University Competition and Grading Standards

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 28, 2009
Not for circulation

Abstract

Some universities have more students with high GPAs than others, and claim that it is because they have better students. We consider a setting in which universities internalize how their grading standards affect the equilibrium wages of graduates, and seek to maximize the total wage bill of their graduates. Universities are distinguished by the distribution of student abilities. We show that in equilibrium, universities with better distributions of students set softer grading standards — the marginal "A" student at better universities is less able than the marginal student at lesser schools. Indeed, better universities set grading standards that are below the social optimum, while worse universities set excessively strict grading standards. Improving the distribution of student abilities at lesser universities causes the better universities to raise their grading standards, but can have an ambiguous effect on grading standards at the lesser school.

1 Introduction

Universities award grades to measure the performance of students in courses. In turn, important decisions by third parties weigh GPAs in their decisions — firms tend to offer higher wages to students with high GPAs, and graduate schools tend to admit high GPA students. The information content of a high GPA depends on how hard it is to earn — the better the ability composition of the student body, and the fewer high grades offered, the more informative and valuable a high GPA is. In this paper, we characterize how universities choose their grading standards when they care about the decisions made by third parties based on GPAs, i.e., when universities care about the job assignments that their graduates obtain, or the schools that they gain admission to. In this context, our goal is to understand how the ability composition of student bodies at different schools affects the equilibrium grading standards that they set.

In particular, we want to glean insights into why better universities award a higher fraction of high grades to their students than others. For example, the GPA at private schools in 2006-2007 is 0.3 higher than at public schools (Rojstaczer (2003)); and Healy (2002) documents that 91% of Harvard students graduated cum laude in 2001 while only 8% did so at Cornell. The top universities would argue that the high proportion of high grades simply reflects the better composition of their student body. In fact, our central result is that better universities strategically set softer grading standards than is socially optimal, while lesser universities compete by setting excessively demanding standards. As a result, the marginal "A" student at a top university is less able than the marginal "A" student at a lesser university.

In our model, firms base job assignments in part on the information contained in a students GPA. The intuition for why top universities set softer grading standards is simple: because third parties cannot disentangle whether a student is from the top or bottom end of those receiving "A" grades, worse students can piggy back on the presence of more top students at their school to get a good job assignment. It is the competition for more good job assignments for graduates that generates this result — while grade inflation at a school (giving more "A" grades), lowers the expected productivity and hence wages of those students who receive good job assignments, additional well-placed students offset that. In contrast, lesser schools must compete for better job assignments by raising the average ability composition of students who receive "A" grades, setting an excessively high grading standard.

***argue less good schools gain more from assigning grades because it lets them distinguish their top students, thereby permitting them to compete against better schools for top jobs.

We want to learn whether giving more As is a consequence of being a better university; and if it is, are there any other consequences. In this sense, the closest paper to ours is probably a free-rider paper of Yang and Yip (2003), since they also predict abundance of good grades in better schools. Their prediction, however, is that the damage to society comes from universities explicitly lying about students' abilities to perform on a job by deviating from honest reports — schools intentionally destroy value. The prediction of their paper that is not entirely clear to us is that all students get the same wage; so essentially there is no point for grading or taking a grade into consideration when employing — is that so? Why grades are needed? How amount of grades is determined by a university? How distribution of students across universities affect relative grading standards? How universities take into account the existence of other universities? To answer these questions, we study the competition of universities on the placement market.

Some universities have more students with higher GPAs than others, and claim that it is because they have a better distribution of students. We consider a setting in which universities internalize how their grading standards affect the equilibrium wages of graduates, and seek to maximize the total wage bill of their graduates. Universities are distinguished by the distribution of student abilities. We show that in equilibrium, universities with better distributions of students set lower grading standards — the marginal "A" student at better universities is less able than the marginal student at lesser schools. Indeed, better universities set grading standards that are below the social optimum, while worse universities set excessively strict grading standards. Improving the distribution of students at worse universities causes the better universities to raise their grading standards, but can have an ambiguous effect on grading standards at the worse university.

The trend of increasing GPA in the late 1900s is well-documented (see Rojstaczer (2003), and his website GradeInflation.com; and Mansfield (2001)); authors claim that GPA inflation leads to lower motivation levels and general depreciation of grades. However, the trend *per se* is not so interesting, it's the inference problems for third parties that are created by grade inflation that matter. Chen et al. (2007) model intentional loss of academic reliability, that is: the grading standard is not fixed so that otherwise identical people who take identical actions might not receive the same grade. They consider a framework where the abundance of good students is random, observed by schools but not firms; and they argue that this is why grading standards have varied over time. Dubey and Geanakoplos (2009) investigate how relative rank influence 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) look extensively on how the structure of the schooling market affects the tradeoff between studying effort, wealth and leisure, particularly investigating income effects (namely, whether income inequality will strengthen) and educational matching efficiency (namely, whether better students attend better schools).

A body of literature studies the behavior of grading standard from the point of view of central planner. Costrell (1994) studies how different policy towards standards can affect student effort (particularly, he makes a statement that egalitarian central planner is likely to pick lower grading standards than total earnings maximizer), and provides a review of the grading literature. Betts (1998) makes an opposite argument.

2 Modeling The Competition

The world contains two types of *universities*, $u \in \{H, I\}$ that supply *students* on the job market. The two types of 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 $f_I(\theta)$, where the densities are continuous and strictly positive on their support, $[0, 1]$. We capture the notion that the student body at a type H school is better than the distribution at a type I school with the concept of conditional first-order stochastic dominance. That is, $f_H(x|x > t)$ first order stochastically dominates $f_I(x|x > t)$ for all $t \in [0, 1)$, and $f_H(x|x < t)$ first-order stochastically dominates $f_I(x|x < t)$ for all $t \in (0,]$, written $f_H(x) \succeq_C f_I(x)$. In particular, the associated cumulated distribution functions satisfy $F_I(x|x > t) > F_H(x|x > t)$, for all $t \in (0, 1)$ and $x \in (t, 1)$. The total measure of students is normalized to one, and measure $\alpha \in (0, 1)$ of students attend type H universities. To capture that each university admits a negligible portion of the entire pool of students, we assume that there is a continuum of each type of university.

A student is distinguished by (i) his university type, (ii) his productive ability, θ, and (iii) his social skill, μ. Firms observe μ, but not θ. We assume that μ is distributed independently from θ according to the distribution $G(\cdot)$ with full concave unbounded from above support. We assume that the distribution of social skills is the same at all schools.

There are two types of firms, or jobs. There is a positive fraction Γ of jobs where the productivity of a student with ability (θ, μ) is $S(\theta + k\mu)$, and many jobs where the student's productivity is $s(\theta + k\mu)$, where $S > s > 0$; that is, there is a measure Γ of "good" jobs, and all other students receive "bad" jobs[1]. k is a constant capturing the relative importance of social skills. There are many firms, and we assume that wage offers are competitively determined with firms earning zero expected profits.

Universities know the abilities of each of their students, and their problem is to assign a grade $g \in \{A, B\}$ to each student[2]. We assume that universities assign grades to maximize the expected sum of wages accruing to their graduates; that turns out to be equivalent to maximizing the total product of students employed in good jobs. We also assume that universities' treasure academic integrity in a sense that only θ can affect the grade, and μ does not.

Firms do not observe the ability of students, but know the university that each student attended, the distribution of student abilities at each school, and the grading standard $\hat{\theta}$ that determines who

[1]This is a difference from Yang and Yip (2003), that essentially creates a reason to not give away too many good grades.

[2]Lizzeri (1999) argues why universities are not interested in revealing too much information; Dubey and Geanakoplos (2009) suggest a story why having coarse signal structure might help students' motivation. For instance, a lot of graduate programs formulate their admission requirements in form of thresholds, and these thresholds are more or less consistent among departments.

received "A" grades at each school. The zero profit condition implies that at a student receiving a good job will receive wage $S\left(E[\theta|g,u,\hat{\theta}]+k\mu\right)$, whereas a student receiving a bad job earns wage $s\left(E[\theta|g,u,\hat{\theta}]+k\mu\right)$.[3]

Denote a student with grade g from university u as ug student. Denote $E_{ug}\theta=\frac{\int_0^1 I(\text{grade is g})\theta f_u(\theta)d\theta}{\int_0^1 I(\text{grade is g})f_u(\theta)d\theta}$.[4] Notice that increasing θ_U would increase both $E_{uA}\theta$ and $E_{uB}\theta$. Notice also that if grading standard is the same, university H would have more "A"-graded students than university I will by CFOSD.

Good firms will employ a student demonstrating social skills of μ from university u with grade g if her total expected productivity $E_{ug}\theta+k\mu$ is big enough. Let K denote the lowest total expected productivity among employed students, and $\mu_{ug}=\frac{K-E[\theta|u,g,\hat{\theta}]}{k}$ will denote the minimum social skills required from the ug student. Each university is too small to affect the choice of K, but each university can see how choice of $\hat{\theta}$ will affect μ_{ug}. So K and demand curve summarize the response of good firms to the choice of grading strategies by universities.

Naturally, universities need to have some motivation to choose θ_H and θ_I. We posit that universities choose grading standards to maximize their students' total income (one can think of that as maximizing the average student's wage). University of type u choose $\hat{\theta}_u$ to maximize

$$\pi_u=S\int_{\hat{\mu}_{uA}}^{+\infty}\int_{\hat{\theta}_u}^{1}(k\mu+\theta)\,dF_u(\theta)dG(\mu)+s\int_{-\infty}^{\hat{\mu}_{uA}}\int_{\hat{\theta}_u}^{1}(k\mu+\theta)\,dF_u(\theta)dG(\mu)+$$

$$+S\int_{\hat{\mu}_{uB}}^{+\infty}\int_{0}^{\hat{\theta}_u}(k\mu+\theta)\,dF_u(\theta)dG(\mu)+s\int_{-\infty}^{\hat{\mu}_{uB}}\int_{0}^{\hat{\theta}_u}(k\mu+\theta)\,dF_u(\theta)dG(\mu)\rightarrow_{\hat{\theta}\in[0,1]}\max!$$

$$\text{s.t. }k\mu_{uA}+\frac{\int_{\hat{\theta}}^{1}\theta dF_u(\theta)}{\int_{\hat{\theta}}^{1}dF_u(\theta)}=k\mu_{uB}+\frac{\int_{0}^{\hat{\theta}}\theta dF_u(\theta)}{\int_{0}^{\hat{\theta}}dF_u(\theta)}=K$$

Universities know how labor market reacts to their choices, but the grading standard of a peers is taken as given. Then maximizing π_U is equivalent to maximizing the product of employed in good jobs[5], or

[3]Though a lot of literature, including Yang and Yip (2003) and Coate and Loury (1993), insist on same wage for same job, we do not: even though monetary remuneration might be required to be the same, nonmonetary benefits like on-site gym access or easy vacation can and will be used by firms to attract expected-better employees.

[4]If denominator happens to be 0, set $E_{uB}\theta=0$, $E_{uA}\theta=1$ to keep everything continuous.

[5]To see that, notice that maximizing total revenue is equivalent to maximizing total revenue minus a total revenue a university would get if everyone were employed in bad jobs (a constant), divided by $S-s$ (a positive constant).

$$\Pi_u = \int_{\hat{\mu}_{uA}}^{+\infty} \int_{\hat{\theta}_u}^{1} (k\mu + \theta)\, dF_u(\theta) dG(\mu) + \int_{\hat{\mu}_{uB}}^{+\infty} \int_{0}^{\hat{\theta}_u} (k\mu + \theta)\, dF_u(\theta) dG(\mu) \rightarrow_{\hat{\theta} \in [0,1]} \max!$$

Maximizing the total wage bill is equivalent to maximizing the total product of people employed in good jobs. Change of $\hat{\theta}$ that leads to either higher average productivity of employed students or with a higher quantity of employed students of university u, others being equal, will increase Π_u.

Proposition 1 *A positive mass of "A" students from university u is always employed by good firms.*

Proof. Obviously, expected productivity is higher among "A" students of university u than among "B" students of same university. Not employing "A" students in good jobs means not employing anyone, and getting value of $\Pi_U = 0$. Not employing anyone is worse than employing a positive mass of students, getting $\Pi_U > 0$. Therefore, university which does not place its students in good jobs will always change its grading standard. Can it change its' grading standard to something that will give positive employment?

Fix a university u where no "A" students are employed. It cannot be the case that there is no university where "A" students are not employed, since "B" students would not be employed as well. Take any other university where a positive mass of "A" students is employed. Set θ_u to be equal to expected θ for the second university. Then $E(\theta | f_u, \theta > \theta_u)$ is bigger than second university's, and therefore a positive mass of u students get employed. University u has a profitable deviation. ■

2.1 Equilibrium

Define symmetric equilibrium in pure strategies in this economy as a triple $(K^*, \theta_H^*, \theta_I^*)$, that satisfy following conditions:

- θ_H^* maximizes the utility of university of type H, $\Pi_H(\hat{\theta})$, subject to the lowest total expected productivity of K^* among employed students;

- θ_I^* maximizes the utility of university of type I, $\Pi_I(\hat{\theta})$, subject to the lowest total expected productivity of K^* among employed students;

- K^* is such that good job capacity constraint holds subject to θ_H^* and θ_I^*:

$$\alpha\left[\left(1-G(\mu_{HA})\right)\left(1-F_H(\theta_H)\right)+\left(1-G(\mu_{HB})\right)F_H(\theta_H)\right]+$$

$$+(1-\alpha)\left[\left(1-G(\mu_{IA})\right)\left(1-F_I(\theta_I)\right)+\left(1-G(\mu_{IB})\right)F_I(\theta_I)\right]=\Gamma.$$

$$k\mu_{ug}+E_{ug}\theta(\theta_u^*)=K \quad \forall ug\in\{HA,IA,HB,IB\}.$$

Equilibrium exists by Kakutani's fixed point theorem (universities' best response correspondences are upper hemicontinuous by Maximum Theorem; K^* is continuous in θ_H and θ_I by continuity of equations of the system).

2.2 Digression: $k=0$

In case of zero productivity of social skills only expected productivity of a *ug* group matter. This specific case is extremely useful to illustrate the mechanics of future results. First, assume there is a social planner that can set grading standards for schools, and assume she can perfectly enforce her standards; she cannot, however, unite the universities. She wants to maximize the total product of the economy.

Proposition 2 *Social planner will be choosing* $\theta_H^{\mathcal{P}}=\theta_I^{\mathcal{P}}$.

Proof. Take any social planner's strategy. Then, university by university, giving "A" to best students that total the same employment mass, gives her at least the same product. Thus, social planner's choice will be such that only "A" students are employed.

Social planner will be choosing "A" grades so that there's exactly Γ of "A" grades; otherwise, she can make standards stricter, thus increase the quality of employed students and increase the total product of the economy.

Evidently, same type universities will have same standards. Assume different university types have different standards. Then increasing standards in less-demanding university with decreasing standards in another type to keep total employment of Γ will necessarily increase the average product of students from that university, thus improving social planner's maximized function. ■

Next proposition will narrow down the space of variables we should look at when searching for an equilibrium.

Proposition 3 *In symmetric equilibrium with $k = 0$, if some students of a group ug are employed, group ug should be employed completely.*

Proof. If by increasing θ_u a university can increase the acceptance of its students (specifically, ug group), university will do it: both admittance and average productivity will increase. This rules out cases when group ug's productivity is equal to some other group's (slight increase in θ_u will render ug dominating the other group); therefore, ug can only be the lowest-productivity group among employed. If $g = A$, university can increase θ_u to increase $E_{uA}\theta$; even if the amount of employed students is not increasing, the average productivity of employed persons will increase, thus increasing Π_u.

If $g = B$, consider a decrease in θ_u so that the amount of people receiving grade "A" from u is equal to the amount of currently employed students from this university; denote $\widetilde{uA}$ the group of students with As after this change. Notice that $E_{\widetilde{uA}}\theta$ is going to be larger than average productivity of employed people from u: honest grading will make the best of uB to get "A"s under new standards. Also, it will be larger than E_{uB}, therefore other university's employment structure does not change; therefore, everyone from $\widetilde{uA}$ will be employed (they are all better than the group following ug, and their mass is equal to mass of previously employed people from u). Increased average productivity of employed people combined with same amount of employed people yields larger value of Π_u; therefore, previous θ_u was not an equilibrium one. ■

If an equilibrium contains not only HA and IA students, by Proposition 3 it has to be the case that all graduates of one university are employed.

Lemma 1 *Equilibrium with someone from IB employed and no one from HB not employed does not exist.*

Proof. Assume such equilibrium exists; then, for H universities to not deviate, $E_{HA}\theta \leq E_{IB}\theta$. However, by CFOSD and properties of conditional distributions, $E_{HA}\theta \geq E_H\theta > E_I(\theta) \geq E_{IB}\theta$, a contradiction. ■

Proposition 3 states that Γ has to be bigger than α for such an equilibrium to exist. Equilibrium where someone from every group is employed requires $\Gamma = 1$. Therefore, only "A" students are getting employed in equilibrium when Γ is no more than α.

Consider an equilibrium where only HA and IA students are employed. If one of the groups has higher average productivity than another, that first group will earn more by laxing its standards by a small amount: more people from this group will be hired; therefore, expected productivity of A students have to be equal in equilibrium. Combined with the capacity constraint, equilibrium conditions are:

$$\begin{cases} \frac{\int_{\theta_H^*}^1 \theta f_H(\theta)d\theta}{\int_{\theta_H^*}^1 f_H(\theta)d\theta} = \frac{\int_{\theta_I^*}^1 \theta f_I(\theta)d\theta}{\int_{\theta_I^*}^1 f_I(\theta)d\theta} \\ \alpha \int_{\theta_H^*}^1 f_H(\theta)d\theta + (1-\alpha)\int_{\theta_I^*}^1 f_I(\theta)d\theta = \Gamma \end{cases}$$

Proposition 4 *If $\Gamma \leq \alpha$, in equilibrium $\theta_H^* < \theta_I^*$.*

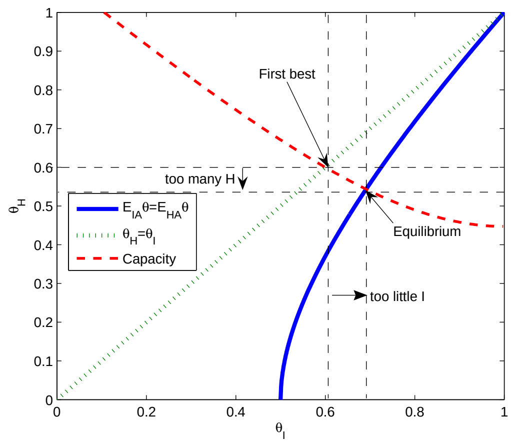
Proof. If $\Gamma \leq \alpha$ only equilibrium with employment of HA and IA students exists.

Equilibrium requires $E_{HA}\theta = E_{IA}\theta$. Define $G_U(x) = \frac{\int_x^1 t f_U(t)dt}{\int_x^1 f_U(t)dt}$ for $x \in (0,1)$, zero at 0 and one in 1; it is trivially strictly increasing in x. By CFOSD, $G_H(x) > G_I(x)$ for all x. Therefore, for every x the value y defined by $G_H(y) = G_I(x)$ is always less than x. ■

This result states that if good jobs are scarce enough, university H will have laxer grading standards, resulting in existence of positive mass of students of H with As that would not get an A in university I. This asymmetry, not a grading inflation, is what should be a topic of grading standard investigations: better universities have an incentive to dilute the mass of good students with As by people who would get a B in other universities: even after dilution, larger amount of good students make an average student of H as attractive as a student from I with an A.

Corollary 1 *If $\Gamma \leq \alpha$, there's too much of H students employed, and too little of I students.*

Reason for Corollary 1 is evident from Figure 1, depicting an equilibrium outcome. CFOSD makes solid line of equal expected productivities lie below 45° line; the intersection of dashed line and solid line is below 45° line is the statement of Proposition 4; capacity line is obviously always negatively sloped. Unlike Yang and Yip (2003), an ability to pick the grading standard not only makes some universities grade laxer than in first best scenario, but also forces other universities to have stricter grading standards. This outcome is still preferred by I university to no-grades outcome, which would render no I alumni in good jobs.



Note: $\Gamma = 0.25$, $\alpha = 0.5$. $f_H(x) = 2x$, $f_I(x) = 2 - 2x$.

Figure 1: Comparison of First Best and Equilibrium outcomes.

Improvement of I's distribution of students make standards in H stricter, and have an ambiguous effect on I's standard. Denote $E_U(x) = \int_x^1 \theta dF_U$, and $m_U(x) = \int_x^1 dF_U$. Parameterize I's distribution as $F_\lambda(x) = P(\theta < x|\lambda, I) = \lambda F_H(x) + (1-\lambda)F_I(x)$. Then $E(\theta|\lambda, I, \theta > x) = \lambda E_H(x) + (1-\lambda)E_I(x)$, and $\int_x^1 \theta dF_\lambda = \lambda m_H(x) + (1-\lambda)m_I(x)$. Let $(\theta_H^*(\lambda), \theta_I^*(\lambda))$ denote the equilibrium pair of grading standards for a value of $\lambda \in [0,1]$ when I's distribution of students' productivity is F_λ; let $(\theta_H^*, \theta_I^*) = (\theta_H^*(0), \theta_I^*(0))$. Then taking a derivative of the equilibrium conditions with respect to λ when $\lambda = 0$ yields

$$\begin{cases} E_H'(\theta_H^*)\theta_H^*(0)' = E_I'(\theta_I^*)\theta_I^*(0)' + (E_H(\theta_I^*) - E_I(\theta_I^*)) \\ -f_H(\theta_H)\theta_H^*(0)' - f_I(\theta_I)\theta_I^*(0)' + (m_H(\theta_I^*) - m_I(\theta_I^*)) = 0 \end{cases} \Rightarrow$$

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

$$\begin{pmatrix} \theta^*_H(0)' \\ \theta^*_I(0)' \end{pmatrix} = \begin{pmatrix} \frac{CE+BF}{AE+BD} > 0 \\ \frac{AF-CD}{AE+BD} \lesseqgtr 0 \end{pmatrix}$$

C and F, defined in above system, are positive by CFOSD. Similar result one can obtain in a similar experiment with H's students quality deterioration — it eases grading in I and leads to ambiguous effect in H.

If average student quality increases in a sense of α increasing, that would correspond to dashed line on Figure 1 becoming flatter (a little change in grading standard by H universities would allow for bigger change in I universities' grading standard), thus decreasing the grading discrepancy. Intersection of dashed line with $\theta_I = 1$ line will move up with increase in α, and intersection with $\theta_H = 1$ will move left; thus, effect on grading standards in both universities' types is ambiguous.

Naturally, if $\Gamma = 1$, there is no point in grading. What if there were an abundance of good jobs?

Proposition 5 *In equilibrium with not only "A" students getting employed $\theta^*_H > \theta^*_I$.*

Proof. By Lemma 1 and Proposition 1, if $\Gamma > \alpha$, employed groups are HA, HB and IA. Equilibrium implies $E_{HB}\theta = E_{IA}\theta$. Then

$$\theta^*_H > E_H(\theta|\theta < \theta^*_H) = E_I(\theta|\theta > \theta^*_I) > \theta^*_I$$

Thus, grading standard in H is greater than grading standard in I. ■

Solving for three-group equilibrium when $1 > \Gamma > \alpha$ is simple: first solve $\int_{\theta^*_I}^1 dF_I(\theta) = \frac{\Gamma-\alpha}{1-\alpha}$ for θ^*_I, and then pick θ^*_H by solving $E_H(\theta|\theta < \theta^*_H) = E_I(\theta|\theta > \theta^*_I)$. Solution always exists.

Summing up, when $k = 0$ better universities use their better student bodies to press some of worse universities' students off the labor market, thus having laxer grading standards than would be in first-best scenario, which is achievable by picking a grading standard that would make exactly Γ of "A" students. We now look at the case when μ matters.

2.3 General Case: $k > 0$

The case when both A and B students are employed is particularly interesting, since that is the scenario closest to the real world. It will happen since $G'(x) > 0$ on the whole support and $k > 0$: no matter how high are demands, there is a positive mass of students satisfying them.

Almost immediately conditions for both social planner and equilibrium tell us that $\mu_{uA} < \mu_{uB}$. Social planner's choice is a solution of a following system of equations (see Appendix A):

$$\theta_u = K - kE\left[\mu | \mu \in [\mu_{uA}, \mu_{uB}]\right]$$

$$k\mu_{uA} + E(\theta | \theta > \theta_u) = K$$

$$k\mu_{uB} + E(\theta | \theta < \theta_u) = K$$

Second and third conditions in both systems are representing a decreasing function in (θ_u, μ_{ug}) space. $E\left[\mu | \mu \in [\mu_{uA}, \mu_{uB}]\right]$ is an increasing function of both μ_{uA} and μ_{uB}; thus, implicit function $U(\mu_{uA}, \mu_{uB}) \triangleq Q$, where $U(\mu_{uA}, \mu_{uB}) \triangleq K - kE\left[\mu | \mu \in [\mu_{uA}, \mu_{uB}]\right]$, for some constant Q is a decreasing function in a space of (μ_{uA}, μ_{uB}). Also, it does not depend upon the distribution of θ; thus, it is the same for both universities in the space of (μ_{uA}, μ_{uB}). U decreases in both arguments.

Proposition 6 *When equilibrium is unique, social planner's choice is* $\theta_H^{\mathcal{P}} > \theta_I^{\mathcal{P}}$.

Proof. Substitute first social planner condition into the second and third ones:

$$\underbrace{kx + E(\theta | \theta > U(x, y)) = K}_{\text{Condition A}} \qquad \underbrace{ky + E(\theta | \theta < U(x, y)) = K}_{\text{Condition B}}$$

Consider Condition A evaluated for two different university types at the same point of x:

$$kx + E(\theta | H, \theta > U(x, y_H)) = K \qquad kx + E(\theta | I, \theta > U(x, y_I)) = K$$

By conditional first-order stochastic dominance, $U(x, y_H) < U(x, y_I)$, and therefore $y_H > y_I$. That makes Condition A steeper in (μ_{uA}, μ_{uB}) space than Condition A for I (see Figure 2 for illustration).

Condition B analogously implies an increasing curve in (θ, μ) space as long as $E[\theta|\theta < t]$ is increasing slower than t (which is not guaranteed by first-order stochastic dominance); if $E[\theta|\theta < t]$ is increasing faster for H than for I, slope of Condition B would be flatter.

Condition A equations for two types of universities intersect when $E_H(\theta|\theta > U(x,y)) = E_I(\theta|\theta > U(x,y))$; it is trivial to establish that this can only happen when $U(\mu_{uA}, \mu_{uB}) = 1$, and it is trivially satisfied at $\mu_{uB} = \mu_{uA} = \frac{K-1}{k}$. Condition B for two types of universities intersect when $U(\mu_{uA}, \mu_{uB}) = 0$ and it is trivially satisfied at $\mu_{uB} = \mu_{uA} = \frac{K}{k}$. By their intersection satisfying $\mu_{uA} < \mu_{uB}$, intersection of Condition A equations is in south-west part of (μ_{uA}, μ_{uB}) space compared to intersection of Condition B equations.

$$(\mu_{HA}, \mu_{HB}) < (\mu_{IA}, \mu_{IB}) \Rightarrow U(\mu_{HA}, \mu_{HB}) < U(\mu_{IA}, \mu_{IB}) \Rightarrow \theta_H > \theta_I$$

Thus, social planner will be more demanding to H students than to I students. ■

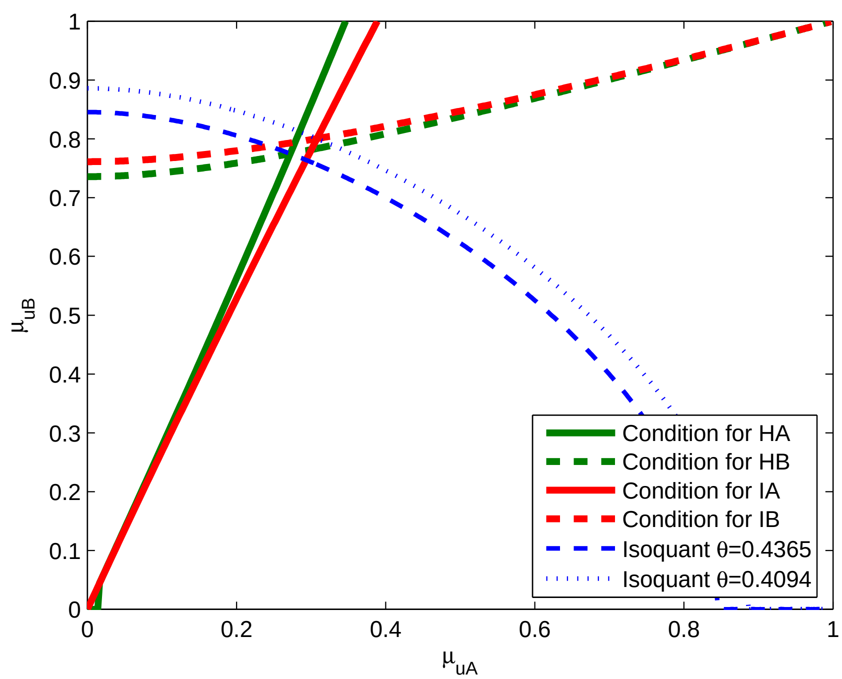


Figure 2: Social planner's choice when $k > 0$.
Note: $G(x) = x^2$, $k = 1$, $f_H(x) = 2(0.4 + 0.2x)$, $f_I(x) = 2(0.6 - 0.2x)$. The example distribution does not have a bounded support.

Corollary 2 *Under conditions of Proposition 6, improvement of the student distribution in one I-*

type university distribution will increase it's grading standard and lower its both μ_A and μ_B cutoffs. Reverse is true for single H-type university getting worse distribution.

Think of a distribution of students' abilities of a given university as a mixture of $F_I(\cdot)$ with weight γ and $F_H(\cdot)$ with a complimentary weight. Then for $\gamma \in (0,1)$ it is true that

$$E(\theta|\gamma, \theta \in \Theta) = \gamma E(\theta|I, \theta \in \Theta) + (1-\gamma)E(\theta|H, \theta \in \Theta),$$

which will make the position of Conditions A and B to be between the Conditions for H-typed universities and I-typed universities, from which the result is immediate.

The equilibrium conditions, obtained in Appendix A, are

$$\theta_u = V(\mu^*_{uA}, \mu^*_{uB}) \triangleq \frac{\frac{K}{k}\frac{g(\mu^*_{uA})\left(K-k\mu^*_{uA}\right)-g(\mu_{uB*})\left(K-k\mu^*_{uB}\right)}{G(\mu^*_B)-G(\mu^*_A)} - kE(\mu|\mu \in [\mu^*_{uA}, \mu^*_{uB}])}{\left(1+\frac{K}{k}\frac{g(\mu^*_{uA})-g(\mu^*_{uB})}{G(\mu^*_{uB})-G(\mu^*_{uA})}\right)}$$

$$k\mu_{uA} + E(\theta|\theta > \theta_u) = K^*$$

$$k\mu_{uB} + E(\theta|\theta < \theta_u) = K^*$$

Second and third conditions in both systems are similar to the ones in social planner's choice conditions, a decreasing function in (θ_u, μ_{ug}) space.

Proposition 7 *When equilibrium is unique, and Γ and k are small enough, in equilibrium $\theta^*_H < \theta^*_I$.*

Proof.

When k is 0, Proposition 4 guarantees $\theta^*_H < \theta^*_I$. Problems of both universities are continuous in k, best responses have a unique intersection; therefore, solution is continuous in k. Therefore, for small enough $k > 0$ the relative position of θ^*_H and θ^*_I should preserve the relation of such under $k = 0$. Therefore, $\theta^*_H < \theta^*_I$. ■

Summing up this section, we conclude that there is a significant incentive for better schools to lower their grading standards, tarring their best students with mediocre ones, exploiting the labor market's job assignment inefficiency to place socially inefficiently many graduates. This will happen when grades matter, and when good jobs are scarce.

Notice also that all propositions in no way utilize the fact that there is only two types of universities. As long as university types can be CFOSD-ordered by their distribution of θ, all results apply, whether types are discrete or continuous.

3 Endogenizing The Effort

Dubey and Geanakoplos (2009), Costrell (1994) and Betts (1998) argue that grading standards affect effort levels. Naturally, others being equal, it is easier to satisfy lower grading standards. How university competition is going to affect the students' choice of effort?

To study the effect of grading standards, we drop the assumption of bounded support for θ (which was useful for illustrative purposes), and assume $f_I(\theta) = f_H(\theta - C) = f(\theta)$ for some positive C, $f(\cdot)$ being positive everywhere; easy to see that $F_H \succ_C F_I$. Since support is unbounded, for every k there will be a positive mass of both A and B students that are going to get a good job.

However, now the good job technology is $S(\theta + k\mu + e)$, with e denoting effort of the student, unobservable by firms, and undistinguishable from θ by universities. So, universities grade based on comparison of a total of $\theta + e$ and θ_u: students with $\theta + e > \theta_u$ get "A".

The timing of the game is organized similarly to the game in previous section. Students join universities, and the admission process is organized so that the abilities of students joining H is distributed as $f_H(\cdot)$; abilities of students who join I is distributed with cdf of $f_I(\cdot)$. Students are not aware of their θ or μ yet; they choose the positive level of effort to maximize the expected wage provided they know what is their university's θ distribution. This makes them all invest equal amount of effort e_u, in the spirit of Bernhardt and Mongrain (2009), thus effectively changing the distribution of $\theta + e$ from $f_H(x) = f(x)$ to $f_H(x|e_H) = f(x + e_H)$ and from $f_I(x) = f(x + c)$ to $f_I(x|e_I) = f(x - e_I + C)$. Students face costs of efforts of $(S - s)c(e)$ (here $(S - s)$ is included to later be cancelled out in students' problem).

Thus, in **effort** equilibrium is

- $\theta_u^*, u \in \{H, I\}$ — the choice of grading standard by university as a best response to K^* and e_u^*.

- K^* and μ_{ug} — the outcome of Bertrand competition by firms based on grade and university type; a function of θ_u^* and e_u^*.

- $e_u^*, u \in \{H, I\}$ — the choice of effort level chosen by students as a best response to θ_u^*, μ_{ug}, K^* and e_u^*.

such that

- $k\mu_{uA} + \underbrace{E[\theta + e_u | u, \theta + e_u > \theta_u]}_{w_{uA}} = K;$
- $k\mu_{uB} + \underbrace{E[\theta + e_u | u, \theta + e_u < \theta_u]}_{w_{uB}} = K;$
- Students solve individual effort-choosing problem $e_u = \text{argmax}_e Q_u(e) - c(e)$, where $Q_u(e)$ denotes the revenue premium from effort level e:

$$Q_u(e) = P(\theta|\theta + e > \theta_u) \left(\int_{\mu_{uA}}^{\infty} [k\mu + e + w_{uA}]\, dG - \int_{-\infty}^{\mu_{uA}} [k\mu + e + w_{uB}]\, dG \right) +$$
$$+ P(\theta|\theta + e < \theta_u) \left(\int_{\mu_{uB}}^{\infty} [k\mu + e + w_{uA}]\, dG - \int_{-\infty}^{\mu_{uB}} [\mu + e + w_{uB}]\, dG \right)$$

- Universities solve universities' problems taking e_u as given.

- Values of $\mu_X \forall X \in \{HA, HB, IA, IB\}$ are such that capacity of good jobs is Γ.

Lemma 2 *Increase in students' effort in a university u leads to lower μ_{ug}s and to lower θ_u.*

Proof. Notice that increase in effort in a university is equivalent to have this university's $\theta + e$ distribution stochastically dominate the distribution before the effort increase. Thus, under same conditions as in Proposition 7, higher e means lower θ and lower μ, others being equal. ■

Next assumption makes students exercise positive effort.

Assumption 1 *$c(e)$ satisfies $c(0) = 0$, $c'(0) = 0$, $c'(x) > 0$, $c''(x) > 0$.*

Lemma 3 *Assume $k = 0$ and $s = 0$. Assume $\theta_H < \theta_I$. In equilibrium, if costs are small enough, $e_H < e_I < e_H + C + (\theta_H^* - \theta_I^*)$; if costs are high enough, and Γ is small enough, $e_H > e_I$.*

Proof. From the point of view of a student, she is expecting a wage premium of $\bar{W}_u$ if her effort is enough to overcome $\hat{\theta}_u$ barrier. Therefore, her problem is

$$\max_e P_u(\theta + e > \hat{\theta}_u)\bar{W} - c(e)$$

yielding the first-order condition of $M_u(e_u^*) = c'(e_u^*)$ where $M_u(e) = W_u f_u(\theta_u - e)$. When $s = 0$ wage premium is actually just the wage on the good job. By equilibrium conditions, when Γ is small enough, wages on the good jobs of students of both universities should be equal. Therefore, in equilibrium $\bar{W}_H = \bar{W}_I = \bar{W}$.

Notice that $M_u(e) \to +\infty$ as $e \to 0^+$ and $M_u(e)$ is decreasing near $e = 0$ by Assumption 1. Also notice that $M_u \to 0$ as $e \to +\infty$.

Since $M_H(e) = M_I(e - C + (\theta_H - \theta_I))$, if $M_I'(0) < 0$, it has to be the case that $M_H(e) < M_I(e)$ by single-peakedness. When $M_H(e) < M_I(e)$, $c'(e)$, an increasing function, will intersect $M_H(e)$ in a point of e_H and $M_I(e)$ in a point of e_I so that $e_H < e_I$. Notice also that $M_I^{(-1)}(M_H(e)) = e + C + (\theta_H^* - \theta_I^*)$, and, since $c'(e)$ is strictly increasing, $e_I - e_H < C + (\theta_H^* - \theta_I^*)$.

The only way to get $M_H(e) > M_I(e)$ is to have $f_H(\hat{\theta}_H) > f_I(\hat{\theta}_I)$, which definitely will happen when both $\hat{\theta}_u$ are big enough; in equilibrium this is only possible when Γ is small enough. By M_H being a horizontal shift of M_I and single-peakedness, there is at most one intersection of these functions; call $\tilde{e}$ the argument where they intersect. If $c'(e)$ is small enough, it will intersect $M_H(e)$ after $\tilde{e}$, and the logic of the case when $M_H(e) < M_I(e)$ applies. If $c'(e)$ is intersecting $M_I(e)$ at or on left of $\tilde{e}$, that makes $e_H \geq e_I$. ■

Proposition 8 *Under assumptions in Lemma 3, in equilibrium $\theta_H < \theta_I$.*

Proof. Assume that in equilibrium $\theta_H = \theta_I$. By Lemma 3, $e_I - e_H < C + (\theta_H - \theta_I)$. Thus, distribution of $\theta + e_H$ across "H" students population still conditionally first-order stochastically dominates distribution of $\theta + e_I$ across "I" students population, implying that best-response to $\theta_H = \theta_I$ has to be $\theta_H < \theta_I$ by Proposition 4, and thus $\theta_H = \theta_I$ cannot be an equilibrium. Same logic rejects $\theta_H > \theta_I$ in equilibrium. Thus, $\theta^*_H < \theta^*_I$. ■

Next two propositions establish that social planner and equilibrium are solving two different problems in the case of $k > 0$, and provide crude conditions for when H university in equilibrium demands lower signals than

Proposition 9 (To prove) *Social planner picks* $e^{\mathcal{P}}_H > e^{\mathcal{P}}_I$, $\mu_{uA} < \mu_{uB}$ *and* $\theta_H > \theta_I$.

Proposition 10 (To prove) *In equilibrium, when* Γ *and* k *are small, universities pick* $e_H < e_I$, $\mu_{uA} < \mu_{uB}$ *and* $\theta_H < \theta_I$.

4 Discussion and Conclusion

The main message we attempt to convey is that comparative advantage of one university in terms of student body composition might lead to inefficient in second-best sense equilibrium because better (in a sense of student composition) university is motivated to ease the grading standards, which leads to easier access of lower-able students to the good job. This effect is not necessarily relieved with introduction of observable effects, and it is exacerbated when students attempt to exhibit efforts. Higher amount of good grades in good universities do not mean lower grading standards, but better student bodies combined with tight entrant market imply lower grading standards in better universities, when universities strive to maximize its graduates' total wages. Notice that we do not make specific statements about how some universities get better student bodies: the reason of better student composition, whether it is because students prefer university H over university I due to better "prestige" or because university H teaches better, or some other reasoning for why it happen, does not affect results. We do, however, believe that some universities have better student bodies than other universities, and ongoing empirical research on graduates' wages (see PayScale, Inc. (2009)) confirms our beliefs.

Though we do not model the grading inflation explicitly, we believe that the best explanation for grading inflation in United States in late 1900s is the gradual increase of the size of labor market of good jobs. Our model predicts that for big enough good job market good universities might suddenly change their behavior to much stricter grading standards (see Proposition 5), a process in line with a campaign for stricter grading in early 2000s.

We believe, we are second after Yang and Yip (2003) to explicitly model for university competition on placement market, and we are first to obtain predictions consistent with common sense (that is, in our world an average person with an "A" grade is better that average person with a "B" grade, and giving a job to "A" person gives strictly positive benefit compared to giving a job to a person with a "B" grade). This is achieved a lot with assuming a limited capacity of good jobs, which is not an innocuous assumption. However, it is not an extremely daring assumption as well: the amount of new job openings is usually determined before the new alumni job market when setting up the budgets of departments, and this funding allocation is hard to change.

The asymmetry that arise from significant difference between distributions of students' abilities is detrimental to the society in different ways. First and foremost, equally productive students do not get same wages; this is partially a consequence of information asymmetry. Equally productive students do not get the same job — this result is a direct consequence of unfair competition. Introduction of observable characteristics or endogenous effort of students does not necessarily save the day. In the worst case equilibrium scenario, students of better university put in less effort and get a good grade for lower performance than their peers from university with worse distribution of abilities. Employers, on the other hand, even knowing that students of better university are slacking off, are less demanding to students from better universities. Quantitatively, the size of the effect is not obvious; we expect to accumulate sufficient data to deliver the numerical evaluations.

References

Bernhardt, D. and S. Mongrain (2009): "The Layoff Rat Race," *Scandinavian Journal of Economics*.

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.

Healy, P. (2002): "Harvard Looks To Raise Bar For Graduating With Honors," *The Boston Globe*.

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.

Mansfield, H. C. (2001): "Point of View: Grade Inflation: It's Time to Face the Facts," *The Chronicle of Higher Education*.

PayScale, Inc. (2009): "Top State Universities by Salary Potential," http://www.payscale.com/best-colleges/top-state-universities.asp.

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

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

A Solutions of Equilibrium and Social Planner's Choice Problems

Social Planner solves the following problem:

$$\alpha\left(\int_{\hat{\mu}_{HA}}^{+\infty}\int_{\hat{\theta}_H}^{1}(k\mu+\theta)\,dF_H(\theta)dG(\mu)+\int_{\hat{\mu}_{HB}}^{+\infty}\int_{0}^{\hat{\theta}_H}(k\mu+\theta)\,dF_H(\theta)dG(\mu)\right)+$$

$$+(1-\alpha)\left(\int_{\hat{\mu}_{IA}}^{+\infty}\int_{\hat{\theta}_I}^{1}(k\mu+\theta)\,dF_I(\theta)dG(\mu)+\int_{\hat{\mu}_{IB}}^{+\infty}\int_{0}^{\hat{\theta}_I}(k\mu+\theta)\,dF_I(\theta)dG(\mu)\right)\xrightarrow[\hat{\theta}_H,\hat{\theta}_I,\hat{\mu}_X]{}\max!$$

$$\text{s.t. }\alpha\left(\int_{\hat{\mu}_{HA}}^{+\infty}\int_{\hat{\theta}_H}^{1}dF_H(\theta)dG(\mu)+\int_{\hat{\mu}_{HB}}^{+\infty}\int_{0}^{\hat{\theta}_H}dF_H(\theta)dG(\mu)\right)+$$

$$+(1-\alpha)\left(\int_{\hat{\mu}_{IA}}^{+\infty}\int_{\hat{\theta}_I}^{1}dF_I(\theta)dG(\mu)+\int_{\hat{\mu}_{IB}}^{+\infty}\int_{0}^{\hat{\theta}_I}dF_I(\theta)dG(\mu)\right)=\Gamma.$$

μ_X denotes the quadruple of μ_{HA}, μ_{HB}, μ_{IA}, μ_{IB}. Notice that we do not explicitly put the labor market response constraint — two of social planner's FOCs will exactly replicate it. The problem's Lagrangean is

$$\mathfrak{L}(\hat{\theta}_H,\hat{\theta}_I,\hat{\mu}_X,\lambda)=\alpha\left(\int_{\hat{\mu}_{HA}}^{+\infty}\int_{\hat{\theta}_H}^{1}(k\mu+\theta)\,dF_H(\theta)dG(\mu)+\int_{\hat{\mu}_{HB}}^{+\infty}\int_{0}^{\hat{\theta}_H}(k\mu+\theta)\,dF_H(\theta)dG(\mu)\right)+$$

$$+(1-\alpha)\left(\int_{\hat{\mu}_{IA}}^{+\infty}\int_{\hat{\theta}_I}^{1}(k\mu+\theta)\,dF_I(\theta)dG(\mu)+\int_{\hat{\mu}_{IB}}^{+\infty}\int_{0}^{\hat{\theta}_I}(k\mu+\theta)\,dF_I(\theta)dG(\mu)\right)$$

$$-\lambda\left(\alpha\left(\int_{\hat{\mu}_{HA}}^{+\infty}\int_{\hat{\theta}_H}^{1}dF_H(\theta)dG(\mu)+\int_{\hat{\mu}_{HB}}^{+\infty}\int_{0}^{\hat{\theta}_H}dF_H(\theta)dG(\mu)\right)+\right.$$

$$\left.+(1-\alpha)\left(\int_{\hat{\mu}_{IA}}^{+\infty}\int_{\hat{\theta}_I}^{1}dF_I(\theta)dG(\mu)+\int_{\hat{\mu}_{IB}}^{+\infty}\int_{0}^{\hat{\theta}_I}dF_I(\theta)dG(\mu)\right)-\Gamma\right)\xrightarrow[\hat{\theta}_H,\hat{\theta}_I,\hat{\mu}_X,\lambda]{}\max!$$

First-order conditions, besides the capacity constraint, are:

$$\begin{aligned}
\alpha g(\mu_{HA}) \int_{\theta_H}^{1} (k\mu_{HA} + \theta)\, dF_H &= \alpha g(\mu_{HA})\lambda \int_{\theta_H}^{1} dF_H, && (1)\\
\alpha g(\mu_{HB}) \int_{0}^{\theta_H} (k\mu_{HB} + \theta)\, dF_H &= \alpha g(\mu_{HB})\lambda \int_{0}^{\theta_H} dF_H, && (2)\\
(1-\alpha) g(\mu_{IA}) \int_{\theta_I}^{1} (k\mu_{IA} + \theta)\, dF_I &= (1-\alpha) g(\mu_{IA})\lambda \int_{\theta_I}^{1} dF_I, && (3)\\
(1-\alpha) g(\mu_{IB}) \int_{0}^{\theta_I} (k\mu_{IB} + \theta)\, dF_I &= (1-\alpha) g(\mu_{IB})\lambda \int_{0}^{\theta_I} dF_I, && (4)\\
\alpha \left(\int_{\mu_{HA}}^{\mu_{HB}} (k\mu + \theta_H)\, dG \right) f_H \theta_H &= \alpha\lambda \left(\int_{\mu_{HA}}^{\mu_{HB}} dG \right) f_H(\theta_H), && (5)\\
(1-\alpha) \left(\int_{\mu_{IA}}^{\mu_{IB}} (k\mu + \theta_I)\, dG \right) f_I \theta_I &= (1-\alpha)\lambda \left(\int_{\mu_{IA}}^{\mu_{IB}} dG \right) f_I(\theta_I). && (6)
\end{aligned}$$

λ will be pinned down by capacity constraint. Divide all equations by α term and by density terms:

$$\begin{aligned}
\int_{\theta_H}^{1} (k\mu_{HA} + \theta)\, dF_H &= \lambda \int_{\theta_H}^{1} dF_H, && (7)\\
\int_{0}^{\theta_H} (k\mu_{HB} + \theta)\, dF_H &= \lambda \int_{0}^{\theta_H} dF_H, && (8)\\
\int_{\theta_I}^{1} (k\mu_{IA} + \theta)\, dF_I &= \lambda \int_{\theta_I}^{1} dF_I, && (9)\\
\int_{0}^{\theta_I} (k\mu_{IB} + \theta)\, dF_I &= \lambda \int_{0}^{\theta_I} dF_I, && (10)\\
\left(\int_{\mu_{HA}}^{\mu_{HB}} (k\mu + \theta_H)\, dG \right) &= \lambda \int_{\mu_{HA}}^{\mu_{HB}} dG, && (11)\\
\left(\int_{\mu_{IA}}^{\mu_{IB}} (k\mu + \theta_I)\, dG \right) &= \lambda \int_{\mu_{IA}}^{\mu_{IB}} dG. && (12)
\end{aligned}$$

Take the integrals:

$$
\begin{aligned}
(1-F_H(\theta_H))k\mu_{HA} + \int_{\theta_H}^{1} \theta dF_H &= \lambda(1-F_H(\theta_H)), && (13)\\
F_H(\theta_H)k\mu_{HB} + \int_{0}^{\theta_H} \theta dF_H &= \lambda F_H(\theta_H), && (14)\\
(1-F_I(\theta_I))k\mu_{IA} + \int_{\theta_I}^{1} \theta dF_I &= \lambda(1-F_I(\theta_I)), && (15)\\
F_I(\theta_I)k\mu_{IB} + \int_{0}^{\theta_I} \theta dF_I &= \lambda F_I(\theta_I), && (16)\\
k\int_{\mu_{HA}}^{\mu_{HB}} \mu dG + \theta_H\left(G(\mu_{HB}) - G(\mu_{HA})\right) &= \lambda\left(G(\mu_{HB}) - G(\mu_{HA})\right), && (17)\\
k\int_{\mu_{IA}}^{\mu_{IB}} \mu dG + \theta_I\left(G(\mu_{IB}) - G(\mu_{IA})\right) &= \lambda\left(G(\mu_{IB}) - G(\mu_{IA})\right). && (18)
\end{aligned}
$$

Divide first four equations by the F-related term, and last two equations by $G(\cdot) - G(\cdot)$ term:

$$
\begin{aligned}
k\mu_{HA} + E_H(\theta|\theta > \theta_H) &= \lambda, && (19)\\
k\mu_{HB} + E_H(\theta|\theta < \theta_H) &= \lambda, && (20)\\
k\mu_{IA} + E_I(\theta|\theta > \theta_I) &= \lambda, && (21)\\
k\mu_{IB} + E_I(\theta|\theta < \theta_I) &= \lambda, && (22)\\
kE(\mu|\mu \in [\mu_{HA}, \mu_{HB}]) + \theta_H &= \lambda, && (23)\\
kE(\mu|\mu \in [\mu_{IA}, \mu_{IB}]) + \theta_I &= \lambda,. && (24)
\end{aligned}
$$

Notice that equations (19)-(22) all say that the person from that group with smallest μ has to have the productivity of λ. If one had labor market response constraints in original problem, these constraints would have zero lagrange multipliers. Also, Equations (19)-(22) say that $\mu_{HA} < \mu_{HB}$ and $\mu_{IA} < \mu_{IB}$.

Another thing worth noticing is that equations (19), (20) and (23) are the same as (21), (22) and (24), with H replaced by I. The system can be rewritten in a more compact way by saying that for every $u \in \{H, I\}$ following three equations must hold in social planner's choice:

$$
\begin{aligned}
k\mu_{uA}^{\mathcal{P}} + E_u(\theta|\theta > \theta_u^{\mathcal{P}}) &= K, && (25)\\
k\mu_{uB}^{\mathcal{P}} + E_u(\theta|\theta < \theta_u^{\mathcal{P}}) &= K, && (26)\\
kE(\mu|\mu \in [\mu_{uA}^{\mathcal{P}}, \mu_{uB}^{\mathcal{P}}]) + \theta_u^{\mathcal{P}} &= K, && (27)
\end{aligned}
$$

where K is a value determined by distributions, α, k and Γ, and does not depend on university identity, and $X^{\mathcal{P}}$ denotes the value of variable X chosen by social planner.

In **equilibrium**, universities observe the labor market's minimum required productivity for employment of K, and solve the following problem subject to it:

$$\int_{\hat{\mu}_A}^{+\infty}\int_{\hat{\theta}}^{1}(k\mu+\theta)\,dF(\theta)dG(\mu)+\int_{\hat{\mu}_B}^{+\infty}\int_{0}^{\hat{\theta}}(k\mu+\theta)\,dF(\theta)dG(\mu)\xrightarrow[\hat{\theta},\hat{\mu}_A,\hat{\mu}_B]{}\max!$$
$$\text{s.t. } k\mu_A+E(\theta|\theta>\hat{\theta})=K\quad k\mu_B+E(\theta|\theta<\hat{\theta})=K$$

Problem's lagrangean is

$$\mathfrak{L}(\hat{\theta},\hat{\mu}_A,\hat{\mu}_B,\lambda_1,\lambda_2)=\int_{\hat{\mu}_A}^{+\infty}\int_{\hat{\theta}}^{1}(k\mu+\theta)\,dF(\theta)dG(\mu)+\int_{\hat{\mu}_B}^{+\infty}\int_{0}^{\hat{\theta}}(k\mu+\theta)\,dF(\theta)dG(\mu)+$$
$$+\lambda_1\Big(k\mu_A+E(\theta|\theta>\hat{\theta})-K\Big)+\lambda_2\Big(k\mu_B+E(\theta|\theta<\hat{\theta})-K\Big)$$

Here we omitted the university-specific subscript for choice variables; variables with stars will from now on denote the equilibrium choice of the university. First-order conditions are:

$$g(\mu_A^*)\int_{\theta^*}^{1}(k\mu_A^*+\theta)\,dF = k\lambda_1 \tag{28}$$
$$g(\mu_B^*)\int_{0}^{\theta^*}(k\mu_B^*+\theta)\,dF = k\lambda_2 \tag{29}$$
$$f(\theta^*)\left(\int_{\mu_A^*}^{\mu_B^*}(k\mu+\theta^*)\,dG\right) = \lambda_1\frac{f(\theta^*)\left(E(\theta|\theta>\theta^*)-\theta^*\right)}{1-F(\theta^*)}+\lambda_2\frac{f(\theta^*)\left(\theta^*-E(\theta|\theta<\theta^*)\right)}{F(\theta^*)} \tag{30}$$
$$k\mu_A^*+E(\theta|\theta>\theta^*) = K, \tag{31}$$
$$k\mu_B^*+E(\theta|\theta<\theta^*) = K. \tag{32}$$

Solve the integral in (28) and (29), substitute (31) and (32) to get expressions for λ_1 and λ_2 in terms of K. Solve the integral in (30), substitute λs, get

$$g(\mu_A^*)\underbrace{(k\mu_A^*+E(\theta|\theta>\theta^*))}_{K} = k\frac{\lambda_1}{1-F(\theta^*)} \tag{33}$$
$$g(\mu_B^*)\underbrace{(k\mu_B^*+E(\theta|\theta<\theta^*))}_{K} = k\frac{\lambda_2}{F(\theta^*)} \tag{34}$$
$$kE(\mu|\mu\in[\mu_A,\mu_B])+\theta^* = \frac{K}{k}\frac{g(\mu_A)\left(E(\theta|\theta>\theta^*)-\theta^*\right)+g(\mu_B)\left(\theta^*-E(\theta|\theta<\theta^*)\right)}{G(\mu_B^*)-G(\mu_A^*)} \tag{35}$$
$$k\mu_A^*+E(\theta|\theta>\theta^*) = K, \tag{36}$$
$$k\mu_B^*+E(\theta|\theta<\theta^*) = K. \tag{37}$$

Thus, in equilibrium each university is going to choose grading policies according to

$$\theta^* = \frac{\frac{K}{k}\frac{g(\mu_A^*)\left(K-k\mu_A^*\right)-g(\mu_B^*)\left(K-k\mu_B^*\right)}{G(\mu_B^*)-G(\mu_A^*)} - kE(\mu|\mu \in [\mu_A^*, \mu_B^*])}{\left(1+\frac{K}{k}\frac{g(\mu_A^*)-g(\mu_B^*)}{G(\mu_B^*)-G(\mu_A^*)}\right)} \tag{38}$$

$$k\mu_A^* + E(\theta|\theta > \theta^*) = K, \tag{39}$$

$$k\mu_B^* + E(\theta|\theta < \theta^*) = K. \tag{40}$$