Grade Non-Disclosure

Daniel Gottlieb and Kent Smetters*

September 16, 2014

Abstract

This paper documents and explains why students vote for grade non-disclosure policies in Masters in Business Administration programs, why these policies are fully concentrated in highly-ranked programs, and why these policies are not prevalent in most other professional degree programs such as law, medicine and accounting. Our model accommodates various mechanisms — including honors, awards, and minimum grade requirements — that schools often introduce in order to encourage effort, and which could impact the median student's vote for grade non-disclosure.

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. As shown in Tables 1 and 2, non-disclosure policies are concentrated among highly-ranked U.S. MBA programs.[1] A majority of the most selective 15 U.S. MBA programs have a grade non-disclosure policy, while no school ranked 20 - 50 has such a policy. Moreover, these policies are distinctive in that they mainly exist in MBA programs and not in other professional programs including medicine, law, and accounting.[2]

The prestigious U.S. MBA programs with non-disclosure norms produce over 50,000 graduates per decade, many of whom will go on to lead critical functions within major firms in a range of sectors, including financial services, health care, manufacturing, technology and

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

[1] Outside of the United States, this pattern also appears, as top business schools like INSEAD also have a grade non-disclosure norm.

[2] 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.

global services. A graduate's deficiency in specific areas could easily go undetected for years, causing problems later after promotion. Grade non-disclosure policies could even contribute to a more general culture of hiring and promotion practices that fail to adequately reward training and merit, instead emphasizing other characteristics such as physical appearance and gender, which have been shown in past studies to be predictive indicators of compensation and promotions. Jacquart and Armstrong (2013) provide a review of the extensive literature analyzing the biases in the hiring, promotion and compensation of top executives.

By U.S. federal law, academic grades cannot be released by schools to third-parties, including potential employers, without student permission.[3] Conversely, schools in the United States cannot prevent students from disclosing their grades either. 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.

Support for grade non-disclosure has been met with mixed reactions from faculty across various schools.[4] Many business schools with a grade non-disclosure norm have nonetheless introduced methods of revealing the very best students, usually in the form of honors (the "dean's list") and prestigious awards like the Baker Scholar designation at Harvard and the Palmer Scholar designation at Wharton.[5] At Chicago's Booth School, MBA students also openly compete to earn teaching assistant positions, which are viewed as an indicator of excellent performance.[6] Concurrently, at many schools, including at Wharton, students whose grades fall below some minimum standards are dismissed. In other words, many business schools are still able to "manage the tails" of the performance distribution under the non-disclosure norm but not most of the mass falling in-between.

Many arguments have been made in the past in favor of grade non-disclosure, which we review in Section 2. These arguments include the desire to take more challenging courses, enhance cooperation among students, and to improve the ability of students to "network" with each other. To be sure, these arguments might have some merit in principle. However, the arguments are generally inconsistent with the available data. Some of the arguments also take endogenous choices as exogenous. Each of these arguments also fails to explain the two key facts of this market: (i) grade non-disclosure policies only exist in *top-ranked* MBA

[3]The law is the Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 CFR Part 99), sometimes called the Buckley Amendment.

[4]No MBA program with grade non-disclosure in the U.S. openly markets the potential benefit of the grade non-disclosure policy to potential applicants. As deputy dean of Stanford's Graduate School of Business, David Kreps (2005) wrote a thoughtful memorandum in 2005 to GSB students that discussed the advantages and disadvantages of grade non-disclosure during a period of internal school debate about the topic. Faculty at Wharton have consistently voiced their opposition non-disclosure policies (Jain 1997; 2005).

[5]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. So, for example, 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.

[6]Devin Pope (private communication).

programs and (ii) such policies are not present in *other* professional programs.

In this paper, we present a fairly standard signaling model with students, schools, and employers. Schools are heterogenous in their selectivity, thereby allowing for differences in quality. Under grade *disclosure*, employers can observe both a student's grades and the school's selectivity; under *non-disclosure*, an employer can only observe the school's selectivity. Students are heterogeneous in abilities, and they prefer larger post-school wages but dislike studying. If a majority of students vote to adopt a non-disclosure policy then students effectively pool the wages they receive upon graduation. We derive a pooling condition under which a median student voter prefers (a) a non-disclosure policy that allows her to perform *low* effort while in school and receive the expected (*mean*) wage upon graduation versus (b) a disclosure policy where she exhibits *high* effort while in school and receives the *median* wage upon graduation.

We show that this model produces the key stylized facts in this market. First, in Section 3 (and generalized in Section 4), we show that standard wage distributions, which are commonly used in the labor literature, in fact, imply that pooling (non-disclosure) condition is satisfied *at more selective schools*. These distributions include the log-normal (the most commonly used distribution in the labor literature), the Pareto family (often used in estimates of income inequality) and the Gamma families (consistent with newer evidence using confidential Social Security records). Importantly, our result does *not* require that the ability distribution of students admitted to more selective schools take on a different form than that of students admitted to less selective schools. Simply increasing the students' mean ability (the school's selectivity) is sufficient to produce pooling, reflecting the general nature of the pooling condition that we derive.

Second, in Section 5, we then introduce legal certification exams into the model. Certification exams are prevalent in non-MBA professional occupations, including medicine ("the boards") , law ("the bar") and accounting (the CPA). Passing these exams is legally required to fully practice in those occupations.[7] We show that the presence of these additional legal requirements undercuts the support for grade non-disclosure in these non-MBA programs under fairly general conditions. Intuitively, certification exams require a level of effort that is complementary to studying, reducing the value from pooling to reduce effort.

Finally, in Section 6 we introduce honors / awards and minimum grade requirements, thereby recognizing that schools with MBA programs still maintain the legal ability to introduce some mechanisms to encourage effort and potentially influence the median voter's adoption of grade non-disclosure. Not surprisingly, a minimum grade requirement encourages additional effort from lower ability students. However, rather than undermining support for grade non-disclosure by forcing more effort, this requirement actually *reinforces* the support for non-disclosure in equilibrium. The reason is that a minimum grade requirement raises the expected pooled wage with low effort, but it does not change the median wage under high effort. In contrast, giving honors / awards to top students reduces the support for non-disclosure by allowing the higher-ability students to separate themselves, thereby reducing the expected wage earned by the pool of students who have not obtained this additional signal.

[7] The historical legal basis for certification exams typically focused on consumer protection. To be sure, legal exams also exist in some specific areas of business, mainly security dealing. The corresponding coursework, however, tends to a smaller part of the MBA curriculum.

Our paper, therefore, contributes to a small but growing literature on school grades. Several recent papers examine conditions under which schools may reduce the informativeness of grade signals through "grade inflation" or other mechanisms (e.g. Yang and Yip 2003; Chan, Hao, and Suen 2007; Ostrovsky and Schwarz 2010). Dubey and Geanakoplos (2010) examine the impact of grades on student effort when students care about their relative ranking ("status"). In contrast, our model focuses on the *students*' collective action in reducing the informativeness of their signals.

2 Common Explanations of Grade Non-Disclosure

Before turning to our own theory of grade non-disclosure policies, we first examine several common explanations in support of these policies.

2.1 Taking More Challenging Courses

As the former president of the student association at Chicago Booth School of Business, April Park, argued, "grade non-disclosure allows students to take more challenging courses instead of taking classes in which they are over-qualified."[8] Similar statements are routinely made by other student associations in schools with grade non-disclosure policies. However, this argument is problematic in at least four ways.

First, empirically, self-reported amount of time spent on academics fell by close to a third during the first five years after grade non-disclosure was fully implemented at Wharton.[9] As one Wharton student put it: "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."[10] The same experience has been felt at other top programs as well. 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."[11] Akram Zaman, former co-president of the Harvard's Student Association's Executive Committee, notes 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."[12] At Stanford GSB, a student put it this way: "[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."[13]

[8]The Economist, ibid.

[9]Jain (2005) reports that time spent studying fell by 32% for first-semester Wharton MBA students, by 34% for students in their second semester, 16% during the third semester, and 8% during the fourth semester. Third- and fourth-semester students already started from a lower base.

[10]Wise and Hauser, ibid. "Campus Confidential." Business Week, September 12, 2005.

[11]The Economist, ibid.

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

[13]Wise and Hauser (2007).

But have hours studying maybe simply decreased over time across all education disciplines? Apparently not. During the same period in which non-disclosure norms were being adopted in many top MBA programs, total hours spent studying by both college freshman and seniors remained flat (National Survey of Student Engagement 2007, Figure 2). Conversations with fellow professors at various law schools and medical schools, while informal, also indicate that effort exerted in those programs have not changed much either. One academic dean at a top law school, speaking to us confidentially, noted, that "our students work just as hard today as they did ten years ago. We have not seen any change in how hard they work." In contrast, professors at MBA programs, who taught before and after grade non-disclosure policies were implemented, often do not share the same opinion. In sum, the fall in study hours by MBA students is the opposite outcome of what would be expected if students were voting for grade non-disclosure in order to take accept a greater challenge. In contrast, hours spent studying did not fall in other education programs at the same time.

Second, as it turns out, the administrative records indicates that the introduction of grade nondisclosure at Wharton did not actually affect the pattern of courses waived by students.[14] If grade non-disclosure really existed to take on more challenging courses, we would have expected students wanting a more challenging course load to waive out of more basic courses.

Third, there is no reason why the benefit from taking more difficult classes would be restricted to *elite* MBA programs. Why wouldn't students from non-elite MBA programs also want to take on more challenging courses? If it is challenging to offer an explanation that does not depend on a tradeoff between signaling and effort itself.

Fourth, it is not clear why the value from taking more difficult classes would be largely restricted to *MBA* programs at all. Why don't students enrolled in law schools, accounting programs and medical schools also see the value in reducing the risk of reporting poor grades so that they can take on more challenging courses? To be sure, the range of elective coursework is often more narrow within a given field of concentration within these non-MBA programs relative to MBA programs. But, these other professional degree programs still offer considerable choice, especially in the form of selecting more challenging concentrations. In any case, the same basic argument for taking more challenging courses should still apply even if the option set is smaller.

2.2 Enhancing Student Cooperation

The Student Handbook of Stanford GSB mentions that grade non-disclosure also improves student cooperation. But, this hypothesis is also problematic for several reasons. First, as Kreps (2005) points out: "Helping 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?" Second, even if there are some gains from cooperation, the time-series evidence noted above suggests that this cooperation is coming at the expense of a substantial reduction in studying. While it is plausible that cooperation could lead to some efficiency gains and less study time for some students, increased cooperation should plausibly lead to more average study time.

[14] Jain (1997).

Third, as before, this theory still does not explain why only students in *elite* MBA programs want to cooperate. Fourth, why don't students who are training to become medical doctors, lawyers, and accountants, also want to cooperate?

2.3 Building Networks

Closely related, students sometimes say that part of the value of getting an MBA is the ability to network with each other, which is easier to do without competition. This rationale, however, is subject to the same basic critiques as the last argument. The odds that a few friends can materially impact one's future is small. It is also again unclear why only students in *elite* MBA programs value networking — one could even imagine that they need it less — or why networking does not have value in the study of law.

This argument might also be confusing "cause and effect" since the desire to network is endogenous and it depends on the availability of other signals. Networking becomes relatively more important after the primary signal of quality (grades) has been removed. Indeed, ex ante, it is not even clear that the value of networking itself is improved by adopting grade non-disclosure. It apparently all depends on what is really meant by "networking." On one hand, it could mean socializing with new friends, consistent with the time series data showing a declining in study effort after grade non-disclosure policies are adopted. To be sure, there might be value to such relaxation, especially since many students in MBA top programs previously worked long hours on Wall Street, at consulting firms, or at a startup firm. On the other hand, if the purpose of networking is to assess the value of potential future business partners, then grade non-disclosure likely reduces the value of networking by reducing the group dynamics around studying for those skills.[15]

3 Basic Model

This section derives the conditions that support grade non-disclosure under a non-overlapping property of the relationship between effort and grades, explained below. This property allows us to give the most straightforward interpretation of our results. Section 4 then generalizes the results to the case where the non-overlapping property is relaxed.

3.1 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 $e \in \{0, 1\}$. Schools are characterized by a selectivity parameter $\alpha \in \mathbb{R}$,

[15] Of course, one might argue that the education offered by MBA programs is of little future business value altogether. But this argument would then raise the question of why students spend so much money (including lost wages) on paying for elite MBA degrees when other mechansims for networking exist and are much more affordable. Moreover, if grades really "don't matter" in MBA programs for future productivity, one wonders why students in elite programs with grade non-disclosure policies appear to be so passionate about passing these policies, especially when employers ask for grades at schools without these policies.

which affects the distribution of accepted students.[16] 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 $\mu_{\alpha'}$. 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.

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 $c(1) := c > 0$, $c(0) = 0$. Without loss in generality, we have arbitrarily normalized the minimum effort level to zero to denote the amount of effort a student would choose in the absence of explicit incentives. Of course, we are not claiming that students would literally not study at all in the absence of explicit incentives; all of our results would maintain under different normalizations.

Let $g(\theta, e)$ equal the grade received by student with ability θ who provides effort e.

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

Remark 1. The 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 weak inequality 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 4, 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.[17] 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, which is the policy at several top programs.[18] However, our results would also be robust to assuming that the policy of grade non-disclosure must be approved by a super-majority, as is the case at Columbia. Schools

[16] 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.

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

[18] 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 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.

cannot legally decide the voting procedures since the grades are owned by students (Section 1). In Section 6, we consider mechanisms that schools might use to influence the median voter.

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, the grades of students. The competitive market offers a wage w equal to the expected productivity of the student.

A few related remarks are in order related to time $t = 2$. First, modeling effort as a binary choice simplifies the model without much loss in generality for our particular results. In the Appendix, we generalize our key results to a continuum of efforts.

Second, notice that we also assume that students know their own productivity θ. Our results would maintain, however, if students had non-degenerate beliefs about their own productivity. Since the mean wage for each school is observable (indeed, published each year in major news outlets), we only require that students have a general sense of how their private productivity compares to the mean wage.

Third, we model the grading technology $g(\theta, e)$ as deterministic since agents are risk neutral. We can also take $g(\theta, e)$ as exogenous at a given school since our results will still maintain in a continuation game where our model is nested inside of a competitive school model. In other words, once admitted, the median student in our model will vote according to the conditions that we derive herein.

3.2 Equilibrium

We study the Perfect Bayesian Equilibrium in pure strategies of the game. Denote G as the grade disclosed by a student if the grade non-disclosure policy is rejected.

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:

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

2. Employers earn zero profits given beliefs:

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

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.[19]

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) For any attainable grade $G \in g([\theta_0, \theta_1], \{0, 1\})$, $\beta_D(\cdot \mid G)$ 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 after the disclosure policy is determined to satisfy Bayes' rule. Condition (4c) states that beliefs conditional on a grade cannot attach a positive mass to types that are unable to reach this grade.

We study the equilibrium of the game backwards.

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:

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

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

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

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

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

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 if $\kappa \leq c$. 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.

3.3 Majority Rule and the Role of Selectivity

Turning now to the majority rule outcome, for a given level of school selectivity α, the median voter balances 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 more productive students.

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 (see 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 $\theta_\alpha^{Median} < E_\alpha[\theta]$. Then, there exists $\bar{\kappa}_\alpha > c$ such that, in any PBE, non-disclosure is chosen if $\kappa \leq \bar{\kappa}_\alpha$ and disclosure is chosen if $\kappa \geq \bar{\kappa}_\alpha$.*

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. Then:

Corollary 2. *Suppose $G(\alpha)$ is increasing. Then, 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. Greater selectivity raises the quality of students by attracting a disproportionately larger amount of very good students. We will refer to the condition "$G(\alpha)$ is increasing" as the "pooling condition" since it implies that the conditions in Corollaries 1 and 2 are both met.

3.4 Distributions that Produce an Increasing $G(\alpha)$

In realistic wage distributions, means exceed medians. In the United States, the mean wage is about 32% larger than the median wage, as averaged across the more than 800 occupations tracked by U.S. Bureau of Labor Statistics (BLS 2013). For the more than 100 business related occupations tracked by the BLS, the mean wage exceeded the mean wage in every one. To directly test whether $G(\alpha)$ is increasing for *MBA graduates*, however, it would be ideal to observe the distribution of initial total compensation received by students before and after grade non-disclosure policies are adopted. But getting that data proved to be very challenging.[20]

However, the pooling condition is satisfied by the most common parametric distributions that are used to study income dynamics in the labor economics literature: lognormal, Pareto, and Gamma. Most studies, including those estimating the education wage premium, assume that wages are lognormally distributed (e.g., Grogger and Eide 1995; Gouskova 2014). This assumption is also consistent with plausible formulations of heterogeneity in abilities and productivity gains from education (Card 2001). Pareto distributions, however, are routinely used in the study of income inequality because of their ability to capture heavier tails. More recently, using confidential employer-reported wage data from the U.S. Social Security Administration, Guvenen et al (2014) estimate heavy tailed wage distributions that appear to better described by the Gamma family. We now consider each distribution with a series of examples.

[20] Alternatively, one could try to compare the distribution of wages of closely-ranked schools that differ in their grade non-disclosure policies. However, most schools not willing to share that level of detailed information, and schools typically only observe starting base salaries and not contingent incentives that are common in many industries. Moreover, schools that differ in their grade non-disclosure policies also differ in the types of industries in which many of their students enter upon graduation. For example, Wharton (with grade non-disclosure) has historically sent more students into finance relative to Northwestern (with grade disclosure).

Example 1. [Log Normal] 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 so 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

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

Remark 2. Notice that the pooling condition does *not* require that the distribution of θ have a different variance, skewness or kurtosis at different values of the school's selectivity,α. The variance (equal to σ^2), the skewness (equal to $\left(e^{\sigma^2} + 2\right)\sqrt{e^{\sigma^2} - 1}$), and kurtosis (equal to $e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$) are all independent of α. In other words, the wages facing the graduates of all schools, including those with and without disclosure, can be modeled with the same distribution where more selective schools attract higher ability students on average. An increasing average alone produces an increasing $G(\alpha)$.

Example 2. [Pareto] Suppose θ is distributed according to a Pareto (α, k) distribution, where $\alpha \geq 0$ is the scale parameter and $k > 1$ is the Pareto index. The mean and median of θ are $\frac{k}{k-1}\alpha$ and $2^{\frac{1}{k}}\alpha$.[21] The mean-median gap,

$$G(\alpha; k) = \left(\frac{k}{k-1} - 2^{\frac{1}{k}}\right)\alpha,$$

is positive and increasing in α. Therefore, the conditions from Corollaries 1 and 2 are also satisfied for Pareto distributions. Indeed, the equilibrium disclosure policy can be calculated explicitly. Grade non-disclosure is always chosen if studying effort is inefficient, $\kappa \leq c$. If effort is efficient, $\kappa > c$, non-disclosure is chosen if

$$\alpha \geq \frac{\kappa - c}{\frac{k}{k-1} - 2^{\frac{1}{k}}}.$$

Remark 3. Similar to Remark 2, notice that all schools across all levels of selectivity can share the same distribution k-Pareto distribution for the pooling condition to hold. Selective schools are only differentiated by having a large value of α.

Example 3. [Gamma] Because the median of the Gamma distribution does not have a closed form solution, we verified the conditions numerically across a wide range of parameter values. Details are available from the authors.

[21] Under a Pareto (α, k) distribution, the proportion of a population whose income exceeds $\theta \in [\alpha, +\infty)$ is $\left(\frac{\alpha}{\theta}\right)^k$. The scale parameter α shifts the whole distribution to the right and, therefore, orders distributions in terms of first-order stochastic dominance. The Pareto index k, which has to exceed one for the first moment to exist, measures the inequality of the distribution. A smaller index is associated with a larger proportion of individuals having higher incomes. While the Pareto index does not induce a ranking in terms of first-order stochastic dominance, the mean-median gap is also increasing in k. Therefore, the results from Corollaries 1 and 2 would also hold if we assumed that more selective schools chose from Pareto distributions with larger Pareto indices k.

3.5 Summary

In sum, our model predicts that selective schools would adopt a grade non-disclosure policy whereas less selective schools would adopt a policy of disclosure under a wide range of standard wage distributions. This pattern is consistent with the evidence from MBA programs (Tables 1 and 2). Our model also predicts that studying effort is (weakly) lower with grade non-disclosure, which is also consistent with the evidence reported earlier from Jain (2005).

4 Overlapping Grades

Thus far, we have assumed that effort had such a strong effect on grades that even 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:[22]

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

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.

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 if the type space is unbounded ($\theta_1 = +\infty$). 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: $(\theta + \gamma) - \theta > c$.

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

4.1 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 3 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. Moreover, $e_D(\theta) = 1$ for all θ. Furthermore, this PBE survives the intuitive criterion.*[23]

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.

It remains to be shown that the equilibrium 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 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 satisfies the intuitive criterion. □

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

[23] See Cho and Kreps (1987) for a presentation of the intuitive criterion. Because we focus on PBE in pure strategies, the existence of an equilibrium that survives the intuitive criterion is not immediate.

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

4.2 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.[24] 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*

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

Moreover, this PBE survives the intuitive criterion.

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

Intuitively, when grades are not sufficiently responsive to effort, a type-θ student loses 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 the equilibrium has non-decreasing effort, 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 PBE with non-decreasing effort,* $e(\theta) = 0$ *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 the equilibrium has non-decreasing effort 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 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. □

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

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. A type in the third interval who deviates to $e = 0$ obtains grades in the interval $(\theta_1 - \gamma, \theta_1]$, which is also off the equilibrium path. Let off-the-equilibrium-path beliefs $\mu(\theta|G)$ assign a unit mass to 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 productivity 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:

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

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

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

$$c \geq \gamma \implies G + \kappa - c \leq G - \gamma + \kappa, \text{ and}$$

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

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 contained at the set of undominated types $\{G - \gamma, G\}$. Substituting the student's equilibrium payoff and the market's best response, the first condition becomes

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

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

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

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

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 $\theta_\alpha^{Median} + \kappa - c \geq E_\alpha[\theta]$ 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 non-decreasing 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.[25] For concreteness, in the rest of the paper, we will select the non-increasing PBE when considering the model with sufficiently unresponsive grades. Nevertheless, all of our results can be generalized for other PBE. If the PBE is fully separating, the results remain exactly as stated. If there is pooling, the relevant conditions are in terms of the students' equilibrium wages, rather than their skills.

5 Certification

Business schools are unique in that most other professional programs — including medicine, law and accounting — allow for grade disclosure, even at the same prestigious universities with MBA programs that have adopted the grade non-disclosure norm.[26] These other professional programs, however, also have some uniform certification (medical licensing examination, legal bar, CPA) that is required to practice at the fullest level. Despite the presence of a uniform certification, grades still play an important role on the students' job market outcomes.[27] 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

[25] 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 $\phi_\alpha^{Median} \leq E_\alpha[\theta]$ and disclosure is chosen if $\phi_\alpha^{Median} \geq E_\alpha[\theta]$.

[26] Limited exceptions were noted earlier.

[27] Sander and Yakowitz (2010), for example, argue that law school grades are "the most important predictor of career success" and "decisively more important than law school eliteness." The survey from the 2008 National Residency Matching Program reports that grades from medical school are consistently among the top-5 most important criteria for selecting candidates into residency programs. The other important criteria were the Medical Licensing Exam score, the School Performance evaluation (which is sometimes referred to as the "Dean's letter," and reviews the student's academic performance, including grades in all coursework), the personal statement, and the letter of recommendation.

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.

To be sure, non-MBA programs at prestigious universities presumably have different values of κ and η, as well as a different cost function $c(e, s)$, relative to MBA programs at these same universities. These parameters, along with the wage distributions of non-MBA professional schools under the counterfactual economy where external exams do *not* exist, are not empirically observable. Hence, it is certainly possible that students in non-MBA programs might have voted to reject grade non-disclosure even without the presence of external exams. Nonetheless, our analysis in this section shows that if students had voted in favor of the grade non-disclosure norm without external exams, the presence of external exams decreases the support for grade non-disclosure.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 3 and 4, 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 comparative static result shows 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 3 or Section 4 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. With certification, a type-θ student obtains expected payoff

$$\theta + \eta + \max\{\kappa - c(1,1);\ -c(0,1)\}$$

if there is grade disclosure and

$$E_\alpha[\theta] + \eta - c(0,1)$$

if there is grade non-disclosure. The condition for grade non-disclosure to win is then

$$\theta_\alpha^{\text{Median}} - E_\alpha[\theta] \leq \min\{c(1,1) - c(0,1) - \kappa;\ 0\}. \tag{4}$$

Without certification, the condition for grade non-disclosure to win is

$$\theta_{\alpha}^{\text{Median}} - E_{\alpha}[\theta] \leq \min\{c(1,0) - c(0,0) - \kappa;\ 0\}. \tag{5}$$

By decreasing differences, $c(1,1) - c(0,1) \leq c(1,0) - c(0,0)$. Hence, whenever inequality (4) holds, (5) must also hold. □

6 Minimum Grade Requirements and Awards

In the presence of grade non-disclosure, schools still have some limited tools available to encourage effort and even influence the vote whether students adopt grade non-disclosure. The two most common instruments are awards and honors as well as some minimum performance requirement.[28] 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.

6.1 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 $\bar{w} < E_{\alpha}[\theta] - c$ 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 the model with non-overlapping grades (Section 3), 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 a 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

[28]More drastic tools to encourage the median voter to reject non-disclosure include reducing the grade granularity by, for example, giving the highest grade A to 51% or more of its students.

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 ($E_\alpha[\theta] \geq \theta_\alpha^{Median} + \kappa - c$), this 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 4, 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 welfare as measured by the utilitarian criterion.[29] Therefore, any mechanism that determines the minimum grade policy by a utilitarian 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 Sections 3 or 4. 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 utilitarian 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 3 or Section 4. Since μ_α 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

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

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

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

[29] The utilitarian welfare criterion maximizes the sum of all payoffs. It corresponds to the expected utility computed at time "$t = 0$," before students know their ability θ.

which establishes that the utilitarian welfare is higher under the minimum grade policy. $\square$

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.

6.2 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:[30]

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

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. Any mixed strategy equilibrium, a positive mass of types chooses $e = 1$ with strictly positive probability.[31] If this were not the case, by the previous argument, all types slightly below the top ϕ of the type distribution

[30] If no type can be sure to receive the prize with low effort, i.e. $\int_{g(\theta,1) \geq g(\theta_1,0)} d\mu_\alpha\left(\theta\right) < \phi$, let $\theta_\alpha^{**} = \theta_1$.

[31] It is straightfoward, but not very insightful, to characterize the equilibria in mixed strategies. Any such equilibrium partitions the type space into three (possibly empty) intervals. In the lowest interval, all types choose low effort and never win the prize. In the highest interval, all types choose high effort and always win the prize. In the intermediate interval, types play strictly mixed strategies, win the prize with positive probability, and are indifferent between exerting high and low efforts.

would benefit from playing $e = 1$. Since a positive mass of students exerts high effort, the utilitarian 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 3 or 4. Suppose it is efficient to exert high effort $\kappa > c$ and equilibria have grade non-disclosure $\theta_\alpha^{med} + \kappa - c < E_\alpha[\theta]$. In any equilibrium, introducing an award strictly increases utilitarian welfare.*

Proof. Existence of a mixed strategy equilibrium for the continuation game follows from Dasgupta and Maskin (1986). 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 utilitarian welfare is then $E_\alpha[\theta] + \lambda (\kappa - c)$, which is greater than the utilitarian welfare without the award $E_\alpha[\theta]$. $\square$

Whenever the prize is given to less than half of the 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 (since he is now pooled with a worse pool of students). 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 even without effort $(\bar{\theta_\alpha}, \theta_1]$. 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 the support for grade non-disclosure.

7 Conclusion

Grade non-disclosure is prevalent among elite MBA programs but not commonly found in lower ranked MBA programs or in other professional degree programs with external certification. Common explanations by proponents of this norm are typically inconsistent with the data and also fail to explain why students in non-elite MBA programs do not also adopt this norm. Our model explains why students in elite MBA 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 increase 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 is both legal and is not generally practiced, as in other professional occupations. Future work could explore this issue in more detail.

Appendix

General Production Functions

This section considers the model of Section 3 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 3, the payoff of type θ under grade disclosure is

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

whereas the payoff under grade non-disclosure is $E\left[f\left(\theta, 0\right)\right]$. 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 $\phi(\theta_\alpha^{Median}) \geq E_\alpha\left[f\left(\theta, 0\right)\right]$ and non-disclosure is chosen if $\phi(\theta_\alpha^{Median}) \leq E_\alpha\left[f\left(\theta, 0\right)\right]$.*

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

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 $\tilde{G}\left(\alpha\right) \equiv E_\alpha[f\left(\theta, 0\right)] - \phi\left(\theta_\alpha^{Median}\right)$ 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 $\tilde{G}(\alpha) \geq 0$, 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. □

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

$$f(\theta_\alpha^{Median}, 1, 1) - E[f(\theta, 0, 1)] \leq c(1, 1) - c(0, 1), \text{ and} \tag{6}$$

$$f(\theta_\alpha^{Median}, 1, 0) - E[f(\theta, 0, 0)] \leq c(1, 0) - c(0, 0). \tag{7}$$

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 (6) 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 (6) , and the third follows from the decreasing differences property of c. Hence, (7) 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,

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

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.[32]

[32] The equilibrium of the continuation game under grade disclosure features excessive effort (see, e.g. Riley, 1975): $\frac{\partial f}{\partial g}\left(\theta, g^*(\theta)\right) < \frac{\partial C}{\partial g}\left(\theta, g^*(\theta)\right)$. 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).

References

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

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

[3] Bureau of Labor Statistics, 2013. "Occupational Employment Statistics: May 2013 National Occupational Employment and Wages Estimates – United States." Available at: http://www.bls.gov/oes/current/oes_nat.htm.

[4] Card, David. "Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems," Econometrica 69 (2001), 1127–60.

[5] Chan, W., Hao, L. and Suen, W. (2007). A Signaling Theory of Grade Inflation. *International Economic Review*, 48, 1065-1090.

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

[7] 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.

[8] Dasgupta, P. and Maskin, E. (1986). The Existence of Equilibrium in Discontinuous Economic Games, I: Theory. *Review of Economic Studies*, 53(1): 1-26.

[9] Dubey, P. and J. Geanakoplos. (2010). Grading Exams: 100, 99,...,1 or A, B, C? Incentives in Games of Status. *Games and Economic Behavior*, 69(1): 72-94.

[10] The Economist, 2005. "News from the schools." December 16.

[11] INSEAD MBA Career Services, 2013. "Grade Non Disclosure Policy," insidemba.insead.edu/careers/whatsnew/GradeNonDisclosure.pdf

[12] Gouskova, Elena (2014). "Parameter estimates comparison of earnings functions in the PSID and CPS data, 1976–2007." Economics Letters 122(2): 353-357

[13] Grogger, J., & Eide, E. (1995). Changes in college skills and the rise in the college wage premium. Journal of Human Resources, 30(2), 280-310.

[14] Guvenen, Fatih, Fatih Karahan, Serdar Ozkan and Jae Song, 2014. "What Do Data on Millions of U.S. Workers Say about Life Cycle Income Risk?" University of Michigan Retirement Research Center, Working Paper 2013-302.

[15] Jain, Anjani (1997). “Rethinking Grade Nondisclosure,” An address to faculty and students, University of Pennsylvania.

[16] Jain, Anjani (2005). “Proposed Expansion of Academic Honors: Context and Rationale,” *Wharton Journal*.

[17] Jacquart, P and J.S. Armstrong, 2013. “The Ombudsman: Are Top Executives Paid Enough? An Evidence-Based Review.” Interfaces, 43(6): 580 - 589.

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

[19] Kreps, David (2005) “On Grade Non-Disclosure at the Stanford Graduate School of Business,” *Mimeo*, Stanford University.

[20] National Resident Matching Program, Data Release and Research Committee (2009). Results of the 2008 NRMP Program Director Survey. *www.nrmp.org/data/programresultsbyspecialty.pdf*.

[21] National Survey of Student Enagement (2007), Annual Report 2007. School of Education, Indiana University, Bloomington, IN.

[22] Ostrovsky, Michael and Schwarz, Michael (2010). Information Disclosure and Unraveling in Matching Markets. *American Economic Journal: Microeconomics*, 2(2): 34-63.

[23] Riley, John G. (1975) “Competitive Signaling,” *Journal of Economic Theory*,10:175- 186.

[24] Riley, John G. (1979) “Informational Equilibrium,” *Econometrica*, 47: 331-359.

[25] Sander, Richard and Yakowitz, Jane. (2010) “The Secret of My Success: How Status, Prestige and School Performance Shape Legal Careers,” *Mimeo*, UCLA.

[26] Spence, Michael (1973) “Job Market Signaling,” *Quarterly Journal of Economics*, 87 (3): 355–374.

[27] Spence, Michael (1974) “Market Signalling: Information Transfer in Hiring and Related Processes,” Harvard University Press, Cambridge, MA.

[28] Yang, H. and Yip, C. S. (2003) “An Economic Theory of Grade Inflation,” Working Paper, Ohio State University.

[29] Wise, C. C. and Hauser S. (2007) “The Business School Buzz Book 2007,” Vault Editors.

Table 1: Top 25 MBA Programs, Ordered by GMAT

School	Non-Disclosure Policy	GMAT	US News Rank
Stanford GSB	Yes	728	1
Harvard HBS	Mixed(1)	724	2
Yale	Yes	722	11
Penn (Wharton)	Yes	718	4
MIT (Sloan)	No	718	3
UC Berkeley (Haas)	Partial(2)	718	8
Dartmouth (Tuck)	No	716	7
U Chicago (Booth)	Yes	715	6
NYU (Stern)	Yes	715	10
Northwestern (Kellogg)	No	714	5
Columbia	Yes	712	9
UCLA (Anderson)	No	710	14
Michigan (Ross)	Yes	704	15
UVA (Darden)	No	699	13
Duke (Fuqua)	No	697	12
Wash U St. Louis (Olin)	No	695	20
Carnegie Mellon (Tepper)	Partial(2)	694	18
Minnesota (Carlson)	No	694	21
U of Florida (Hough)	No	694	47
UC Davis	No	692	29
USC (Marshall)	No	690	22
Cornell (Johnson)	No	687	16
UNC (Kenan-Flagler)	No	686	19
Notre Dame (Mendoza)	No	685	38
UT Austin (McCombs)	No	684	17

Source: US News and World Report (2011) and authors' calculations.

Notes: (1) Harvard traditionally had grade non-disclosure with the strong support of 87% of the class (Harvard Business School Alumni Bulletin, 2006). Harvard's current 1-2-3 point system effectively maintains non-disclosure by pooling 75% of students into grade 2, making it similar to a non-disclosure school with honors and a minimum grade requirement. (2) Grade disclosure not allowed until the second interview.

Table 2: Next 25 MBA Programs, Ordered by GMAT

School	Non-Disclosure Policy	GMAT	US News Rank
Georgetown (McDonough)	No	684	25
Boston U	No	681	35
U of Washington (Foster)	No	681	39
Emory U (Goizueta)	No	680	23
Georgia Institute of Tech	No	678	28
Rochester (Simon)	No	677	46
Ohio State (Fisher)	No	676	26
U Wisconsin - Madison	No	675	30
Brigham Young U (Marriott)	No	675	32
Tulane (Freeman)	No	674	41
Vanderbilt (Owen)	No	673	31
UC Irvine (Merage)	No	673	42
Arizona State (Carey)	No	672	27
Rice U (Jones)	No	672	36
U of Maryland (Smith)	No	670	45
UT Dallas	No	668	44
Indiana U (Kelley)	No	664	24
Boston College (Carroll)	No	662	34
Iowa (Tippie)	No	657	43
Purdue (Krannert)	No	654	50
Wake Forest (Babcock)	No	653	48
Penn State (Smeal)	No	650	40
Texas A&M (Mays)	No	646	33
Illinois Urbana-Champaign	No	641	37
Michigan State (Broad)	No	636	49

Source: US News and World Report (2011) and authors' calculations.