G(\hat{\mu}_u)}\hat{\mu}_u^2} \equiv R(\hat{\mu}_u). \tag{6.3}$$

$R(\cdot)$ is an increasing function since $g(\mu) + \mu g'(\mu)$ is positive for all μ in the support:

$$sign\frac{\partial R(x)}{\partial x} = sign\left\{\int_x^{\bar{\mu}} \mu\frac{g(\mu)}{g(x)}d\mu + g'(x)x\int_x^{\bar{\mu}} \mu\frac{g(\mu)}{(g(x))^2}d\mu\right\} = sign\left\{(xg(x))'\right\}.$$

Equilibrium is governed by equations (6.2) and (6.3). The latter is a decreasing curve in $(\hat{\mu}, \hat{\theta})$ space (see Figure 2) so there is a unique equilibrium described by their intersections. If the optimal solution is on the boundary with $\hat{\mu} = \underline{\mu}$, the optimality condition becomes $R(\underline{\mu}) \geq \hat{\theta}$; if it is at $\hat{\theta} = \underline{\theta}$, the optimality condition becomes $R(\hat{\mu}) \leq \underline{\theta}$.

The CFOSD assumption implies that $E_I[\theta|\theta > \hat{\theta}] < E_H[\theta|\theta > \hat{\theta}]$, for every $\hat{\theta} < \bar{\theta}$. Hence, the efficient job assignment equation (6.2) for a type H school is everywhere to the left of that for a type I school; and since $R(\hat{\mu})$ increases in $\hat{\mu}$, it intersects the job assignment equation for H below that for I, so that $\hat{\mu}^*_{HA} \leq \hat{\mu}^*_{IA}$. Equality only occurs if the intersections occur at $\hat{\mu} = \underline{\mu}$, i.e., all "A" students get good jobs. Further, since $R(\hat{\mu})$ is increasing, $\hat{\mu}^*_{HA} < \hat{\mu}^*_{IA}$ implies $\hat{\theta}^*_H < \hat{\theta}^*_I$. See Figure 2. □

Proof of Proposition 7. The social planner chooses grading standards to maximize expected output subject to the constraint that job placement decisions are based on student social skills and the information contained in grades:

$$\max_{\hat\theta_H,\hat\theta_I,\hat\mu_H,\hat\mu_I} \alpha \int_{\hat\mu_{HA}}^{\bar\mu}\int_{\hat\theta_H}^{\bar\theta} (\mu\theta)\, dF_H(\theta)dG(\mu) + (1-\alpha)\int_{\hat\mu_{IA}}^{\bar\mu}\int_{\hat\theta_I}^{\bar\theta} (\mu\theta)\, dF_I(\theta)dG(\mu)$$

$$\text{s.t. } \alpha \int_{\hat\mu_{HA}}^{\bar\mu}\int_{\hat\theta_H}^{\bar\theta} dF_H(\theta)dG(\mu) + (1-\alpha)\int_{\hat\mu_{IA}}^{\bar\mu}\int_{\hat\theta_I}^{\bar\theta} dF_I(\theta)dG(\mu) = \Gamma.$$

The first-order conditions (assuming an interior solution) for a type u school are:

$$\begin{aligned} \alpha g(\hat\mu_{uA})\hat\mu_{uA}\int_{\hat\theta_u}^{\bar\theta}\theta dF_u &= \alpha g(\hat\mu_{uA})\lambda\int_{\hat\theta_u}^{\bar\theta} dF_u, \\ \alpha\left(\int_{\hat\mu_{uA}}^{1}\mu dG\right)\hat\theta_u f_u \hat\theta_u &= \alpha\lambda\left(\int_{\hat\mu_{uA}}^{1} dG\right) f_u(\hat\theta_u). \end{aligned}$$

The densities and α cancel on both sides. After canceling and isolating λ on the right-hand side, the equations for both types of universities become

$$\hat\mu_u E_u[\theta|\theta > \hat\theta_u] = \lambda, \tag{6.4}$$

$$E[\mu|\mu > \hat\mu_u]\hat\theta_u = \lambda. \tag{6.5}$$

Optimality condition, equation (6.5), is the same for both university types. If the optimal solution is on the boundary, $\hat\mu_u = \underline{\mu}$, the condition is $\hat\theta_u E[\mu] \geq \lambda$; if it is at $\hat\theta_u = \underline{\theta}$, the condition becomes $\underline{\theta}E[\mu|\mu < \hat\mu_u] \leq \lambda$. Denote the implicit function for $\hat\theta_u$ from equation (6.4) by $R_u(\hat\mu)$ and that from equation (6.5) by $Q(\hat\mu)$. Their derivatives are

$$R_u'(x) = -\frac{\lambda}{x^2}\overbrace{\frac{1}{(E_u[\theta|\theta>t])'_{t=\frac{\lambda}{x}}}}^{>1}, \quad Q'(x) = -\frac{\lambda}{(E[\mu|\mu>x])^2}\overbrace{(E[\mu|\mu>x])'_x}^{<1}.$$

Because both $F_u(\cdot)$ and $G(\cdot)$ feature increasing hazards, the derivatives of both expectations are less than 1 (Bagnoli and Bergstrom (2005)). As $E[\mu|\mu > x] > x$ for all $x < \bar\mu$, we have $0 > Q'(x) > R_u'(x)$. By CFOSD, the efficient job assignment equation (6.4) for a type H school is everywhere to the left of that for a type I school. Therefore, the intersection of equations (6.4) and (6.5) in $(\hat\mu, \hat\theta)$ space that determine $(\hat\theta_H, \hat\mu_{HA})$ occur above and to the left of the intersection that determines $(\hat\theta_I, \hat\mu_{IA})$. Thus, $\hat\theta_H^{\mathcal{P}} \geq \hat\theta_I^{\mathcal{P}}$ and $\hat\mu_{HA}^{\mathcal{P}} \leq \hat\mu_{IA}^{\mathcal{P}}$, and $\hat\mu_{HA}^{\mathcal{P}} = \hat\mu_{IA}^{\mathcal{P}}$ occurs only when lower boundary on support of μ binds for both types of universities, so that all "A" students get good jobs.

References

Bagnoli, M. and T. Bergstrom (2005): "Log-concave probability and its applications," *Economic Theory*, 26, 445–469.

Bagues, M., M. S. Labini, and N. Zinovyeva (2008): "Differential Grading Standards and University Funding: Evidence from Italy," *CESifo Economic Studies*, 54, 149–176.

Bar, T. and A. Zussman (2010): "Partisan grading," Working paper.

Betts, J. R. (1998): "The Impact of Educational Standards on the Level and Distribution of Earnings," *The American Economic Review*, 88, 266–275.

Chen, W., L. Hao, and W. Suen (2007): "A Signaling Theory of Grade Inflation," *International Economic Review*, 48, 1065–1090.

Coate, S. and G. C. Loury (1993): "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 83, 1220–1240.

Costrell, R. (1994): "A Simple Model of Educational Standards," *American Economic Review*, 84, 956–971.

Dubey, P. and J. Geanakoplos (2009): "Grading Exams: 100, 99, ..., 1 Or A, B, C? Incentives In Games Of Status," Yale University.

Lizzeri, A. (1999): "Information Revelation and Certification Intermediaries," *The RAND Journal of Economics*, 30, 214–231.

MacLeod, W. B. and M. Urquiola (2009): "Anti-Lemons: School Reputation and Educational Quality," Working Paper 15112, National Bureau of Economic Research.

Rojstaczer, S. (2003): "Where All Grades Are Above Average," *Washington Post*.

Rose, H. and J. R. Betts (2004): "The Effect of High School Courses on Earnings," *Review of Economics and Statistics*, 86, 497–513.

Yang, H. and C. S. Yip (2003): "An Economic Theory Of Grade Inflation," University of Pennsylvania.