NBER WORKING PAPER SERIES

GRADE NON-DISCLOSURE

Daniel Gottlieb
Kent Smetters

Working Paper 17465
http://www.nber.org/papers/w17465

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2011

Grade Non-Disclosure
Daniel Gottlieb and Kent Smetters
NBER Working Paper No. 17465
September 2011
JEL No. D0,D82,I21,J24

ABSTRACT

This paper documents and explains the existence of grade non-disclosure policies in Masters in Business Administration programs, why these policies are concentrated in highly-ranked programs, and why these policies are not prevalent in most other professional degree programs. Related policies, including honors and minimum grade requirements, are also consistent with our model.

Daniel Gottlieb
University of Pennsylvania
SH-DH 3303
3620 Locust Walk
Philadelphia, PA 19104
dgott@wharton.upenn.edu

Kent Smetters
University of Pennsylvania
SH-DH 3303
3620 Locust Walk
Philadelphia, PA 19104
and NBER
smetters@wharton.upenn.edu

Grade Non-Disclosure

Daniel Gottlieb and Kent Smetters*

September 22, 2011

Abstract

This paper documents and explains the existence of grade non-disclosure policies in Masters in Business Administration programs, why these policies are concentrated in highly-ranked programs, and why these policies are not prevalent in most other professional degree programs. Related policies, including honors and minimum grade requirements, are also consistent with our model.

Journal of Economic Literature Classification Numbers: D80, D71, D81

Key Words: Signaling, education, disclosure

1 Introduction

The signal value of education has been well studied ever since Michael Spence's seminal 1973 contribution. It is, therefore, interesting that students in many leading Master of Business Administration (MBA) programs vote to reduce the accuracy of their own signal by passing grade non-disclosure policies. These policies are distinctive in that they mainly exist in MBA programs and not in other professional programs including medicine, law, and accounting.[1] Moreover, grade non-disclosure is most prevalent in schools where the faculty are largely trained economists. In some cases, faculty seems to support these policies – or, at least, not heavily discourage them.[2]

*The Wharton School, 3620 Locust Walk, SH-DH 3303, Philadelphia, PA. Daniel Gottlieb: dgott@wharton.upenn.edu. Kent Smetters: Smetters@wharton.upenn.edu.

[1] Some medical and under-graduate programs (e.g., MIT) have limited grade non-disclosure, for example, covering first-year grades. Yale Law has only limited disclosure as well. But these cases are exceptional and typically don't limit the signal as much as many MBA programs.

[2] As deputy dean of Stanford's Graduate School of Business, David Kreps wrote a thoughtful memorandum in 2005 that discussed the advantages and disadvantages of grade non-disclosure during a period of school debate. He expressed his personal support for non-disclosure, with some caveats. But he explained that the school's official policy under the law is to remain neutral, thereby making non-disclosure a norm of students rather than a policy of the school. Michael Spence was Stanford's dean during this period. At Harvard Business School, grade non-disclosure was considered an official school policy. Students were required by the administration to accept grade non-disclosure before matriculation, and those who did not follow these

By U.S. federal law, academic grades cannot be released by schools to third-parties, including potential employers, without student permission.[3] The law allows for "directory" information to be disclosed, including characteristics such as name, address, field of study, date of attendance, and degrees as well as "awards" and other "honors." The exact level of granularity for awards and honors is not defined precisely by law. But the conventional legal wisdom is that information pertaining to awards and honors should be limited to exceptional performance. An award or honor system whose main purpose is to substitute for traditional grading marks (e.g., "Platinum," "Gold," "Silver," and "Bronze" instead of "A," "B," "C," and "D/F"), would not be considered to be "directory" information.

Since grades are the property of students and not schools, students can vote to create a "social norm" of grade non-disclosure to potential employers. While individual students are legally allowed to break ranks with this norm and disclose their individual grades, they generally do not. Moreover, employers (often including alumni) typically do not ask for grade information at schools where non-disclosure has been endorsed by students.

As shown in Tables I and II, grade non-disclosure policies in MBA programs are concentrated within highly-ranked schools. A majority of the most selective 15 schools have a grade non-disclosure policy, while no school ranked 20 - 50 has such a policy. In some cases, school faculty do not appear to openly oppose the non-disclosure policy approved by students. In other cases, like at Wharton, the faculty is fairly vocal in their opposition to the policy, which is generally approved each year by a large majority (about 95%) of students. Students at Wharton and many other top-tier programs argue that non-disclosure allows them to take greater risks and harder courses without fear of having an embarrassing transcript. In practice, self-reported levels of learning effort have fallen significantly since the introduction of grade non-disclosure (Jain 2005).

However, business schools with a grade non-disclosure norm typically have some method of revealing the very best students. This revelation usually takes the form of awards and honors. Moreover, at Chicago's Booth School of Business, where grade non-disclosure is the norm, MBA students also openly compete to earn teaching assistant positions, which are viewed as an indicator of excellent performance.[4] Concurrently, at many schools, including Wharton, students whose grades fall below some minimum standards are also dismissed.

This paper sets out to explain the main stylized facts about non-disclosure policies. First, what are conditions that lead to majority support of non-disclosure policies among students? Second, why is support for non-disclosure mainly concentrated within highly selective programs? Third, why isn't non-disclosure prevalent within other professional degree granting programs such as medicine, law, and accounting? Fourth, how does the introduction of awards and minimum grade requirements, which reintroduce some informativeness, impact the support for nondisclosure and welfare?

---

policies were subject to disciplinary action (Ciolli, 2007). There was concern about the legality of these requirements, which led the school to eventually remove them by 2005 [The Economist, "News from the schools." (Dec 16, 2005)]. At INSEAD, grade non-disclosure is still an official school policy and students and recruiters found in breach of it are subject to disciplinary action. [INSEAD MBA Career Services, "Grade Non Disclosure Policy," insidemba.insead.edu/careers/whatsnew/GradeNonDisclosure.pdf.

[3]The Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 CFR Part 99), sometimes called the Buckley Amendment, is a Federal law that protects the privacy of student education records.

[4]Devin Pope (private communication).

We construct a model with students, schools, and employers. Students prefer larger post-school wages but dislike studying. Schools are heterogenous in their selectivity (reputation). Under disclosure, employers can observe both a student's grades and the school's selectivity; under non-disclosure, an employer can only observe the partial signal of the school's selectivity.

Section 2 then derives the conditions that support grade non-disclosure under a non-overlapping property of the relationship between effort and grades. This assumption allows for the most straightforward interpretation of our results. We show that students at elite schools are the most likely to adopt a non-disclosure policy, subsequently reducing their effort. Intuitively, a non-disclosure policy allows the median voter to study less and then pool to receive the expected (mean) wage, which might be more valuable to her than receiving the median wage with effort. For plausible wage distributions, the desire to pool becomes more valuable at more selective schools. Section 3 then generalizes the results to the case where the non-overlapping property is relaxed. Section 4 explains why grade non-disclosure is not common in other professional degree programs, including medicine, law and accounting. In particular, minimum certification requirements undercut the support for grade non-disclosure in those programs by requiring a level of effort that is complementary to studying, thereby reducing the value from pooling to reduce effort. Section 5 extends the analysis to allow for awards (similarly, honors) as well as minimum grades. Interestingly, while a minimum grade requirement encourages additional effort from lower ability students, the requirement also reinforces the support for non-disclosure by raising the expected pooled wage with low effort without changing the median wage with high effort. In contrast, awards reduce the support for non-disclosure by allowing higher-ability students to separate, thereby reducing the expected wage earned by the pool of students who have not obtained this additional signal. Section 6 then draws some conclusions, including a discussion of some mechanisms that schools could potentially deploy to undermine student majority support for grade non-disclosure.

2 Basic Model

2.1 The Environment

Consider a model of students, schools, and employers. Students have ability types θ distributed in a (possibly unbounded) closed interval $[\theta_0, \theta_1]$ with $0 \leq \theta_0 < \theta_1 \leq +\infty$. Effort at school is binary $e \in \{0, 1\}$.[5] Schools are characterized by a selectivity parameter $\alpha \in \mathbb{R}$, which affects the distribution of accepted students.[6] A school with selectivity α accepts a continuum of students distributed according to an atomless cumulative distribution function μ_α. We assume that more selective schools pick a better distribution of students. Formally, $\alpha > \alpha'$ implies that μ_α first-order stochastically dominates (FOSD) $\mu_{\alpha'}$. Let E_α denote the expectation operator with respect to the distribution μ_α and let θ_α^{Median} denote the median type under distribution μ_α.

[5] In the Appendix, we show that our results generalize to the case of a continuum of effort levels.

[6] To determine admission, schools can observe some pre-school signal of performance (e.g., undergraduate transcript, GMAT, etc.) to determine if it exceeds the school's α.

Each student has a utility function

$$U(w,e) = w - c(e),$$

where w is the student's wage and e is the studying effort. The costs of effort are $c(1) := c > 0$, $c(0) = 0$.[7]

We make the following assumptions about the grade technology:

Assumption 1 (Non-overlapping grades). g is continuous, strictly increasing, and satisfies $g(\theta_0, 1) \geq g(\theta_1, 0)$.

Continuity of g is a technical assumption. The assumption that g is increasing states that individuals with greater ability and higher effort get higher grades. The third assumption states that even the highest skilled student gets a lower grade than the lowest ability student if the highest type doesn't study (i.e., the sets of grades under high and low efforts do not overlap). This assumption, which will be dropped in Section 3, simplifies the analysis of the equilibrium with disclosure.

Upon graduation, students obtain jobs in a competitive market of employers. A student with ability θ who exerts studying effort e has productivity $f(\theta, e) = \theta + \kappa e$, where $\kappa \geq 0$ parametrizes the human capital component of education.[8] When $\kappa = 0$, studying effort does not affect productivity and grades have purely a signaling aspect. When $\kappa > 0$, education also has a human capital dimension and studying increases the student's productivity.

As discussed in the introduction, students determine the school's disclosure policy through a voting procedure. We formalize this mechanism by assuming that the voting procedure selects the policy preferred by the majority of students.[9]

The timing of the game is as follows:

t=1. *Voting:* A pool of students with types θ distributed according to μ_α joins the school. The disclosure policy $d \in \{D, ND\}$ is determined by majority voting, where D denotes a policy of disclosing grades and ND denotes not disclosing grades.

t=2. *Effort:* Each student chooses a level of effort $e \in \{0, 1\}$ and obtains a grade $g(\theta, e)$.

t=3. *Market Wage:* A competitive market of employers observes the school's selectivity α, the school policy's policy on grade non-disclosure, and, if allowed, grades. The competitive market offers a wage w equal to the expected productivity of the student.

[7] The minimum effort level denotes the amount a student would choose in the absence of explicit incentives. While we have normalized it to zero, we do not claim that students would not study at all in the absence of explicit incentives.

[8] In the Appendix, we characterize the solution of our model for general production functions $f(\theta, e)$.

[9] When there are finitely many students, this procedure picks the unique equilibrium outcome from the strategic voting game. With infinitely many students, each student has mass zero and, therefore, is indifferent between voting in either policy. As a result, all policies can be an equilibrium of the the voting game. The procedure then picks the equilibrium in which students vote on their preferred strategies. This is the only outcome that can be approached taking the limit of the game with finitely many players as the number of players grows.

2.2 Equilibrium

We study the Perfect Bayesian Equilibrium of the game.

Definition 1. A Perfect Bayesian Equilibrium (PBE) of the game is a profile of strategies $\{e_D(\theta), e_{ND}(\theta), w_D(G), w_{ND}\}$, a disclosure policy $d \in \{D, ND\}$, and beliefs $\{\beta_D(\cdot \mid G), \beta_{ND}(\cdot)\}$ such that

1. Each student's strategy is optimal given the wage schedule and the disclosure policy:

$$\begin{aligned} e_D(\theta) &\in \arg\max_e w_D(g(\theta, e)) - c(e), \\ e_{ND}(\theta) &\in \arg\max_e w_{ND} - c(e). \end{aligned}$$

2. Employers earn zero profits given beliefs:

$$\begin{aligned} w_D(G) &\equiv \int_\theta f(\theta, e(\theta))\, d\beta_D(\theta | g(\theta, e(\theta)) = G), \\ w_{ND} &\equiv \int_\theta f(\theta, e(\theta))\, d\beta_{ND}(\theta). \end{aligned}$$

3. The disclosure policy satisfies a majority rule:

$$\int_\theta \mathbf{1}(U_D(\theta) > U_{ND}(\theta))\, d\mu_\alpha > \frac{1}{2} \implies d = D, \text{ and}$$

$$\int_\theta \mathbf{1}(U_{ND}(\theta) > U_D(\theta))\, d\mu_\alpha > \frac{1}{2} \implies d = ND,$$

where $\mathbf{1}$ denotes the indicator function and $U_d(\theta)$ denotes the payoff of type θ under disclosure policy d.[10]

4. Beliefs are consistent:

(a) $\beta_{ND}(\cdot)$ is derived from the student's strategy using Bayes' rule,

(b) $\beta_D(\cdot \mid G)$ is derived from the student's strategy using Bayes' rule whenever $G = g(\theta, e_D(\theta))$ for some θ, and

(c) $\beta_D(\cdot \mid G)$ assigns mass zero to all types for which $g(\theta, 0) \neq G$ and $g(\theta, 1) \neq G$.

Conditions (1) and (2) are the standard perfection requirements, stating that players choose their actions optimally given beliefs and other players' actions. Condition (3) is the majority rule requirement for the disclosure policy. Conditions (4a) and (4b) require beliefs on the equilibrium path of the continuation games after the disclosure policy is determined to satisfy Bayes' rule.[11] Condition (4c) states that beliefs conditional on a grade cannot attach a positive mass to types for whom this grade is unattainable.

We study the equilibrium of the game backwards.

[10] Formally, $U_{ND}(\theta) \equiv w_{ND} - c(e_{ND}(\theta))$ and $U_D(\theta) \equiv w_D(g(\theta, e_D(\theta)) - c(e_D(\theta)))$.

[11] This game has two proper subgames, each starting after a disclosure policy is determined. Consistency of beliefs for the continuation games after the disclosure policies are determined, therefore, amounts to imposing subgame perfection.

t=3. *Market Wages:* Conditions (2) and (4) require that, in the case of non-disclosure, employers offer wages equal to the student's expected productivity:

$$w_{ND} = E_\alpha \left[\theta + \kappa e_{ND}(\theta)\right]. \tag{1}$$

In the case of disclosure, the market wage is equal to the student's expected productivity for all grades on the equilibrium path:

$$w_D(G) = E_\alpha \left[\theta + \kappa e_D(\theta) | g(\theta, e_D(\theta)) = G\right],$$

whenever $G = g(\theta, e_D(\theta))$ for some θ.

Because g is strictly increasing in θ and $g(\theta_0, 1) \geq g(\theta_1, 0)$, each grade uniquely identifies the student's skills. Condition (4c) implies that the market must assign probability 1 to type θ when an off-equilibrium-path grade is only achievable by type θ. Therefore, the wage schedule is determined by

$$w_D(g(\theta, e)) = \theta + \kappa e, \tag{2}$$

for $e \in \{0, 1\}$ and $\theta \in [\theta_0, \theta_1]$.

t=2. *Effort:* Under no disclosure, students are offered the same wage regardless of their grades. Then, because effort is costly and does not affect wages, condition (1) implies that all types choose zero effort: $e_{ND}(\theta) = 0$.

Under disclosure, a type-θ student who exerts effort e gets utility $\theta + \kappa e - c$. Thus, the high effort is chosen if $\kappa \geq c$ and the low effort is chosen in $\kappa \leq c$. The student's utility under disclosure is

$$\theta + \max\{\kappa - c, 0\}$$

t=1. *Voting:* Consider the students' preferences over disclosure and non-disclosure. Type θ prefers *disclosure* if

$$\theta + \max\{\kappa - c, 0\} \geq E_\alpha[\theta]. \tag{3}$$

The student's voting decision balances her own disclosed productivity under *disclosure* (the left-hand side of equation (3)) against the expected pooled wage with low effort that she would receive under *non-disclosure* (the right-hand side of equation (3)). Of course, that decision depends on whether, under disclosure, it is efficient for the student to actually study ($\kappa \geq c$), thereby earning $\theta + (\kappa - c)$ rather than just θ.

2.3 Majority Rule and the Role of Selectivity

Turning now to the majority outcome, the decision made by the median voter, for a given level of school selectivity α, must, therefore, balance the *median* wage under *disclosure* ($\theta_\alpha^{Median} + \max\{\kappa - c, 0\}$) against the pooled *mean* wage with low effort under *non-disclosure* ($E_\alpha[\theta]$). A vote for disclosure allows the median voter to reveal his or her actual productivity; the median voter will then choose high effort if it is efficient to do so ($\kappa \geq c$). A vote for non-disclosure, however, allows the median voter to essentially "free ride" off of the expected pooled wage, which will be advantageous when there are enough students who are more productive than the median voter.

Proposition 1. *In any PBE, disclosure is chosen if* $\theta_\alpha^{Median} + \max\{\kappa - c, 0\} \geq E_\alpha[\theta]$ *and non-disclosure is chosen if* $\theta_\alpha^{Median} + \max\{\kappa - c, 0\} \leq E_\alpha[\theta]$.

The disclosure policy, therefore, depends on the skewness of the distribution of ability. When the distribution of ability is symmetric, the median is equal to the mean. Then, there always exists an equilibrium with grade disclosure. Moreover, when high effort is efficient $\kappa > c$, no equilibrium features grade non-disclosure.

However, actual wage distributions exhibit the empirical property that the median wage is below the mean (consistent with example distributions considered below). It is, therefore, reasonable to assume that the median ability is lower than mean ability: $\theta_\alpha^{Median} < E_\alpha[\theta]$. Therefore, as long as the human capital parameter κ is "not too large," non-disclosure is chosen.

Corollary 1. *Suppose* $\theta_\alpha^{Median} < E_\alpha[\theta]$. *There exists* $\kappa_\alpha > c$ *such that, in any PBE, non-disclosure is chosen if* $\kappa \leq \kappa_\alpha$ *and disclosure is chosen if* $\kappa \geq \kappa_\alpha$.

Greater school selectivity α, however, has an ambiguous impact on the support for non-disclosure. A larger value of α increases both median and mean abilities. If the mean ability is more responsive to a change in selectivity than the median, increasing selectivity would raise the proportion of people voting for non-disclosure. Then, the most selective schools would be the ones whose students vote to implement a grade non-disclosure policy. Formally, let $G(\alpha) \equiv E_\alpha[\theta] - \theta_\alpha^{Median}$ denote the mean-median gap. We, therefore, have:

Corollary 2. *Suppose* $G(\alpha)$ *is increasing. There exists* $\bar{\alpha} \in \mathbb{R} \cup \{-\infty, +\infty\}$ *such that, in any PBE, non-disclosure is chosen if* $\alpha > \bar{\alpha}$ *and disclosure is chosen in* $\alpha < \bar{\alpha}$.

Thus, more selective schools will adopt grade non-disclosure policies while less selective schools will not if the mean-median gap is increasing in school selectivity. The increasing mean-median gap assumption states that higher selectivity increases the quality of students by attracting a disproportionately higher amount of very good students. As we show in Examples 1 and 2, the conditions of Corollaries 1 and 2 are satisfied by the most commonly used wage distributions: those belonging to lognormal and the truncated normal families.

Example 1. Suppose $\theta \sim \text{lognormal}(\alpha, \sigma^2)$. Then, the mean $E[\theta] = e^{\alpha + \frac{\sigma^2}{2}}$ is always greater than the median $\theta_\alpha^{Median} = e^\alpha$ and the gap $G(\alpha) = e^\alpha \left(e^{\frac{\sigma^2}{2}} - 1\right)$ is increasing in α. Hence, the conditions of Corollaries 1 and 2 are satisfied. In fact, the equilibrium disclosure policy can be calculated analytically. Grade non-disclosure is selected if

$$\alpha \geq \begin{cases} \ln(\kappa - c) - \ln\left(e^{\frac{\sigma^2}{2}} - 1\right) \text{ if } \kappa > c \\ -\infty \text{ if } \kappa \leq c \end{cases}.$$

Example 2. Suppose θ is distributed according to a truncated normal with parameters (μ, σ^2, α), where α is the truncation parameter. Then, $E[\theta] = \mu + \frac{\pi\left(\frac{\alpha-\mu}{\sigma}\right)}{1-\Pi\left(\frac{\alpha-\mu}{\sigma}\right)}\sigma$, and $\theta_\alpha^{Median} = \mu + \Pi^{-1}\left(\frac{1+\Pi\left(\frac{\alpha-\mu}{\sigma}\right)}{2}\right)\pi$, where π and Π are the probability and cumulative distribution functions of the standard normal distribution. This distribution displays a positive mean-median gap for all truncation parameters α located within approximately 8 standard deviations from the mean μ. For realistic parameters μ, wages 8 standard deviations below the mean would give negative wages, and therefore, are unreasonable.

The mean-median gap is increasing in selectivity in the truncated normal case except when the truncation occurs in the middle of the right tail. Such high values for the truncation parameter, though, would produce an unusual shape for the wage distribution, with the mode being equal to the lowest value in the support and with most of the mass being concentrated very close to the truncation. Since typical wage distributions seem to have an interior mode, we find it reasonable to assume that $\alpha < \mu$, which implies that the gap is always increasing in selectivity.[12] Thus, for realistic truncation parameters, the conditions from Corollaries 1 and 2 are satisfied.

Therefore, under typical wage distributions, our model predicts that selective schools would adopt a grade non-disclosure policy whereas less selective schools would adopt a policy of disclosure. This pattern is consistent with the evidence from MBA programs (Tables I and II). Our model also predicts that studying effort is (weakly) lower with grade non-disclosure, which is also consistent with the evidence reported in Jain (2005).

3 Overlapping Grades

Thus far, we have assumed that education had such a strong effect on grades that the lowest ability student obtained a higher grade by studying than the highest ability student who did not study (non-overlapping grades). This assumption simplified the analysis because each grade fully identifies the student's skill. In practice, however, the highest ability student may be able to obtain relatively better grades even without exerting much effort. This section, therefore, allows grades to overlap. For simplicity, we focus on an additive grade technology:[13]

$$g(\theta, e) = \theta + \gamma e.$$

Assumption 2 (Overlapping Grades). $\theta_1 > \theta_0 + \gamma$ where $\gamma > 0$.

Under Assumption 2, which replaces Assumption 1, the set of possible grades under high and low efforts are now allowed to overlap. A high effort allows a type-θ student, at a cost of c, to obtain the same grade of a type $\theta + \gamma$ who chooses low effort.

It is helpful to partition the grade space into three intervals:

1. In the lowest interval, $[\theta_0, \theta_0 + \gamma)$, each grade G can only be obtained by the type $\theta = G$ under low effort. Consistency (Condition 4(c)) requires beliefs $\beta_D(\theta|G)$ to assign a unit mass at type G.

2. In the intermediate interval, $[\theta_0 + \gamma, \theta_1]$, grades can be obtained by two different types: $\theta = G$ under low effort and $\theta = G - \gamma$ under high effort. Consistency requires beliefs to assign zero mass to all other types.

[12] More precisely, the mean-median gap is positive if $\alpha \in [\mu - 8.437\sigma, \mu + 7.7073\sigma]$ and it is increasing if $\alpha < 0.210\sigma$.

[13] The Appendix presents results for general grade technologies.

3. In the highest interval, $(\theta_1, \theta_1 + \gamma]$, each grade G can only be obtained by the type $\theta = G - \gamma$ under high effort. Consistency requires beliefs to assign a unit mass to type $G - \gamma$.

The first two intervals are always non-empty whereas the third interval is empty when $\theta_1 = +\infty$. In this section, we will also assume that it is efficient to exert high effort: $\kappa > c$.

Definition 2. We will say that grades are *sufficiently responsive to effort* if $\gamma > c$, and we will say that they are *not sufficiently responsive to effort* if $\gamma < c$.

In words, grades are sufficiently responsive to effort if exerting high effort allows a student to pool with someone whose productivity exceeds the student's own productivity by an amount greater than the cost: $(\theta + \gamma) - \theta > c$.

3.1 Equilibrium with responsive grades

We first characterize the unique equilibrium of the continuation game after grade disclosure has been selected in the case of sufficiently responsive grades. Since this unique equilibrium involves full separation of types, the results from Section 2 remain unchanged in any PBE of the game.

Proposition 2. *Suppose grades are sufficiently responsive to effort. There exists a unique PBE of the continuation game conditional on grade disclosure in which $e_D(\theta) = 1$ for all θ. Furthermore, this PBE survives the intuitive criterion.*[14]

The proof of Proposition 2 will be presented through a series of lemmata. Intuitively, deviating from a high to a low effort causes the student to be pooled with someone with much lower productivity, and so this cannot be a profitable deviation when grades are sufficiently responsive to effort. Similarly, deviating in effort from low to high causes the student to be pooled with someone with higher productivity, which *is* a profitable deviation under sufficient responsiveness, ruling out low effort as an equilibrium. Therefore, only equilibria in which everyone studies hard can be sustained.

Lemma 1. *Suppose grades are sufficiently responsive to effort. There exists a PBE that survives the intuitive criterion in which $e_D(\theta) = 1$ for all θ.*

Proof. Let $e_D(\theta) = 1$ for all types and consider a deviation by a type θ to $e = 0$. Since this is a separating equilibrium, each type gets payoff $u^*(\theta) = \theta + \kappa - c$. If $\theta < \theta_0 + \gamma$, deviating fully identifies the type and the market offers wage $w = \theta$. The deviation is not profitable since $\kappa > c \implies \theta + \kappa - c > \theta$. If a type belongs to either the second or the third interval, deviating leads to grade θ, which is the equilibrium grade of type $\theta - \gamma$, who exerts high effort. Therefore, the market offers wage $w = \theta - \gamma + \kappa$. This deviation is also not profitable since $\gamma > c \implies \theta + \kappa - c > \theta + \kappa - \gamma$. Hence, $e_D(\theta) = 1$ is an equilibrium.

But it remains to be shown that it survives the intuitive criterion. The set of off-equilibrium-path grades is $[\theta_0, \theta_0 + \gamma)$. Any grade G in this interval can only be obtained by type $\theta = G$. Let $v(G, w, \theta)$ denote the payoff of type θ when she obtains grade G (or, in

[14] See Cho and Kreps (1987) for a presentation of the intuitive criterion.

the standard signaling terms, sends message G) and the market offers wage w. Message G is dominated by the equilibrium payoffs since

$$u^*(G) = G + \kappa - c > v(G, G, G).$$

Thus, the equilibrium trivially satisfied the intuitive criterion. $\square$

Lemma 2. *Suppose grades are sufficiently responsive to effort. There is no PBE in which $e_D(\theta) \neq 1$ for some θ.*

Proof. Suppose, by contradiction, we have a PBE in which $e_D(\theta) = 0$ for some type θ. Now consider a deviation to $e = 1$. There are 3 possibilities: she separates herself, she obtains a grade already being taken by type $\theta + \gamma$, or she obtains a grade that is off-the-equilibrium path. If she separates herself, she obtains payoff $\theta + \kappa - c > \theta = u^*(\theta)$. Thus, the deviation is profitable. If she takes the grade already taken by type $\theta + \gamma$, she obtains $\theta + \gamma - c > \theta$, which is also a profitable deviation. Finally, if the student obtains an off-the-equilibrium-path grade, she obtains

$$\lambda(\theta + \kappa) + (1 - \lambda)(\theta + \gamma) - c = \lambda(\theta + \kappa - c) + (1 - \lambda)(\theta + \gamma - c),$$

where $\lambda \in [0, 1]$ are the market's beliefs about the probability of the deviant type being θ. Since $\theta + \kappa - c > \theta$ and $\theta + \gamma - c > \theta$, it follows that this term is greater than the equilibrium payoff $u^*(\theta) = \theta$ for any belief λ. Therefore, we cannot have a PBE in which $e(\theta) = 0$. $\square$

From Condition 4 of Definition 1, beliefs $\mu(\theta|G)$ must assign a unit mass at $\theta = G - \gamma$ for all $G \geq \theta_0 + \gamma$ and a unit mass at $\theta = G$ for all $G < \theta_0 + \gamma$. Furthermore, in any PBE, the market must offer the following wage schedule:

$$w(G) = \begin{cases} G - \gamma + \kappa \text{ if } G \geq \theta_0 + \gamma \\ G \text{ if } G < \theta_0 + \gamma \end{cases}.$$

Thus, Lemmata 1 and 2 imply that the PBE is unique. Because the unique PBE of the continuation game has full separation, each type obtains payoff $\theta + \kappa - c$ under grade disclosure. Proceeding as in the previous section, we obtain the following results:

Proposition 3. *Suppose grades are sufficiently responsive to effort.*

1. In any PBE, disclosure is chosen if $\theta_\alpha^{Median} + \kappa - c > E_\alpha[\theta]$ and non-disclosure is chosen if $\theta_\alpha^{Median} + \kappa - c \leq E_\alpha[\theta]$.

2. Suppose $G(\alpha)$ is increasing. There exists $\bar{\alpha} \in \mathbb{R} \cup \{-\infty, +\infty\}$ such that, in any PBE, non-disclosure is chosen if $\alpha > \bar{\alpha}$ and disclosure is chosen in $\alpha < \bar{\alpha}$.

Therefore, the results from the previous section immediately generalize when we allow for overlapping grades if we assume that grades are sufficiently responsive to effort. Next, we consider the case of non-responsive grades.

3.2 Equilibrium with non-responsive grades

When grades are not sufficiently responsive to effort, there may be multiple equilibria depending on the distribution of types. We will say that a PBE is *essentially unique* in a given class of equilibria if all PBE in that class have the same grade and wage schedules, and feature the same beliefs for all grades on the equilibrium path. The following proposition states that there exists an essentially unique PBE with non-decreasing effort.[15] Moreover, this equilibrium satisfies the intuitive criterion refinement.

Proposition 4. *Suppose grades are not sufficiently responsive to effort. There exists an essentially unique PBE with non-decreasing effort, in which*

$$e(\theta) = \begin{cases} 1 \text{ if } \theta > \theta_1 - \gamma \\ 0 \text{ if } \theta \leq \theta_1 - \gamma \end{cases}.$$

Moreover, this PBE survives the intuitive criterion.

Intuitively, when grades are not sufficiently responsive to effort, a type θ student looses relatively little in the form of wages by choosing low effort and pooling with a lower type $\theta - \gamma$. Therefore, all types have an incentive to reduce their effort and imitate a slightly less productive type while saving the cost of effort. Only types at the upper interval $(\theta_1 - \gamma, \theta_1]$ are able to choose high effort in equilibrium because there are no higher types to pool with them. When ability is unbounded ($\theta_1 = +\infty$), this upper interval does not exist and all types exert low effort despite the fact that it is efficient for all of them to choose high effort.

The formal proof of the proposition will be presented through a series of lemmata. The first lemma establishes that all types that can obtain grades in the highest interval $(\theta_1, \theta_1 + \gamma]$ do so:

Lemma 3. *Suppose grades are not sufficiently responsive to effort. In any PBE with non-decreasing effort, $e(\theta) = 1$ for all $\theta > \theta_1 - \gamma$.*

Proof. Consider a PBE in which a type $\theta > \theta_1 - \gamma$ chooses $e = 0$. Since this equilibrium is nondecreasing, we cannot have $e(\theta - \gamma) = 1$. Therefore, type θ must be separated and gets payoff $u^*(\theta) = \theta$. Suppose this type deviates to $e = 1$. Because $g(\theta, 1) = \theta + \gamma > \theta_1$, consistency of beliefs (Condition 4(c) of Definition 1) implies that the market would assign probability one to his true type and would offer wage $\theta + \kappa$. Thus, the student would get payoff $\theta + \kappa - c > \theta = u^*(\theta)$, contradicting the assumption of this being an equilibrium. □

Next, we show that types who are unable to obtain grades in the highest interval choose low effort:

Lemma 4. *Suppose grades are not sufficiently responsive to effort. In any non-decreasing PBE, $e(\theta) = 0$ for all $\theta \leq \theta_1 - \gamma$.*

Proof. Consider a PBE in which $e(\underline{\theta}) = 1$ for some type $\underline{\theta} < \theta_1 - \gamma$. Since this equilibrium is nondecreasing and, by the previous lemma, $e(\theta_1) = 1$, we must also have $e(\theta) = 1$ for all types in $[\underline{\theta}, \theta_1]$. In particular, $e(\theta_1 - \gamma) = 1$. Since both θ_1 and $\theta_1 - \gamma$ choose $e = 1$, they are both

[15] A PBE features non-decreasing effort if $e(\theta)$ is a non-decreasing function.

separated in equilibrium and obtain payoffs $u^*(\theta_1) = \theta_1 + \kappa - c$ and $u^*(\theta_1 - \gamma) = \theta_1 - \gamma + \kappa - c$. If type θ_1 deviates to $e = 0$, he obtains the same grade as type $\theta_1 - \gamma$, thereby obtaining a payoff of $\theta_1 - \gamma + \kappa > \theta_1 + \kappa - c = u^*(\theta_1)$. Thus, this is a profitable deviation. $\square$

Therefore, the only candidate for a non-decreasing equilibrium effort schedule is the one in which only types greater than $\theta_1 - \gamma$ exert high effort. The following lemma establishes that this schedule can be supported in equilibrium, and that such an equilibrium survives the intuitive criterion:

Lemma 5. *Suppose grades are not sufficiently responsive to effort. There exists a PBE in which* $e(\theta) = \begin{cases} 1 \textit{ if } \theta > \theta_1 - \gamma \\ 0 \textit{ if } \theta \le \theta_1 - \gamma \end{cases}$. *Moreover this PBE survives the intuitive criterion.*

Proof. First, we will show that such an equilibrium exists. Given the effort schedule specified in the statement of the lemma, it is useful to partition the type space in 3 intervals: $[\theta_0, \theta_1 - 2\gamma]$, $(\theta_1 - 2\gamma, \theta_1 - \gamma]$, and $(\theta_1 - 1, \theta_1]$.

Students choose $e(\theta) = 0$ in the first and second intervals and choose $e(\theta) = 1$ in the third interval. A type θ in the first interval who deviates to $e = 1$ obtains the equilibrium grade of type $\theta + \gamma \in [\theta_0 + \gamma, \theta_1 - \gamma]$. A type θ in the second interval who deviates to $e = 1$ obtains grade $\theta + \gamma \in (\theta_1 - \gamma, \theta_1]$, which is off the equilibrium path. Similarly, a type in the third interval who deviates to $e = 0$ obtains grades in the interval $(\theta_1 - \gamma, \theta_1]$, which is off the equilibrium path. Let off-the-equilibrium-path beliefs $\mu(\theta|G)$ assign a unit mass at type G (i.e., the market assigns probability 1 to the highest of the two possible types when an off-equilibrium grade is chosen), and define wages as the expected product given beliefs. We will verify that none of these possible deviations are profitable.

If a type θ in the first or second intervals deviates to $e = 1$, she is perceived to be type $\theta + \gamma$, yielding a payoff of $\theta + \gamma - c < \theta = u^*(\theta)$. If a type in the third interval deviates to $e = 0$, she is perceived to be type θ, yielding a payoff of $\theta < \theta + \kappa - c = u^*(\theta)$. Thus, there are no profitable deviations.

Next, we verify that this equilibrium survives the intuitive criterion. Recall that the set of off-equilibrium-path grades is $(\theta_1 - \gamma, \theta_1]$. Consider a grade G in this interval. There are now two types that can obtain such a grade: G and $G - \gamma$. Grade G is undominated for types G and $G - \gamma$, respectively, if the following inequalities hold:

$$u^*(G) = G + \kappa - c \le max\left\{\lambda G + (1 - \lambda)(G - \gamma + \kappa) | 0 \le \lambda \le 1\right\}, \text{ and}$$

$$u^*(G - \gamma) = G - \gamma \le max\left\{\lambda G + (1 - \lambda)(G - \gamma + \kappa) | 0 \le \lambda \le 1\right\} - c.$$

Since $\kappa > c \ge \gamma$, the maximum term is equal to $G - \gamma + \kappa$. Then, these conditions are both satisfied since

$$c \ge \gamma \implies G + \kappa - c \le G - \gamma + \kappa, \text{ and}$$

$$\kappa > c \implies G - \gamma \le G - \gamma + \kappa - c.$$

Hence, both types are undominated for message G.

The PBE fails the intuitive criterion if either of the following conditions hold:

$$min_{BR(G)} u(G, w, G) > u^*(G), \text{ and}$$

$$min_{BR(G)} u(G - \gamma, w, G) > u^*(G - \gamma),$$

where $BR(G)$ denotes the market's best response to grade G for some beliefs with support at the undominated types $\{G - \gamma, G\}$. Substituting the student's equilibrium payoff and the market's best response, the first condition becomes

$$G = min\{\lambda G + (1 - \lambda)(G - \gamma + \kappa)|0 \leq \lambda \leq 1\} > G + \kappa - c.$$

Since $\kappa - c > 0$, this inequality is false. The second condition becomes

$$G - c = min\{\lambda G + (1 - \lambda)(G - \gamma + \kappa)|0 \leq \lambda \leq 1\} - c > G - \gamma.$$

Because $c \geq \gamma$, this is also false. Therefore, the PBE survives the intuitive criterion. □

Since an effort schedule pins down beliefs and wages on the equilibrium path, the lemmata above establish the result from Proposition 4.

Proposition 4 implies that:

Corollary 3. *When grades are not sufficiently responsive to effort, in any PBE with non-decreasing effort:*

1. If $\theta_\alpha^{Median} > \theta_1 - \gamma$, disclosure is chosen when $\theta_\alpha^{Median} + \kappa - c > E_\alpha[\theta]$ and non-disclosure is chosen when $\theta_\alpha^{Median} + \kappa - c \leq E_\alpha[\theta]$;

2. If $\theta_\alpha^{Median} \leq \theta_1 - \gamma$, disclosure is chosen when $\theta_\alpha^{Median} \geq E_\alpha[\theta]$ and non-disclosure is chosen when $\theta_\alpha^{Median} \leq E_\alpha[\theta]$.

If the distribution of skills is unbounded ($\theta_1 = +\infty$), non-disclosure is adopted if the mean-median gap is positive and disclosure is adopted if it is negative.

The nondecreasing effort restriction is not innocuous. Depending on the distribution of types, other equilibria may exist. For example, when types are uniformly distributed and $c \in \left[\frac{3\gamma - \kappa}{2}, \frac{\kappa + \gamma}{2}\right]$, there exists a PBE in which $e(\theta) = \begin{cases} 0 \text{ if } \theta \in [\theta_0 + \gamma, \theta_1 - \gamma] \\ 1 \text{ if } \theta \in [\theta_0, \theta_0 + \gamma) \cup (\theta_1 - \gamma, \theta_1] \end{cases}$.

This (non-monotonic) equilibrium also survives the intuitive criterion.[16]

In the rest of the paper, we will select the non-increasing PBE when considering the model with sufficiently unresponsive grades.

[16] In equilibria with pooling, the condition for grade disclosure to be chosen becomes slightly different. For any α pick one (possibly non-monotone) equilibrium and denote by $\phi_\alpha(\theta)$ the payoff of type θ in continuation game after disclosure is chosen. Let ϕ_α^{Median} denote the median payoff. Non-disclosure is chosen if $\phi_\alpha^{Median} < E_\alpha[\theta]$ and disclosure is chosen if $\phi_\alpha^{Median} \geq E_\alpha[\theta]$.

4 Certification

Business schools are unique in that most other professional programs – including medicine, law and accounting – allow for grade disclosure.[17] These other programs, however, also have some uniform certification (medical boards, legal bar, CPA) that is required to practice at the fullest level. We now show how the existence of these external minimum standards makes it *harder* to sustain an equilibrium with grade *non-disclosure*.

Although the results are more general, we consider a simple formulation of certification. There are two types of effort: studying for classes, denoted by e, and studying for the certification exam, denoted by s. For simplicity, we maintain the assumption of binary efforts, $e, s \in \{0, 1\}$, and keep the assumption of an additive production function:

$$f(\theta, e, s) = \theta + \kappa e + \eta s,$$

where $\eta \in \mathbb{R}$ captures the effect of studying for the exam on the student's productivity. When $\eta = 0$, studying for the exam does not affect the student's productivity.

The cost of effort is represented by the function $c(e, s)$, satisfying the following properties:

Assumption 3. c is strictly increasing and satisfies decreasing differences.

The assumption that c is strictly increasing means that both efforts are costly, while decreasing differences states that studying for classes makes it easier to study for the exam so that e and s are "cost-complements." For example, there is usually some overlap between the material covered in class and the material tested in the certification exam, which would make obtaining the certification easier if one studies for class.

In the presence of certification, all students are required to exert effort $s = 1$ in order to work. In the absence of certification, students do not exert such effort (it is not observable by firms and, therefore, it is costly but does not raise their wages). The game is exactly the same as in the model of Sections 2 and 3, with the exception that in the presence of certification workers are required to exert effort $s = 1$ in order to pass the certification exam.

The following proposition states that it is easier to support grade non-disclosure when there is no certification exam. Intuitively, when studying for classes and for the certification exam are cost complements, certification reduces the incremental cost of studying for classes, thereby increasing the value of disclosure.

Proposition 5. *Consider the model of either Section 2 or Section 3 and suppose Assumption 3 holds. If there exists an equilibrium with grade disclosure under non-certification, there also exists an equilibrium with grade disclosure under certification.*

Proof. Under Assumption 3, the relevant conditions for grade non-disclosure under certification and non-certification are

$$\theta_\alpha^{Median} - E_\alpha[\theta] \le c(1,1) - c(0,1), \text{ and} \tag{4}$$

$$\theta_\alpha^{Median} - E_\alpha[\theta] \le c(1,0) - c(0,0). \tag{5}$$

[17] Limited exceptions were noted earlier.

Suppose inequality (4) is satisfied. Then,

$$\begin{aligned} \theta_\alpha^{Median} - E_\alpha[\theta] &\leq c(1,1) - c(0,1) \\ &\leq c(1,0) - c(0,0), \end{aligned}$$

where the second inequality follows from the decreasing differences property of c. Hence, (5) is also satisfied. $\square$

5 Minimum Grade Requirements and Awards

In the presence of grade non-disclosure, schools still have some limited tools available to encourage effort. The two most common instruments are awards and honors as well as some minimum performance requirement. Based on our conversations with MBA offices at most of these schools, awards tend to be more emphasized at schools with a non-disclosure policy, often as an explicit attempt to challenge grade non-disclosure. "Implicit" distinctions also exist in other forms, including winning a teaching assistant position. Concurrently, some MBA programs also impose certain minimum requirements that require some non-trivial amount of effort. For example, students at Wharton are dismissed if they score in the bottom decile in at least five credit unit courses during their first year or eight credit unit courses over two years, a rule passed during the 1998 school year by faculty in response to grade non-disclosure. We will address the minimum requirements first.

5.1 Minimum Grades

Imposing a minimum grade has two effects. On the one hand, it may prevent individuals with the lowest skills from being able to graduate. On the other hand, it induces those with intermediate skills to exert high effort. Formally, let $\bar{g}$ denote the minimum grade. A type-θ student is able to obtain the degree under effort e if $g(\theta, e) \geq \bar{g}$.

We assume that the market cannot determine if a student attended a school but was unable to obtain the minimum grade or if the student never attended the school, and denote the expected productivity of someone who did not attend school by $w < E_\alpha[\theta] - c$ for all α. Assume that exerting high effort is efficient $\kappa > c$ (otherwise there would be no gains from incentivizing effort).

In the case of grade non-disclosure, students who are able to meet the minimum grade requirement exert the minimum effort needed to do so:

$$e_{ND}(\theta) = \begin{cases} 0 \text{ if } g(\theta,0) \geq \bar{g} \text{ or } g(\theta,1) < \bar{g} \\ 1 \text{ if } g(\theta,0) < \bar{g} \leq g(\theta,1) \end{cases}.$$

Model with non-overlapping grades. In in the model with non-overlapping grades (Section 2), it is possible to induce all students to exert high effort by setting the minimum grade $\bar{g} \in (g(\theta_1, 0), g(\theta_0, 1)]$, that is, at level above the grade that would be obtained by a highest ability student under low effort. When the equilibrium of the model without minimum grades features disclosure (i.e., $\theta_\alpha^{Median} + \kappa - c \geq E_\alpha[\theta]$), the minimum grade is innocuous since all students would already choose a high effort.

When the equilibrium features non-disclosure ($E_\alpha[\theta] \geq \theta_\alpha^{Median} + \kappa - c$), the minimum grade policy shifts the equilibrium effort of all students from low to high, which increases their payoffs by $\kappa - c > 0$. Moreover, they still vote for non-disclosure since

$$E_\alpha[\theta] + \kappa - c > E_\alpha[\theta] \geq \theta_\alpha^{Median} + \kappa - c.$$

Hence, any equilibrium with a minimum grade $\bar{g} \in (g(\theta_1, 0), g(\theta_0, 1)]$ is preferred by all students relative to the equilibrium without minimum grade. The minimum grade requirement eliminates "free riding" off of the reduced signal under non-disclosure, leading to a Pareto improvement (it increases all students' payoffs while leaving firms with the same profit as before).

Model with overlapping grades. In the model of Section 3, it is impossible to simultaneously ensure that all types choose high effort and all types achieve the minimum grade under non-disclosure. If the minimum grade is set below $g(\theta_1, 0)$, some types will choose low effort. If it is set above $g(\theta_0, 1)$, some types will be unable to achieve the minimum grade. Since $g(\theta_1, 0) < g(\theta_0, 1)$, no minimum grade is able to simultaneously avoid both issues. Nevertheless, a minimum grade that is high enough to require effort from the lowest types but low enough to make sure that all students are able to pass increases ex-ante welfare.[18] Therefore, any mechanism that determines the minimum grade policy by an efficiency criterion would select an interior minimum grade. Such a mechanism may be the outcome of the school maximizing student welfare, profits, or a combination of both. Moreover, there exists an interior minimum grade that is preferred by the majority of students and would, therefore, be selected by a majority rule voting procedure.

Proposition 6. *Consider the model of either Section 2 or 3. Suppose it is efficient to exert high effort $\kappa > c$ and all equilibria have grade non-disclosure $\theta_\alpha^{Median} + \kappa - c < E_\alpha[\theta]$. Implementing a minimum grade $\bar{g} \in \left(g(\theta_0, 0), g\left(\theta_\alpha^{Median}, 0\right)\right]$ strictly increases the ex-ante welfare $E_\alpha[u^*(\theta)]$, and is strictly preferred by the majority of students (relative to a policy of no minimum grades).*

Proof. Consider either the model of Section 2 or Section 3. Since μ is an atomless distribution, $\theta_\alpha^{Median} > \theta_0$ for all α. Let $\bar{g} \in \left(g(\theta_0, 0), g\left(\theta_\alpha^{Median}, 0\right)\right]$. Under no minimum grades, all PBE have grade non-disclosure and low effort. Under the minimum grade policy, the median type still chooses low effort in case of non-disclosure but all types θ such that $g(\theta, 0) < \bar{g}$ exert high effort in order to achieve the minimum grade. Let $\theta^* \in (\theta_0, \theta_\alpha^{Median})$ be the first type who is able to pass with low effort: $g(\theta^*, 0) = \bar{g}$. Payoffs in any PBE without minimum grade are $E_\alpha[\theta]$ whereas payoffs in any PBE with the minimum grade $\bar{g}$ are

$$\begin{cases} E_\alpha[\theta] + \mu(\theta^*)\kappa - c \text{ if } \theta < \theta^* \\ E_\alpha[\theta] + \mu(\theta^*)\kappa \text{ if } \theta \geq \theta^* \end{cases}.$$

Since $\theta^* < \theta_\alpha^{Median}$, it follows that the median type is better off under minimum grade: $E_\alpha[\theta] + \mu(\theta^*)\kappa > E_\alpha[\theta]$. Taking the expectation of payoffs with respect to θ, yields

$$E_\alpha[\theta] + \mu(\theta^*)(\kappa - c) > E_\alpha[\theta],$$

which establishes that ex-ante welfare is higher under the minimum grade policy. ∎

[18] Ex-ante welfare is the expected utility computed at time "$t = 0$," before students know their ability θ. It is also the utilitarian welfare criterion.

A minimum grade $\bar{g} \in \left(g\left(\theta_0, 0\right), g\left(\theta_\alpha^{Median}, 0\right)\right]$ increases the payoff of the median student in the presence of grade non-disclosure by inducing effort from types with lower ability, thereby increasing the mean wage. It does not, however, affect payoffs in the presence of a grade disclosure policy. Therefore, *a minimum grade requirement increases the support for grade non-disclosure*. Paradoxically, the same minimum grade policies that have been enacted as a reaction to grade non-disclosure may be helping grade non-disclosure to perpetuate.

5.2 Awards and Prizes

Although schools are not allowed to disclose grades to potential employers, they are allowed to distribute awards and honors to students with "exceptional performance." Moreover, because the law treats awards and honors as directory information, schools may disclose this information publicly.

Consider an award or honor given to a fraction of students with the highest grades. Formally, the award is modeled as a binary signal distinguishing the students with grades in the top ϕ percentile of the grade distribution from other students. Since this signal does not reveal any additional information when grades are disclosed, it does not affect the equilibrium of the continuation game after a grade disclosure policy has been selected. However, in the case of grade non-disclosure, allowing students with the highest grades to separate themselves reduces the mean wage of the students who have not received such a distinction.

More formally, consider the continuation game after non-disclosure has been selected (the continuation game after disclosure is selected is trivial). There are no pure strategy equilibria. To see why, let θ_α^* denote the lowest type in the the top ϕ percentile of the *type distribution* and let $\theta_\alpha^{**} > \theta_\alpha^*$ denote the lowest type that can be sure to receive the prize even with low effort:

$$\mu(\theta_\alpha^*) = 1 - \phi \text{ and } \int_{g(\theta,1) \geq g(\theta_\alpha^{**},0)} d\mu_\alpha\left(\theta\right) = \phi.$$

If a positive mass of types $\theta \in (\theta_\alpha^*, \theta_\alpha^{**})$ chooses low effort, types slightly below θ_α^* prefer to exert high effort and get the prize. However, if types slightly below θ_α^* choose high effort, all types in $(\theta_\alpha^*, \theta_\alpha^{**})$ prefer to choose high effort as well and guarantee that they will get the prize. But if all types in $(\theta_\alpha^*, \theta_\alpha^{**})$ choose high effort, those below θ_α^* have no chance of getting the prize and, therefore, choose low effort. Yet, if all types below θ_α^* choose low effort, all types above θ_α^* can win the prize even with low effort and, therefore, choose low effort. Thus, we cannot have an equilibrium in pure strategies.

There exist, however, equilibria in mixed strategies. We characterize these equilibria in the Appendix but discuss their intuition here. Any mixed strategy equilibrium partitions the type space in three intervals. In the first interval, $[\theta_0, \underline{\theta}_\alpha)$ all types choose low effort and never win the prize. In the second interval, $[\underline{\theta}_\alpha, \bar{\theta}_\alpha]$, almost all types play strictly mixed strategies. They get the prize when they play $e = 1$ and they do not get the prize when they play $e = 0$. In the third interval, $(\bar{\theta}_\alpha, \theta_1]$, all types choose the low effort but still get the prize.

Since a positive mass of students exerts high effort, the ex-ante welfare in all of these equilibria is strictly greater than in the equilibria of the model with no awards when effort is efficient.

Proposition 7. *Consider the model of either Section 2 or 3. Suppose it is efficient to exert* *high effort $\kappa > c$ and all equilibria have grade non-disclosure $\theta_\alpha^{med} + \kappa - c < E_\alpha[\theta]$. Introducing an award strictly increases the ex-ante welfare.*

Proof. Existence of a mixed strategy equilibrium for the continuation game follows from standard arguments. Since the game does not have an equilibrium in which almost all types exert low effort, there must be a positive mass of types who exert high effort. Let $\lambda > 0$ denote the mass of such types. The ex-ante payoff is then $E_\alpha[\theta] + \lambda(\kappa - c)$ which is greater than the ex-ante payoff without the award $E_\alpha[\theta]$. $\square$

Whenever the prize is given to less than half of students, the majority will oppose it. Because the median voter will not be able to obtain the prize without exerting high effort, he must be at most indifferent between exerting high and low efforts. However, when the median voter exerts low effort, he does not win the prize and obtains a strictly lower wage than the wage he would obtain in the absence of this policy. Thus, a prize makes the median voter worse off.

More generally, the exclusion of the top of the distribution lowers the payoff under non-disclosure for all but the extreme types who can guarantee themselves to be in the top of the distribution $(\bar{\theta}_\alpha, \theta_1]$. Therefore, except when the distribution is sufficiently concentrated at the top and the award is given to a large proportion of students, an award policy reduces support for grade non-disclosure.

6 Conclusion

Grade non-disclosure is prevalent among elite MBA programs but not commonly found in lower ranked MBA programs as well as other professional degree programs with external certification. Students suggest that grade non-disclosure policies allow them to take more difficult classes. For example, the former president of the student association at Chicago Booth School of Business, April Park, argues that "grade non-disclosure allows students to take more challenging courses instead of taking classes in which they are over-qualified."[19]

However, self-reported levels of curriculum effort have actually fallen since grade non-disclosure policies were adopted. According to the chairman of Harvard's MBA program Richard Ruback, "numerous students had claimed that the non-disclosure policy resulted in little motivation to excel."[20] Akram Zaman, former co-president of the Harvard's Student Association's Executive Committee, claims that "there is a perception that general academic motivation and rigor has gone down [since the adoption of grade non-disclosure]. (...) People think that they won't be in the top fraction of the class and that they won't fail out, so many of them take on the attitude that they don't need to work as hard."[21] As a Stanford GSB student puts it, "[t]he grade non-disclosure policy is somewhat of a curse because it inspires a noticeable amount of apathy among the students. Class participation is good enough, but not as great as it could be if students were a little more compelled to prepare."[22]

[19] The Economist, ibid.

[20] The Economist, ibid.

[21] Harvard Crimson. "HBS Rethinks Grade Policy," November, 2005.

[22] Wise and Hauser (2007).

At Wharton, the amount of time spent on academics fell by 22% in the first four years after grade non-disclosure was implemented.[23] Moreover, the pattern of courses waved by students was not affected by the introduction of grade nondisclosure.[24] Therefore, the idea that grade non-disclosure allows students to take more challenging courses cannot account for the trend in average studying effort among MBA students. It is also not clear why the benefit from taking more difficult classes would be restricted to elite MBA programs, which is needed to explain the cross-sectional pattern of grade non-disclosure policies.

The Student Handbook of Stanford GSB mentions that grade non-disclosure improves student cooperation. This hypothesis is also problematic. As Kreps (2005) argues, "[h]elping one or two or ten classmates here probably does not materially affect your chances of landing a desirable job (...). There are 370 or so of you. How likely is it that you will compete for a given job with one of the people in your circle?" Even if there are gains from cooperation, the time-series evidence suggests that they are coming at the expense of reduced studying effort. Moreover, it is again unclear why the importance of cooperation is restricted to elite MBA programs and not other MBA or other professional degree programs.

Our model explains why students in elite programs might pass non-disclosure norms that reduce both their signal and their level of effort if they dislike studying. Our model also explains why non-disclosure appears to be fairly unique to elite MBA programs. Interestingly, minimum grade requirements, which were implemented as a way of combating non-disclosure, actually increases its support in equilibrium, precisely by enhancing the wage of colluding students. In contrast, awards reduce the support for non-disclosure by reducing the wage of colluding students who are not able to earn the distinction. External certification reduces the support for non-disclosure by reducing the incremental cost of studying.

Our results, therefore, identify two mechanisms – certification exams and awards – that can reduce support for non-disclosure in equilibrium. Most schools with non-disclosure norms currently utilize some form of prizes or honors. But MBA students are not subject to certification requirements in a way that would change the level of effort of the median student, which is likely substantially more powerful. Certification (or a minimum grade threshold that is more stringent than current practice) is both legal and is not generally practiced, as in other professional occupations. Future work could explore this issue in more detail.

Appendix

Equilibrium of Game with Awards

This section characterizes the mixed strategy PBE of the game with prizes under grade non-disclosure discussed in Subsection 5.2. For simplicity, will assume that the grade technology is additive as in Section 3: $g(\theta, e) = \theta + \gamma e$. We will omit α from the distributions for notational clarity.

[23] Business Week, "Campus Confidential." (Sep 12, 2005). As a Wharton student argues, "Wharton students don't always see their studies as their top priority, but instead look to balance their time between academics, career search, and social fun. This is made possible by the school's grade nondisclosure policy, which prevents students from sharing their grades with recruiters." Wise and Hauser, ibid.

[24] Jain (1997).

Let $a(\theta) \in [0,1]$ denote the probability of a type θ student choosing $e=1$. We assume that there is a continuum of students of each type. By the monotonicity of the grade technology, any equilibrium is characterized by two thresholds, $\underline{\theta}$ and $\bar{\theta}$ such that $a(\theta)=0$ if $\theta<\underline{\theta}$ or $\theta>\bar{\theta}$, and $a(\theta)\in(0,1)$ for almost all types $\underline{\theta}<\theta<\bar{\theta}$.[25] Assume that the distribution of types μ has a continuous probability distribution function, denoted by $\mu'(\theta)=\frac{\partial\mu(\theta)}{\partial\theta}$. Given the thresholds $\underline{\theta}$ and $\bar{\theta}$, we can compute the probability distribution of grades:

$$\pi(G)=\begin{cases} \mu'(G) & if \quad G\geq\bar{\theta}+\gamma \\ \mu'(G)+\mu'(G-\gamma)a(G-\gamma) & if \quad \bar{\theta}\leq G<\bar{\theta}-\gamma \\ \mu'(G)\left[1-a(G)\right]+\mu'(G-\gamma)a(G-\gamma)a(G-\gamma) & if \quad \underline{\theta}+\gamma\leq G<\bar{\theta} \\ \left[1-a(G)\right]\mu'(G) & if \quad \underline{\theta}\leq G<\underline{\theta}+\gamma \\ \mu'(G) & if \quad G<\underline{\theta} \end{cases}.$$

Integrating, we obtain the cumulative distribution of grades $\Pi(G)$. Types in the interval $[\underline{\theta},\bar{\theta}]$ must be indifferent between high and low efforts in order to play mixed strategies.

Let G^* denote the minimum grade required to get the prize: $\phi=1-\Pi(G^*)$. If $G^*>\underline{\theta}+\gamma$, types in $(\underline{\theta},G^*-\gamma)$ never win the prize and, therefore, cannot be indifferent between high and low effort. On the other hand, if $G^*<\underline{\theta}-\gamma$ types in the interval $(G^*+\gamma,\underline{\theta})$ win the prize even with low effort, and therefore cannot be indifferent between high and low efforts. Therefore, we must have $G^*=\underline{\theta}+\gamma$, i.e. all types that play mixed strategies win the prize if they choose $e=1$ and do not win if they choose $e=0$. Substituting in the cumulative distribution function of grades, we obtain:

$$1-\Pi(G^*)=\int_{G^*}^{1}\pi(G)dG=1-\mu(G^*)+\int_{G^*-\gamma}^{G^*}a(G)d\mu(G).$$

Therefore, the minimum grade G^* and the mixed strategies $a(G)$ must satisfy:

$$\phi=1-\mu(G^*)+\int_{G^*-\gamma}^{G^*}a(G)d\mu(G). \tag{6}$$

Let w_P and w_{NP} denote the wages offered for those who win and do not win the prize:

$$w_{NP}=E[\theta|G<G^*]=\frac{\int_{\theta_0}^{G^*}\theta d\mu(\theta)-\int_{G^*-\gamma}^{G^*}a(\theta)d\mu(\theta)}{1-\phi}, \text{ and}$$

$$w_P=E[\theta+\kappa e(\theta)|G\geq G^*]=\frac{\int_{G}^{\theta_1}\theta d\mu(\theta)+\int_{G^*-\gamma}^{G^*}a(\theta)(\theta+\kappa)d\mu(\theta)}{\phi}.$$

Types in the interval $[\underline{\theta},\bar{\theta}]$ must be indifferent between high and low efforts: $w_P-c=w_{NP}$. Substituting the expressions for w_P and w_{NP}, we obtain:

$$\frac{\int_{G}^{\theta_1}\theta d\mu(\theta)+\int_{G^*-\gamma}^{G^*}a(\theta)(\theta+\kappa)d\mu(\theta)}{\phi}-\frac{\int_{\theta_0}^{G^*}\theta d\mu(\theta)-\int_{G^*-\gamma}^{G^*}a(\theta)d\mu(\theta)}{1-\phi}=c,$$

[25] As we will show, all types in this intermediate interval are indifferent between both actions. Therefore, a measure zero of types may play pure strategies, which does not affect the probability of getting a prize.

which can be simplified as

$$\int_{G^*}^{\theta_1}\theta d\mu(\theta)+\int_{G^*-\gamma}^{G^*}\theta a(\theta)d\mu(\theta)-\phi\int_{\theta_0}^{\theta_1}\theta d\mu(\theta)+\kappa(1-\phi)\int_{G^*-\gamma}^{G^*}a(\theta)d\mu(\theta)=\phi(1-\phi)c. \quad (7)$$

Any real number $G^* \in (\theta_0+\gamma,\theta_1)$ and function $a(\theta):(G^*-\gamma,G^*)\to(0,1)$ satisfying conditions (6) and (7) is a PBE of the game.

Consider the case of types distributed according to a uniform $[0,1]$ distribution and suppose all types that play mixed strategies play the same strategy: $a(\theta)=\bar{a}\in(0,1)$ for all $\theta\in(G^*,G^*+\gamma)$. In this case, after tedious algebraic manipulations, it can be shown that the minimum grade to obtain the prize G^* solves the following equation:

$$G^2+\left[\frac{2}{3}(\kappa-1)(1-\phi)-\frac{\gamma}{3}\right]G+\frac{2(1-\phi)}{3}\left[\frac{\gamma}{2}-(1-\phi)\kappa-c\phi\right]=0$$

For $\phi\le\frac{\kappa-\gamma/2}{\kappa-c}$, the unique solution is the positive root of this equation. For example, when $\kappa=2$, $c=1$, $\phi=\frac{1}{10}$, and $\gamma=\frac{1}{4}$, we obtain $G^*=\frac{3\sqrt{16297}-93}{360}\approx 0.8055$ and $\bar{a}=\frac{18}{5}-4G^*\approx 0.3780$. In the unique mixed strategy equilibrium in which all players in the intermediate interval play the same strategy is such that

$$a(\theta)=\begin{cases}0 & if \quad \theta\le\underline{\theta}\approx 0.5555\\ 0.3780 & if \quad 0.5555\approx\underline{\theta}\le\theta\le G^*\approx 0.8055\\ 0 & if \quad \theta>G^*\approx 0.8055\end{cases}.$$

This is the unique PBE in which all types that play mixed strategies play the same strategy $\bar{a}$. However, this equilibrium is not unique. For example, any function $a(\theta)$ taking values between 0 and 1 such that

$$\int_{G^*-\frac{1}{4}}^{G^*}\theta a(\theta)d\theta=\frac{\bar{a}}{2}\left(G^*-\frac{1}{8}\right), \text{ and}$$

$$\int_{G^*-\frac{1}{4}}^{G^*}a(\theta)d\theta=\frac{\bar{a}}{4},$$

where $G^*=\frac{3\sqrt{16297}-93}{360}$ and $\bar{a}=\frac{18}{5}-4G^*$ is also a PBE of the model.

General Production Functions

This section considers the model of Section 2 under nonlinear production functions $f(\theta,e)$. Because students are risk neutral, one can also think of the productivity of type θ given effort e as a random variable with expected value equal to $f(\theta,e)$. We assume that f is non-decreasing in both arguments.

Equilibrium

Proceeding as in Section 2, the payoff of type θ under grade disclosure is

$$\phi(\theta)=\max\{f(\theta,1)-c,f(\theta,0)\},$$

whereas the payoff under grade non-disclosure is $E[f(\theta,0)]$. Since ϕ is strictly monotonic, the median of $\phi(\theta)$ is equal to $\phi(\theta_\alpha^{Median})$. Hence, we have the following result:

Proposition 8. *In any PBE, disclosure is chosen if $\phi(\theta_\alpha^{Median}) \geq E_\alpha[f(\theta,0)]$ and non-disclosure is chosen if $\phi(\theta_\alpha^{Median}) \leq E_\alpha[f(\theta,0)]$.*

In particular, when wages under low effort $f(\theta,0)$ are symmetrically distributed for every α, there always exists a PBE in which disclosure wins. For distributions such that median wage is below the average wage, non-disclosure is chosen if the median type's incremental productivity from effort is not too large:

Corollary 4. *Suppose $f(\theta_\alpha^{Median},0) < E_\alpha[f(\theta,0)]$. There exists $\kappa > c$ such that non-disclosure is chosen in any PBE if $f(\theta_\alpha^{Median},1) - f(\theta_\alpha^{Median},0) \leq \kappa$.*

Proof. If $f(\theta_\alpha^{Median},1) - f(\theta_\alpha^{Median},0) = c$, the result is immediate. Suppose $f(\theta_\alpha^{Median},1) - f(\theta_\alpha^{Median},0) > c$ and define $\varepsilon \equiv f(\theta_\alpha^{Median},1) - f(\theta_\alpha^{Median},0) - c > 0$. Then, we have

$$\begin{aligned}\phi\left(\theta_\alpha^{Median}\right) - f\left(\theta_\alpha^{Median},0\right) &= \max\{f(\theta_\alpha^{Median},1) - f(\theta_\alpha^{Median},0) - c, 0\} \\ &= \varepsilon.\end{aligned}$$

Hence,

$$\phi\left(\theta_\alpha^{Median}\right) = f\left(\theta_\alpha^{Median},0\right) + \varepsilon.$$

Since $f(\theta_\alpha^{Median},0) < E_\alpha[f(\theta,0)]$, setting ε small enough yields $\phi\left(\theta_\alpha^{Median}\right) < E_\alpha[f(\theta,0)]$. □

As in the model with additive production functions, if the mean output under non-disclosure is more responsive to a change in selectivity than the median output given disclosure, increasing selectivity would raise the proportion of people voting for non-disclosure. Let $\tilde{G}(\alpha) \equiv E_\alpha[f(\theta,0)] - \phi\left(\theta_\alpha^{Median}\right)$ denote the gap between the mean payoff under low effort and the median payoff under the optimal effort. Then, we have:

Corollary 5. *Suppose $\tilde{G}(\alpha)$ is increasing. There exists $\bar{\alpha} \in \mathbb{R} \cup \{-\infty,+\infty\}$ such that, in any PBE, non-disclosure is chosen if $\alpha > \bar{\alpha}$ and disclosure is chosen in $\alpha < \bar{\alpha}$.*

Let $\kappa \equiv \sup\{f(\theta,1) - f(\theta,0)\}$. The following proposition shows that, when the productivity given low effort follows a lognormal distribution, non-disclosure is chosen if the school's selectivity is sufficiently high.

Proposition 9. *Let $f(\theta,0) \sim$lognormal(α,σ^2). For any κ, there exists $\bar{\alpha}(\kappa)$ such that non-disclosure is chosen for all $\alpha > \bar{\alpha}(\kappa)$. Moreover, $\bar{\alpha}(\kappa)$* is a *non-decreasing function.*

Proof. Note that $\tilde{G}(\alpha) = E_\alpha[f(\theta,0)] - \max\left\{f\left(\theta_\alpha^{Median},1\right) - c, f\left(\theta_\alpha^{Median},0\right)\right\}$. Using the expression for the mean of the lognormal distribution, we obtain

$$\tilde{G}(\alpha) = e^{\alpha+\frac{\sigma^2}{2}} - \max\left\{f\left(\theta_\alpha^{Median},1\right) - c, f\left(\theta_\alpha^{Median},0\right)\right\}.$$

Since $f(\theta,1) \leq f(\theta,0) + \kappa$ for all θ, it follows that

$$\max\left\{f\left(\theta_\alpha^{Median},1\right) - c, f\left(\theta_\alpha^{Median},0\right)\right\} \leq f\left(\theta_\alpha^{Median},0\right) + \max\{\kappa - c, 0\}.$$

Thus,

$$\tilde{G}(\alpha) \geq e^{\alpha+\frac{\sigma^2}{2}} - f\left(\theta_\alpha^{Median},0\right) - \max\{\kappa - c, 0\} = e^{\alpha+\frac{\sigma^2}{2}} - e^{\alpha} - \max\{\kappa - c, 0\},$$

where the equality uses the expression for the median of a lognormal distribution. Since non-disclosure is chosen whenever $\widetilde{G}(\alpha) \geq 0$, a sufficient condition for grade non-disclosure to be chosen is

$$e^{\alpha+\frac{\sigma^2}{2}} - e^{\alpha} \geq \max\{\kappa - c, 0\}.$$

This expression is always true when $\kappa \leq c$. For $\kappa > c$, it is satisfied if

$$\alpha \geq \ln(\kappa - c) - \ln\left(e^{\frac{\sigma^2}{2}} - 1\right),$$

which completes the proof. $\square$

The generalization of our results on minimum grades and awards is straightforward. On the next subsection, we consider the generalization of our results on the effects of certification on non-disclosure.

Certification

There are two types of effort: studying for classes $e \in \{0,1\}$, and studying for the certification exam, $s \in \{0,1\}$. For simplicity, we assume that the productivity of a student is additively separable between efforts:

$$f(\theta, e, s) = g(\theta, e) + h(\theta, s).$$

The cost of effort is represented by the strictly increasing function $c(e, s)$, satisfying decreasing differences. In the presence of certification, all students are required to exert effort $s = 1$. We maintain the assumption that the distribution of productivities in the case of zero effort for class is skewed to the right so that the median is lower than the mean:

Assumption 4. $f(\theta_\alpha^{Median}, 0, s) \leq E_\alpha[f(\theta, 0, s)]$ for $s \in \{0,1\}$.

We also assume that studying for the certification exam does not increase the mean (expected) productivity by more than it increases the median productivity:

Assumption 5. $E_\alpha[h(\theta,1) - h(\theta,0)] \leq h(\theta_\alpha^{Median}, 1) - h(\theta_\alpha^{Median}, 0)$.

Assumption 5 is satisfied, for example, if certification changes productivity uniformly (i.e., h is constant in θ), or has no effect on productivity. It is also satisfied if certification helps lower types more than higher types (i.e., h has decreasing differences) and the distribution of the benefit of studying for certification is skewed to the right. For example, suppose h is linear:

$$h(\theta, s) = \beta(\theta_1 - \theta)s + \gamma,$$

where $\beta \geq 0$ in order to satisfy decreasing differences and $\gamma \in \mathbb{R}$. Assumption 5 is satisfied if and only if $\theta_\alpha^{Median} \leq E_\alpha[\theta]$.

It is reasonable to assume that a certification technology that ensures that all students have a minimum set of basic skills increases the productivity of unskilled students more than the productivity of skilled students (decreasing differences). Assumption 5 will then be satisfied as long as the distribution of abilities is skewed to the right, consistent with previous examples.

The following proposition states that it is easier to support grade nondisclosure when there is no certification exam. Intuitively, certification raises the median productivity more than the expected productivity, giving the median voter more incentive to want to reveal his own ability even at the cost of more effort.

Proposition 10. *If there exists an equilibrium with grade disclosure under non-certification, there also exists an equilibrium with grade disclosure under certification.*

Proof. Under Assumption 4, the relevant conditions for grade non-disclosure under certification and non-certification are

$$f(\theta_\alpha^{Median}, 1, 1) - E[f(\theta, 0, 1)] \leq c(1, 1) - c(0, 1), \text{ and} \tag{8}$$

$$f(\theta_\alpha^{Median}, 1, 0) - E[f(\theta, 0, 0)] \leq c(1, 0) - c(0, 0). \tag{9}$$

Using additive separability, we obtain

$$f(\theta_\alpha^{Median}, 1, 1) - E[f(\theta, 0, 1)] = h(\theta_\alpha^{Median}, 1) - E[h(\theta, 1)], \text{ and}$$

$$f(\theta_\alpha^{Median}, 1, 0) - E[f(\theta, 0, 0)] = h(\theta_\alpha^{Median}, 0) - E[h(\theta, 0)]$$

Suppose inequality (8) is satisfied. Then,

$$\begin{aligned} h(\theta_\alpha^{Median}, 0) - E[h(\theta, 0)] &\leq h(\theta_\alpha^{Median}, 1) - E[h(\theta, 1)] \\ &\leq c(1, 1) - c(0, 1) \\ &\leq c(1, 0) - c(0, 0), \end{aligned}$$

where the first inequality follows by Assumption 5, the second is due to (8) , and the third follows from the decreasing differences property of c. Hence, (9) is also satisfied. □

Continuum of Efforts

In this section, we consider the model with a continuum of effort levels: $e \in [e_0, e_1]$. Each type θ who exerts effort e obtains grade $g(\theta, e)$. We assume that the grade function is twice differentiable and strictly increasing so that both ability and effort increase grades. Moreover, we assume that g satisfies increasing differences $\left(\frac{\partial^2 g}{\partial\theta\partial e} > 0\right)$, which states that effort has a higher impact on grades for students with greater ability.

As before, the student's utility function is $U(w, e) = w - c(e)$, where c is strictly increasing. Because grades are strictly increasing in ability, there exists an inverse function $g^{-1}(\theta, \cdot)$ such that for all θ,

$$e = g^{-1}(\theta, g(\theta, e)).$$

Letting $C(\theta, g) = c(g^{-1}(\theta, g))$, we can write the utility function of type θ as

$$V(w, g; \theta) = w - C(\theta, g),$$

where $\frac{\partial C}{\partial \theta} < 0$, $\frac{\partial C}{\partial e} > 0$, and $\frac{\partial^2 C}{\partial\theta\partial e} < 0$.

Treating grades as the student's choice variable, the continuation game after grade disclosure has been selected becomes the standard model of Spence (1974) and Riley (1975). Under certain selection criteria, including the reactive equilibrium of Riley (1979), the divinity and universal divinity criteria of Bank and Sobel (1987), the D1 criterion of Cho and Kreps (1987), and the stability criterion of Kohlberg and Mertens (1987), the unique equilibrium is the most efficient separating PBE (Riley outcome).

Therefore, under any of those selection criteria, the unique equilibrium of the continuation game under grade disclosure features full separation. The continuation game under nondisclosure still features $e^* = e_0$, and $w^* = E[f(\theta, 0)]$. Hence, the results from this paper generalize to the model with a continuum of efforts with the appropriate substitution of the payoffs from the continuation games under disclosure and non-disclosure.[26]

References

Araujo, A., D. Gottlieb and H. Moreira (2007). A Model of Mixed Signals with Applications to Countersignalling. *RAND Journal of Economics,* 38:1020-1043.

Banks, J. S. and J. S. Sobel (1987). Equilibrium Selection in Signaling Games. *Econometrica,* 55:647-661.

Cho, I-K. & Kreps, D. M. (1987) Signaling games and stable equilibria. *Quarterly Journal of Economics,* 102:179-221

Ciolli, Anthony. (2007). "Grade Non-Disclosure Policies: An Analysis of Restrictions on M.B.A. Student Speech to Employers," *U. Pa. Journal of Labor and Employment Law*, 9 (3): 709-734.

Jain, Anjani (1997). "Rethinking Grade Nondisclosure," Mimeo., University of Pennsylvania.

Jain, Anjani (2005). "Proposed Expansion of Academic Honors: Context and Rationale," *Wharton Journal*.

Kohlberg, E. and Mertens, J.-F. (1986). On the Strategic Stability of Equilibria. *Econometrica,* 54:1003-1037.

Kreps, David (2005) "On Grade Non-Disclosure at the Stanford Graduate School of Business," *Mimeo*, Stanford University.

Riley, John G. (1975) "Competitive Signaling," *Journal of Economic Theory*,10:175- 186

Riley, John G. (1979) "Informational Equilibrium," *Econometrica*, 47: 331-359.

Spence, Michael (1973) "Job Market Signaling," *Quarterly Journal of Economics*, 87 (3): 355–374.

Spence, Michael (1974) "Market Signalling: Information Transfer in Hiring and Related Processes," Harvard University Press, Cambridge, MA.

Wise, C. C. and Hauser S. (2007) "The Business School Buzz Book 2007," Vault Editors.

[26]The equilibrium of the continuation game under grade disclosure features excessive effort (see, e.g. Riley, 1975): $\frac{\partial f}{\partial g}(\theta, g^*(\theta)) < \frac{\partial C}{\partial g}(\theta, g^*(\theta))$. Then, the welfare comparison between grade disclosure and grade non-disclosure weights the welfare cost of having the lowest effort (non-disclosure) against the cost of excessive effort (disclosure).