Citation for published version:
Zubrickas, R 2015, 'Optimal grading', *International Economic Review*, vol. 56, no. 3, pp. 751-776.
https://doi.org/10.1111/iere.12121

DOI:
10.1111/iere.12121

Publication date:
2015

Document Version
Peer reviewed version

Link to publication

Optimal Grading*

Robertas Zubrickas[1]

February 25, 2014

Abstract

The teacher-student relationship is modeled as an agency problem, where teachers are concerned with human capital formation and students with ability signaling. We distinguish between two cases depending on whether in ability inference the job market can or cannot observe the grading rule applied. We show that many empirical grading patterns, including grade compression and inflation, are all consistent with optimal ability screening when grading rules are unobservable. With observable grading rules, the teacher perfectly screens students' abilities provided certain conditions hold. We apply the model to discuss policy applications such as "No Child Left Behind."

*Manuscript submitted January 2012; revised January 2013, August 2013.

[1]This work is a revised version of the first chapter of my thesis written at the Stockholm School of Economics. I would like to thank the editor and two anonymous referees for their valuable comments and suggestions. I also thank Tore Ellingsen, Christian Ewerhart, Drew Fudenberg, Magnus Johannesson, Karl Wärneryd, and seminar participants at various institutions. Financial support from the Jan Wallander and Tom Hedelius Foundation and Forschungskredit of the University of Zurich is gratefully acknowledged. Please address correspondence to: Robertas Zubrickas, University of Zurich, Depart-

1 Introduction

In the economics literature, there is a dichotomy of views on the role of education. The traditional view of human capital formation (Becker 1993) is challenged by the idea that education serves the purpose of ability signaling (Spence 1973; Riley 1979; Bedard 2001). In this paper, we take the standpoint that teachers are concerned with human capital, whereas students with ability signaling. In particular, students treat knowledge acquired as abstract, ephemeral, and, therefore, costly compared to the opportunities foregone. Ability signals come in the form of grades, about which, in contrast to students, teachers are impartial *per se*. The differing attitudes toward education create a conflict of interests between teachers and students. This paper is an attempt to model the resulting agency problem that, as we argue, can have a distinctive feature. With asymmetry in information about the grading rules applied, grades can become costless incentives. We show that the theoretical predictions of our model with costless grades closely match many empirical grading patterns.

The teacher-student relationship is modeled as a monopolist screening problem akin to Mussa and Rosen (1978). A teacher teaches a class of students with disparate abilities for the subject taught. A student's ability is his private information, and the higher the ability, the lower the learning effort he needs to exert to attain a certain knowledge level. The assumption is that the teacher has a technology – the exam test – that allows her to assess accurately the knowledge levels attained by her students. The teacher sets up a grading rule that assigns grades to exam scores, i.e., to knowledge levels. We solve for the optimal grading rule that maximizes the expected knowledge level of a grade-minded

ment of Economics, Schönberggasse 1, CH–8001 Zurich, Switzerland. Phone: +41 44 634 35 88, email: robertas.zubrickas@econ.uzh.ch.

student, where the utility value of a grade comes from its ability-signaling properties.

In determining the optimal grading rule, we distinguish between two cases depending on whether the expected distribution of grades has an effect on their ability-signaling values or not. The distribution matters when the job market observes the grading rule applied and, thus, can make a correct inference about a student's expected ability. In this case, we show that the optimal grading rule perfectly screens abilities provided certain conditions hold. But in the other case, related to the situation when the job market does not observe the grading rule applied and makes inference using a commonly perceived grading standard, perfect screening is no longer expected knowledge maximizing. With unobservable grading rules, the signaling values of grades are fixed, which, consequently, renders grades as costless incentives. We show that with costless grades the teacher rather distorts the effort incentives of more able students than those of less able ones, yielding coarse grading.

The predictions of the model with costless grades closely match many empirical grading patterns. First, the phenomenon of grade compression can be explained by the equilibrium pooling that arises in our model. Second, in the dynamic version of the model with many teachers and the job market adaptively adjusting its perception of grading standards, we obtain the gradual deterioration of grading rules resulting in grade inflation. Third, the dynamic model also explains the reduction in time spent for study by full-time students, reported by Babcock and Marks (2011), in concurrence with increasing average grades. Fourth, the model provides a rationale for the discretization of grading scales as we obtain that in the long run no type is screened in equilibrium. Finally, we present a comparative statics result that the more pessimistic a teacher is about her students' abilities, the more lenient she is in grading. This theoretical prediction finds a strong empirical support from the literature on educational and psychological measurement that shows that professors apply more stringent grading criteria in fields with more able students, and *vice versa* (see Goldman and Widawski 1976; Strenta and Elliott 1987; Johnson 2003). The latter evidence suggests that grading behavior is affected by class composition, supporting our approach of ability screening.

The approach of this paper toward the teacher-student relationship and the suggested explanation of the aforementioned empirical grading patterns are novel in the economics literature. Unlike in this paper, many existing explanations tend to rely on arguments that go beyond the classroom. For instance, McKenzie and Tullock (1981, Ch. 17) see the expanding availability of education as a cause of grade inflation. They argue that due to increasing competition universities engage in lowering grading standards in order to attract more students, thus, creating conditions for grade inflation. The proponents of this explanation may possibly overlook other channels for grade inflation that can originate from the expanding availability of education. For instance, with more students enrolled the distribution of abilities can become more skewed toward the lower end of the ability space. According to our comparative statics result, teachers respond to this change by applying more lenient grading rules in order to extract more effort from more numerous low-ability students, thus, producing grade inflation. Without accounting for changes in the composition of students, empirical studies can overestimate the effects of competition among universities on grade inflation. The other related literature is discussed in the next subsection.

Generally, the agency model studied here relates to a broader class of three-tier agency problems with hidden information, where the interim tier's objective is the alleviation of informational asymmetries between the top and bottom tiers (see Lizzeri 1999). When the preferences of the interim and top tiers are misaligned with respect to the actions of the bottom tier, the agency relationship between the interim and bottom tiers can feature asymmetric transfer values. This asymmetry in transfer values can have important implications for the form of incentive schemes at the bottom tier. Besides the example of the teacher-student relationship, another example is job performance appraisal. As later discussed in the text, the compression of appraisal ratings frequently observed in practice (see Murphy and Cleveland 1995) can arise for the reasons similar to those behind grade compression as postulated here, namely, the costlessness of rewards. A line manager may find it rewarding to inflate appraisal ratings when he is not held accountable for the resultant payroll cost.

The remainder of the paper is organized as follows. After a literature review, we present a modeling framework in Section 2. The case of unobservable grading rules is analyzed in Sections 3 and 4. In Section 3, we solve for the optimal grading rule in the static model and in Section 4 we analyze the dynamics of grading rules under adaptive expectations. Section 5 deals with the case of observable grading rules. In Section 6, we relate our main findings to the grading patterns observed in practice and discuss policy applications. The last section concludes the study. All omitted proofs are in the appendix.

Related literature

An agency approach toward the teacher-student relationship is also taken in Dubey and Geanakoplos (2010). They deal with the question of discretization of grades: what grading scale, finer or coarser, a teacher should use in order to elicit maximal learning effort from students. The teacher-student relationship is modeled as a game of status with stochastic output like in a tournament. In their model, a student's utility of a grade depends on the class ranking, status, resulting from the grade obtained. They find that teachers should use coarse grading schemes and "pyramid" the allocation of grades: the highest grade would be available to fewer students than the second-highest grade, and so on. While their model provides a rationale for the discretization of grades, it is not consistent, however, with the compression of highest grades and grade inflation. Thus, besides status incentives, there may be other first-order motivators in the teacher-student relationship.

In the strand of literature on the economic theory of educational standards, an agency approach is applied toward designing high school graduation requirements (see Becker and Rosen 1992; Costrell 1994, 1997; Betts 1998). In crude terms, a policy maker searches for the passing test score that maximizes a specified social welfare function (typically, human capital) taking into account the optimal response by learning-averse students. Like in our model with observable grading rules, analysis is centered on the trade-offs of incentive and sorting effects that a change in educational standards (e.g., a higher passing score or an additional educational credential) can have on students' learning behavior. In this literature, however, the teacher-student relationship is not explicitly modeled, where the assumption is of invariant grading rules applied by teachers. Significantly, a reform of graduation requirements can have an effect on the form of a teacher's objective function and, accordingly, on the grading rules applied, which, in turn, may affect the achievement of the projected goals of the reform. We illustrate this argument when discussing policy applications later in the text.

Recently, the problems of grade compression and grade inflation have been approached from the perspective of the informative content of grades that university administration chooses to disclose. The information disclosure policy is modeled as a mapping from students' abilities into ability signals under the assumption that students' abilities are observed by administration (unlike in our model). With the objective of maximizing students' average job placement, Ostrovsky and Schwarz (2010) show that administration compresses ability signals sent to the job market if the mapping from student ability to job desirability is concave. In the model, they require disclosure policies be consistent so that a signal is equal to the expected ability of students entitled to the signal (as also imposed in our model with observable grading rules). While giving an explanation for compressed grades, their model, however, does not explain the emergence of inflated grades as pointed out by Chan, et al. (2007). In the latter work, the consistency requirement is relaxed so that schools can send inflated messages (as in our model with unobservable grading rules). They show how competition among schools for better student placements can further worsen the problem of compressed grades, ultimately resulting in inflated grades. Yet, their model does not explain grade inflation as a dynamic process, i.e., the gradual increase of grades over time, which our model does produce. Relatedly and similarly to our argument but in the context of information disclosure, Yang and Yip (2003) argue that grade inflation can arise from a free-rider problem between universities as the signaling value of a transcript is based on "collective reputation." However, this literature also assumes the invariance of grading rules and, thus, cannot address the empirically observed mismatch between students' abilities and grades: At the same university, professors of more able classes tend to apply more stringent grading rules notwithstanding identical disclosure policies.

On the theoretical side, our basic model with unobservable grading rules is a version of the monopolist screening model of Mussa and Rosen (1978), modified with free but bounded transfers from the principal to the agent. As we demonstrate, with free transfers it is never optimal to screen most efficient types. Gains from screening do not outweigh losses stemming from suppressed incentives for less able students imposed to ensure incentive compatibility. In the framework of monopolist screening, the scenario that, unlike the agent, the principal is indifferent to transfers is also studied in Guesnerie and Laffont (1984). They distinguish between the "type A" and "type B" preferences of the principal, where the difference is that a transfer enters only the latter (conventional) preferences. The "type A" preferences are primarily used to analyze the social planner's problem of social welfare maximization. A transfer is equivalent, figuratively speaking, to distributing money between two pockets of the same jacket, leaving the social welfare intact. Therefore, the framework of Guesnerie and Laffont (1984) does not apply to the problem studied here. In our model, the principal is actually of "type B" as she cares only about her own utility but does not bear the cost of motivating the agent.

The property of free incentive transfers arises in the model of self-enforcing relational contracts of Levin (2003). In a repeated relationship, compensation paid in a given period does not affect incentives in later play and, as a result, risk-neutral parties can transfer wealth between them at will. The dynamic enforcement constraint, however, puts a limit on wealth transfers, which depends on the total discounted value of the relationship. With a low discount factor and, thus, a low limit on transfers, the principal finds it optimal to pool the most efficient types for reasons similar to those in our model: At the margin, the benefit from screening is zero, but the shadow cost is positive. When the discount factor is sufficiently high, which allows for almost unbounded wealth transfers, it is optimal for the principal to elicit the socially desirable (first-best) effort levels from every agent type. In other words, in effort elicitation the principal is either constrained by wealth transfers or by the first-best effort levels. In our model, however, the optimal effort levels from the principal's perspective do not coincide with the social optimum, taken to be the sum of the teacher's and students' utilities, and are bounded only by available rewards (akin to the wealth constraints in Levin (2003)). It implies that in our model the constraint on transfers (grades) is always binding irrespective of its size.

In the literature quoted above as well as in the current paper, the central question is that of pooling and separation of agent types. The same question arises in various contexts with non-monetary transfers and which policy dominates depends on relative advantages. In wage setting, Fang and Moscarini (2005) show that with workers overconfident in their own skills no wage differentiation is a profit-maximizing policy due to a relatively larger impact of negative morale effects that stem from differentiation. In credit rating agency, Mathis, et al. (2009) show that when reputational concerns are not too strong credit rating agencies find it profitable to apply lax appraisal standards and inflate credit ratings. For the problem of conspicuous consumption Rayo (2013) characterizes necessary and sufficient conditions when a monopolist producer chooses to screen perfectly consumer types. The latter work closely resembles our model with observable grading rules and will be referred to later in the text.

2 Framework

There is a teacher teaching a class of students a particular subject. Students select the subject for exogenous reasons, e.g., it is compulsory in their curriculum.[2] The teacher's

[2]Arcidiacono (2004) empirically argues that students select their majors out of intrinsic preference. At the same time, Sabot and Wakeman-Linn (1991) show that students' choice of classes *not required for their majors* is responsive to their grade expectations. Here, we do not model this possibility.

goal is to pass on knowledge to the students. She has a technology, the exam test, that allows her from a test score $x \in [0, X]$ to assess the knowledge level attained by a student. (The upper bound X is large enough to allow for an interior solution.) We assume that the teacher's technology is perfect in the sense that there is a deterministic and direct relationship between test scores and knowledge levels. Students differ in their privately known learning abilities θ, distributed in the population according to a distribution $F(.)$ over an ability space $[\theta_L, \theta_H]$, $0 < \theta_L < \theta_H$. Achieving a test score x comes to a student of ability θ at an effort cost of $C(x)/\theta$, where $C(.) \geq 0$ is an increasing, strictly convex, and twice differentiable function with $C(0) = 0$ and $C'(0) = 0$.[3]

The teacher uses grades $r \in [0, 1]$ as incentives to motivate her students to perform on the exam. The upper bound of 1 is assumed to be institutionally preset and so is the grading framework: The assessment of student performance is done in the form of grades only.[4] The teacher-student relationship develops as follows. First, the teacher sets up a grading rule, which is a set of direct score-grade allocations, $\{x(\theta), r(\theta)\}$, $\theta \in [\theta_L, \theta_H]$. Then, each student reports a type $\widehat{\theta} \in [\theta_L, \theta_H]$ or, equivalently, chooses an allocation $(x(\widehat{\theta}), r(\widehat{\theta}))$. The teacher's objective is to maximize her expected utility defined as $\int V(x(\theta))dF(\theta)$, where $V'(.) > 0$. A student maximizes his utility of grade r less the

[3]The hidden information assumption can be replaced by a constraint that grades can only be conditioned on a student's exam score but not on his ability even if observed by the teacher (e.g., from previous academic records). For classes of large size, the two approaches are equivalent.

[4]In the model, it is ultimately ability signals that the teacher rewards her students with and the form of appraisal or a grading framework is immaterial as long as it spans the ability space. However, this observation holds only if the teacher cannot type-discriminate her students by, e.g., sending confidential recommendation letters. In this paper, we do not consider the latter possibility, which, though, seems to be relevant for small size graduate classes. To model this situation, the message space of grades would need to be augmented with types, the route left for future research.

effort cost of $C(x)/\theta$ spent to obtain it.

We define the utility of a grade by the ability level signaled by the grade to the job market. We start with the case when the utility of a grade is fixed, i.e., independent of the grading rule applied or of the distribution of grades in the class. In particular, a student's utility of a grade r is given by $\varphi(r)$, where $\varphi : [0,1] \to [\theta_L, \theta_H]$ is a publicly observable grading standard. In the paper, this case is related to the situation when the job market does not observe particular grading rules and relies on the commonly perceived grading standard $\varphi(.)$ for the purpose of ability inference.[5] We also analyze the case of unobservable grading rules in the dynamic setting with adaptive expectations. Then, we consider the case when the job market observes the grading rule applied and can make a correct inference about a student's expected ability. Under this circumstance, the utility of a grade is dependent on the grading rule and is set to be equal to the mean of types receiving the grade.

Here, we impose some further structure on the model. The distribution function $F(.)$ is twice differentiable with its probability density function $f(.) > 0$. In addition, we assume that the hazard rate, $f(\theta)/(1 - F(\theta))$, is non-decreasing in θ. Until further notice, we let the teacher's utility function be linear in test scores, $V(x) = x$. The teacher's linear utility can be interpreted that she equally cares about low- and high-ability students'

[5]This situation can arise when there are too many academic institutions for the job market to distinguish among: The number of degree-granting institutions in the US gradually increased from 563 in 1869–1870 to 4352 in 2007–2008, see Table 188 of 2009 Digest of Education Statistics, http://nces.ed.gov/programs/digest/d09/tables/dt09_188.asp. Another justification for the assumption of fixed grade values is that students can be ignorant about the distribution of abilities in the class or they are of bounded rationality to discern the expected distribution of grades that results from the grading rule set up by the teacher.

performance. When discussing policy applications later in the text, the linear case serves as a benchmark against the cases where 1) $V(.)$ is convex (the teacher puts more weight on high-ability students) and 2) $V(.)$ is concave (the teacher puts more weight on low-ability students). We make a technical assumption that the schedules of score and grade allocations, $\{x(\theta)\}$ and $\{r(\theta)\}$, are piecewise continuous functions. Finally, if a student opts out of the class, then the job market assigns the ability of θ_L, which defines the reservation utility of class participation.

3 Unobservable grading rules: Static model

In this section, we study a static teacher-student agency problem with grading rules unobservable by the job market. Unobservability implies that the ability-signaling values of grades are fixed and, thus, there is no interdependence of utilities across students. As a result, the teacher finds grades not only costless but also abundant. Therefore, the teacher-student relationship in this context turns alike a *single-agent* (i.e., single-student) agency problem with hidden information. In particular, the teacher's problem reduces to screening ability types with the help of a grading rule. However, there is an important difference with the canonical monopolistic screening problem. In our model, the transfer (grade) enters only the student's objective function.

Next, we characterize and solve the teacher's problem of optimal grading, followed by the discussion of the empirically relevant properties of the solution.

3.1 Teacher's problem

By the Revelation Principle, we look for the grading rule that maximizes the student's expected knowledge among those rules that induce the student to reveal his type truthfully:

$$\max_{\{x(\theta),r(\theta)\}} \int_{\theta_L}^{\theta_H} x(\theta)dF(\theta) \tag{1}$$

s.t.

$$\varphi(r(\theta)) - \frac{C(x(\theta))}{\theta} \geq \varphi(r(\widehat{\theta})) - \frac{C(x(\widehat{\theta}))}{\theta}, \tag{2}$$

$$\varphi(r(\theta)) - \frac{C(x(\theta))}{\theta} \geq \theta_L, \tag{3}$$

$$0 \leq r(\theta) \leq 1 \text{ for all } \theta \text{ and } \widehat{\theta} \text{ in } [\theta_L, \theta_H]. \tag{4}$$

Above, (2) is the student's incentive compatibility constraint, (3) is the individual rationality constraint, and the last constraint imposes an upper bound on grades.

As grades do not enter the teacher's objective function directly, we can transform the problem by the one where the teacher rewards the student with ability signals $\widehat{\varphi}$ from the closure of the range $\mathcal{R}(\varphi)$ of the grading standard $\varphi(.)$. (We take the closure in order to ensure the existence of a solution.) Given a signal schedule $\{\widehat{\varphi}(\theta)\}$, the corresponding grade schedule $\{r(\theta)\}$ can be inverted by $\{r(\theta) = r : \varphi(r) = \widehat{\varphi}(\theta)\}$. For analytical convenience, we rewrite constraints (2) and (3) pre-multiplied by θ. Now, the teacher looks for the grading rule $\{x(\theta), \widehat{\varphi}(\theta)\}$ that maximizes her expected utility

$$\max_{\{x(\theta),\widehat{\varphi}(\theta)\}} \int_{\theta_L}^{\theta_H} x(\theta)dF(\theta) \tag{5}$$

s.t.

$$\theta\widehat{\varphi}(\theta) - C(x(\theta)) \geq \theta\widehat{\varphi}(\widehat{\theta}) - C(x(\widehat{\theta})), \tag{6}$$

$$\theta(\widehat{\varphi}(\theta) - \theta_L) - C(x(\theta)) \geq 0, \tag{7}$$

$$\widehat{\varphi}(\theta) \in \overline{\mathcal{R}(\varphi)} \text{ for all } \theta \text{ and } \widehat{\theta} \text{ in } [\theta_L, \theta_H]. \tag{8}$$

In this section, we assume that the range of the grading standard $\varphi(.)$ comprises the whole type space, i.e., $\mathcal{R}(\varphi) = [\theta_L, \theta_H]$.

Applying the standard constraint-simplification techniques, constraints (6) and (7) can be equivalently expressed as

$$\theta\widehat{\varphi}(\theta) - C(x(\theta)) = \theta_L^2 + \int_{\theta_L}^{\theta} \widehat{\varphi}(\tilde{\theta})d\tilde{\theta}, \tag{9}$$

$$\widehat{\varphi}(\theta) \text{ is non-decreasing.} \tag{10}$$

First, in the optimum constraint (7) is binding for $\theta = \theta_L$, i.e., $\theta_L\widehat{\varphi}(\theta_L) - C(x(\theta_L)) = \theta_L^2$. Second, constraint (6) implies both (9) and (10), where the former implication follows from the integral form envelope theorem (see Milgrom and Segal 2002). Lastly, it is straightforward to demonstrate that (9) and (10) imply (6) and (7).

Letting $U(\theta) = \theta\widehat{\varphi}(\theta) - C(x(\theta))$, the teacher's problem becomes

$$\max_{\{\theta_L \leq \widehat{\varphi}(\theta) \leq \theta_H, U(\theta)\}} \int_{\theta_L}^{\theta_H} C^{-1}\left(\theta\widehat{\varphi}(\theta) - U(\theta)\right) f(\theta)d\theta \tag{11}$$

s.t.

$$U'(\theta) = \widehat{\varphi}(\theta), \tag{12}$$

$$\widehat{\varphi}(\theta) \text{ is non-decreasing.} \tag{13}$$

This maximization problem can be analyzed as an optimal control problem with $U(\theta)$ as a state variable and $\widehat{\varphi}(\theta)$ as a control variable. Its solution is presented in the next proposition.

Proposition 1 *Let $\mathcal{R}(\varphi) = [\theta_L, \theta_H]$. The grading rule $\{x(\theta), \widehat{\varphi}(\theta)\}$ that solves optimization problem (5) is characterized by*

$$x(\theta) = C'^{-1}\left(\frac{f(\theta)\theta^2 C'(x(\theta^*))}{f(\theta^*)\theta^{*2}}\right), \tag{14}$$

$$\widehat{\varphi}(\theta) = \theta_L + \frac{C(x(\theta))}{\theta} + \int_{\theta_L}^{\theta} \frac{C(x(\tilde{\theta}))}{\tilde{\theta}^2} d\tilde{\theta} \tag{15}$$

for θ in $[\theta_L, \theta^*)$, where

$$\theta^* = \min\{\theta : \theta f(\theta) - (1 - F(\theta)) \geq 0, \theta \in [\theta_L, \theta_H]\}, \tag{16}$$

and by $x(\theta) = x^*$ *and* $\widehat{\varphi}(\theta) = \theta_H$ *for* θ *in* $[\theta^*, \theta_H]$*, where* x^* *is determined from*

$$\theta_H - \frac{C(x^*)}{\theta^*} - \int_{\theta_L}^{\theta^*} \frac{C(x(\theta))}{\theta^2} d\theta = \theta_L. \tag{17}$$

Before we discuss solution properties, it is worthwhile drawing a parallel to the related problem of single-object pricing. The pricing problem takes a form similar to that of problem (5), with the teacher interpreted as a seller and the student as a buyer. Then, $\{x(\theta)\}$ reads as a payment schedule, and $\{\widehat{\varphi}(\theta)\}$, with its range normalized to $[0, 1]$, as a probability schedule of selling the object to different valuation types θ. There is, however, an important difference. In our model, a score x enters the student's utility in the non-linear way, whereas in the standard pricing model a payment x enters the buyer's utility linearly. This difference has a significant effect on the outcome. In the pricing model, the sale occurs with probability 1 to types $\theta \geq \theta^*$ and with probability zero to the rest. But in our model, the bang-bang outcome is suboptimal as, technically, the Hamiltonian is not linear in the control variable. Intuitively, with increasing marginal costs of effort the gain from eliciting positive effort levels from excluded types outweighs the corresponding loss that comes from the reduction of the effort level of the supramarginal types. Thus, if interpreted in the price-setting language, we obtain that the seller sells the object with probability 1 to the same types, $\theta \geq \theta^*$, but he also sells it to types $\theta < \theta^*$ with a positive probability.

3.2 Solution properties

Pooling at the top

Proposition 1 shows that in the teacher's optimum an interval of types, $[\theta^*, \theta_H]$, is pooled for a uniform allocation.[6] In other words,

Result 1 *The teacher compresses most efficient types for the highest grade.*

More generally, if the principal does not bear or internalize the cost of rewarding the agent, then the optimal contract pools most efficient agent types. This finding is not surprising even though in contrast to the ubiquitous "no pooling at the top" property of optimal contracts with monetary incentives (and also with status incentives, see Moldovanu, et al. 2007; Dubey and Geanakoplos 2010). Suppose it is the case that only the most efficient type receives the highest reward. To make this allocation incentive compatible, the principal has to suppress incentives for other types in order to prevent the most efficient type from misreporting. Now, consider the principal marginally "tilting up" the schedule of score allocations against the corresponding decrease in the score allocation of the most efficient type. As this decrease is weighted with a very small probability, it is certainly outweighed by the sum of increases in test scores of the other types. Thus, the principal finds it optimal to increase the probability mass of agent types that are subject to the highest reward until the gains and losses described offset each other.

Referring back to the teacher-student relationship, we show next that pooling the most efficient types becomes more prevalent when the distribution of abilities is skewed toward the end of low types.

[6]The interval $[\theta^*, \theta_H]$ is not empty as the starting point θ^* is bound to be strictly less than θ_H. Because of $f(.) > 0$ we have $\theta_H f(\theta_H) - (1 - F(\theta_H)) > 0$. Then, due to the monotone hazard rate and the continuity of $F(.)$ there must be some $\theta < \theta_H$ such that $0 \leq \theta f(\theta) - (1 - F(\theta)) < \theta_H f(\theta_H) - (1 - F(\theta_H))$.

Mismatch between grades and abilities

Consider two classes of students, A and B. In class A, students come from a population where abilities are distributed according to $F_A(.)$ and in class B – according to $F_B(.)$ over the same support. Denote the student types from the two classes by θ_A and θ_B, respectively. Let θ_B be smaller than θ_A in the likelihood ratio order:

$$(18) \qquad \frac{f_B(\theta)}{f_A(\theta)} \leq \frac{f_B(\theta')}{f_A(\theta')} \text{ if } \theta < \theta' \text{ for all } \theta, \theta' \text{ in } [\theta_L, \theta_H],$$

where $f_A(.)$ and $f_B(.)$ are the probability density functions of the corresponding distributions. The interpretation of this stochastic dominance condition is that students from class A are held more able than those from class B. Formally, this condition implies that $\int_{\theta'}^{\theta''} \theta f_A(\theta) d\theta / (F_A(\theta'') - F_A(\theta')) \geq \int_{\theta'}^{\theta''} \theta f_B(\theta) d\theta / (F_B(\theta'') - F_B(\theta'))$ for any interval $[\theta', \theta''] \subseteq [\theta_L, \theta_H]$. In words, for any restriction of the ability space the expected student ability in class A is greater than that in class B.

Let $\{x_A(\theta), \widehat{\varphi}_A(\theta)\}$ and $\{x_B(\theta), \widehat{\varphi}_B(\theta)\}$ be the optimal grading rules in the two classes, respectively. We can establish the following relationship

Result 2 *If (18) holds, then* $\widehat{\varphi}_B(\theta) \geq \widehat{\varphi}_A(\theta)$ *for every* θ *in* $[\theta_L, \theta_H]$.

In words, a teacher with lower expectations about her students' abilities should be more lenient in grading. The intuition behind leniency in grading in weaker classes comes from the teacher's attempt to extract more effort from more numerous lower-ability students, and *vice versa*.

Example

Results 1 and 2 are illustrated in Figure 1. It depicts the optimal grading rules for two classes, where ability types in class A are distributed over $[0.5, 1.5]$ according to the distribution with $f_A(\theta) = \theta$ and in class B uniformly. The effort cost function takes the form of $C(x) = x^2/2$. The top diagram shows the optimal score schedules and the bottom diagram the optimal grade schedules. In both classes, the teachers pool most efficient student types, but the teacher of class A pools fewer types but against a higher score. In class B, besides more pooling, we have flatter schedules.

[Insert Figure 1 here]

4 Unobservable grading rules: Dynamic model

According to Result 1, the grade perceived to reveal the highest ability is given to a range of different ability types. Obviously, this outcome cannot be equilibrium in the long run as over time adjustment in grading standards will necessarily take place to account for any mismatch between students' grades and their abilities. Here, we introduce an adjustment process of grading standards into the model. In every period, the ability-signaling value of a grade is set to be equal to the average ability of students who received the grade in the previous period. The assumption is that the job market uses the empirical distribution of students' abilities and grades from the previous period to make inference about students' abilities in the current period. In this section, first, we are interested in the dynamics of the optimal grading rule under the adaptive expectations and, then, in the properties of consistent grading standards, defined as

Definition 1 *A grading standard $\varphi(.)$ is consistent if $\varphi(r)$ is equal to the mean of types θ that receive the grade r under the grading rule $\{x(\theta), \widehat{\varphi}(\theta)\}$ that solves (5) with $\varphi(.)$.*

To begin with, we expand the static model with unobservable grading rules in the following way. There is a large number of teachers. In every period $t \in \{1, 2, ...\}$, each teacher teaches a class of students coming from the same population so that all teachers face the identical grading problem. An individual teacher's grading practice has no effect on the dynamics of grading standards and, therefore, in every period a teacher maximizes her current expected utility only. Specifically, in period t a teacher sets up a grading rule $\{x_t(\theta), \widehat{\varphi}_t(\theta)\}$ that solves (5) under the current grading standard $\varphi_t(.)$. In period $t+1$, grading standards adjust according to

$$\varphi_{t+1}(r) = E(\theta : \widehat{\varphi}_t(\theta) = \varphi_t(r)). \tag{19}$$

As it immediately follows from adjustment process (19), we can equivalently define a consistent grading standard $\varphi_t(.)$ as

Lemma 1 *A grading standard $\varphi_t(.)$ is consistent if and only if $\varphi_{t+1} = \varphi_t$.*

4.1 Dynamics

The job market's temporal adjustments of its perception of grading standards have a direct effect on available rewards at teachers' disposal, so influencing teachers' design of grading rules. To illustrate the dynamics of grading rules, we start from that the initial grading standard $\varphi_1(.)$ is a one-to-one strictly increasing mapping from grades to abilities so that $\overline{\mathcal{R}(\varphi_1)} = [\theta_L, \theta_H]$. A teacher's optimal grading rule $\{x_1(\theta), \widehat{\varphi}_1(\theta)\}$ is given by Proposition 1: pooling the types in $[\theta_1^*, \theta_H]$, where $\theta_1^* = \theta^*$ from (16), and screening the remaining types. In period $t = 2$, the range of rewards under the new grading standard $\varphi_2(.)$ shrinks to $\overline{\mathcal{R}(\varphi_2)} = [\theta_L, \theta_1^*] \cup \overline{\theta_1^*}$, where $\overline{\theta_1^*} = E(\theta \in [\theta_1^*, \theta_H])$ denotes the mean of the types that receive the highest grade under $\{x_1(\theta), \widehat{\varphi}_1(\theta)\}$. Now, we solve problem (5) with $\varphi_2(.)$ for the optimal grading rule $\{x_2(\theta), \widehat{\varphi}_2(\theta)\}$ in period $t = 2$.

Because the range of rewards, $\overline{\mathcal{R}(\varphi_2)}$, consists of an interval and a distinct point, the teacher searches for the starting point of the pooling-at-the-top interval, denoted by

θ_2, and signal allocations $\widehat{\varphi}_2(\theta)$ for $\theta \leq \theta_2$ subject to $\widehat{\varphi}_2(\theta) \leq \theta_1^*$. (Note that θ_1^* is the second-highest reward the teacher can give under the grading standard $\varphi_2(.)$.) By the integral form envelope theorem, the optimal score allocations $x_2(\theta)$ for $\theta \leq \theta_2$ take the form of

$$x_2(\theta) = C^{-1}\left(\theta\widehat{\varphi}_2(\theta) - \int_{\theta_L}^{\theta} \widehat{\varphi}_2(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right). \tag{20}$$

The types in $(\theta_2, \theta_H]$ receive the signal allocation of $\varphi(\theta) = \overline{\theta_1^*}$ and their score allocation is given by

$$\overline{x}_2 = C^{-1}\left(\theta_2\overline{\theta_1^*} - \int_{\theta_L}^{\theta_2} \widehat{\varphi}_2(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right). \tag{21}$$

The Lagrangian of the teacher's problem is

$$\mathcal{L}(\widehat{\varphi}_2(\theta), \theta_2, \mu(\theta)) = \int_{\theta_L}^{\theta_2} x_2(\theta)f(\theta)d\theta + \overline{x}_2(1 - F(\theta_2)) + \mu(\theta)(\theta_1^* - \widehat{\varphi}_2(\theta)), \tag{22}$$

where $\mu(\theta)$ is the Lagrange multiplier on the constraint that $\widehat{\varphi}_2(\theta) \leq \theta_1^*$ for $\theta \leq \theta_2$.

The optimal starting point, θ_2^*, is characterized by the first-order condition

$$\frac{\overline{\theta_1^*} - \widehat{\varphi}_2(\theta_2^*)}{C'(\overline{x}_2)} \frac{1 - F(\theta_2^*)}{f(\theta_2^*)} = \overline{x}_2 - x_2(\theta_2^*). \tag{23}$$

Condition (23) is analogous to pooling condition (16) in the case of the convex range $\mathcal{R}(\varphi)$. Because $C^{-1}(.)$ is strictly concave, it implies that

$$\frac{\theta_2^*\left(\overline{\theta_1^*} - \widehat{\varphi}_2(\theta_2^*)\right)}{C'(\overline{x}_2)} < \overline{x}_2 - x_2(\theta_2^*) < \frac{\theta_2^*\left(\overline{\theta_1^*} - \widehat{\varphi}_2(\theta_2^*)\right)}{C'(x_2(\theta_2^*))}. \tag{24}$$

Then, given (24), the first-order condition in (23) renders

$$\frac{1 - F(\theta_2^*)}{\theta_2^* f(\theta_2^*)} > 1, \tag{25}$$

which in turn implies $\theta_2^* < \theta_1^*$. (Recall that $(1 - F(\theta_1^*))/(\theta_1^* f(\theta_1^*)) = 1$.) Thus, we get more types pooled for the highest grade, which is due to the gap between the highest and second-highest rewards. The reason is that the teacher increases the mass of types subject to the highest reward as inframarginal types are bound for significantly lower rewards and, thus, lower scores.

Now consider the allocations for the remaining types, where the first-order condition with respect to $\widehat{\varphi}_2(\theta)$, $\theta \leq \theta_2^*$, is

$$\text{(26)} \qquad \frac{\theta f(\theta)}{C'(x_2(\theta))} - \int_{\theta}^{\theta_2^*} \frac{f(\tilde{\theta})}{C'(x_2(\tilde{\theta}))} d\tilde{\theta} - \frac{1 - F(\theta_2^*)}{C'\left(\overline{x}_2\right)} = \mu(\theta).$$

The last term of the left-hand side is strictly smaller than $(\theta_2^* f(\theta_2^*))/C'(x_2(\theta_2^*))$ because of (23) and (24). Hence, at $\theta = \theta_2^*$ the left-hand side of (26) is positive, implying that $\mu(\theta_2^*) > 0$ or $\widehat{\varphi}_2(\theta_2^*) = \theta_1^*$. By the monotonicity of a score schedule and the continuity of $F(.)$, the multiplier $\mu(\theta)$ stays positive for a range of types. The gains from a marginal increase in effort levels for types close to θ_2^* outweigh the corresponding reductions of effort levels for higher types. This implies a second interval of types pooled for θ_1^* and, thus, a further coarsening of grading.

To complete the characterization of the optimal grading rule $\{x_2(\theta), \widehat{\varphi}_2(\theta)\}$, let $\theta_{2,2}^*$ denote the starting point of the second pooling interval, $(\theta_{2,2}^*, \theta_2^*]$. All the types in this interval receive the allocation $(x_2(\theta_2^*), \theta_1^*)$. Types $\theta \leq \theta_{2,2}^*$ are screened and their optimal score allocations $x_2(\theta)$ are determined by (26) with $\mu(\theta) = 0$. The type $\theta_{2,2}^*$ is found from (26) and (17) in Proposition 1, where θ^* is replaced with $\theta_{2,2}^*$, x^* with $x_2(\theta_2^*)$, and $x(\theta)$ with $x_2(\theta)$. Finally, the uniform score allocations, $\overline{x}_2$ and $x_2(\theta_2^*)$, are pinned down by (23) and the corresponding incentive compatibility constraint.

In the next period, $t = 3$, the job market adopts a new grading standard, $\varphi_3(.)$ according to (19). Then, teachers choose a new grading rule, $\{x_3(\theta), \widehat{\varphi}_3(\theta)\}$, that solves problem (5) with $\varphi_3(.)$. The range of rewards available to teachers, $\overline{\mathcal{R}(\varphi_3)}$, consists of an interval of types and two distinct points that are equal to the expected types of the two pooling intervals in $\{x_2(\theta), \widehat{\varphi}_2(\theta)\}$, respectively. A teacher looks for the optimal starting points of two pooling intervals, in which types receive the two distinct ability signals, respectively, and for signal allocations for the remaining types. But, as we show in the proof of the next result, the teacher pools an additional interval of types. For the same reason, this pattern reoccurs in subsequent periods: the teacher always pools a range of types for the highest ability signal from the interval of available signals. We numerically illustrate this dynamics later in the text.

The algorithm, described above, is repeated until a consistent grading standard is obtained, i.e., $\varphi_{t+1} = \varphi_t$. The convergence of the adjustment process can be immediately established when the optimal expected score decreases over time, $E(x_t) \geq E(x_{t+1})$. This property, together with the existence of a lower bound (the expected score in any period cannot be lower than the expected score under complete pooling), implies that the sequence $\{E(x_t)\}$ must converge. The convergence of $\{E(x_t)\}$ implies, in turn, the convergence of the adjustment process and, thus, the existence of a convergent grading standard $\varphi(.)$. At the same time, the monotonicity property $E(x_t) \geq E(x_{t+1})$ follows from the revealed preference argument when in consecutive periods there is no improvement in available ability signals at teachers' disposal.[7] Next, we discuss properties of a consistent grading standard, followed by numerical analysis.

[7]In some irregular cases, the monotonicity property can, however, be violated. Consider a current grading standard with two pooling intervals only and thus two grades available. If the cost function $C(.)$ is locally very convex, the teacher may find it optimal to pool fewer types for the highest reward compared to the previous period. This would lead to an improvement in available rewards for the following period and possibly a higher expected score. We do not explore such cases here.

4.2 Consistent grading standards

The next result shows that the observations made when deriving the optimal grading rule in period $t = 2$ can be generalized as the properties of a consistent grading standard. This result will prove useful in characterizing these standards.

Result 3 *Given a consistent grading standard, teachers*

(i) do not separate any interval of types,

(ii) pool more types for the highest grade than in period $t = 1$*.*

As we show in the proof, a grading standard that perfectly screens an interval of types cannot be consistent. If such an interval existed, the teacher would pool some types for the highest ability signal available in this interval. Generally, effort extraction from lower types dominates distortions imposed on higher types. Thus, when individual grading rules are unobservable and separately have no effect on the dynamics of grading standards, the practice of coarse grading will emerge over time. Part (ii) of Result 3 points to the direction of grade inflation as the highest grade is assigned to more students. As already mentioned in the discussion of the grading problem in period $t = 2$, this result is a consequence of a gap that arises between the values of the highest and the second-highest grades and the convexity of the cost function $C(.)$. The dynamics of the pooling-at-the-top interval is numerically explored in the next subsection.

Now, drawing on part (i) of Result 3, we are able to offer a succinct characterization of consistent grading standards as it is sufficient to analyze standards with pooling intervals only. We use a partition $\boldsymbol{\theta} = (\theta_0, \theta_1, ..., \theta_{n-1}, \theta_n)$ of the type space,

(27) $$\theta_L = \theta_0 \leq \theta_1 \leq ... \leq \theta_{n-1} \leq \theta_n = \theta_H,$$

to express a grading standard with n pooling intervals, $[\theta_{i-1}, \theta_i]$, $i = 1, ..., n$. Let Θ^n denote the set of all possible partitions $\boldsymbol{\theta}$ satisfying (27). A grading standard $\boldsymbol{\theta} \in \Theta^n$ with n distinct pooling intervals defines n distinct grades that a teacher finds at her disposal. The ability-signaling values of these grades are given by the function $\boldsymbol{\varphi} : \Theta^n \to [\theta_L, \theta_H]^n$, where

$$\boldsymbol{\varphi}_i(\boldsymbol{\theta}) = \frac{\int_{\theta_{i-1}}^{\theta_i} \theta f(\theta) d\theta}{F(\theta_i) - F(\theta_{i-1})} \tag{28}$$

is the expected ability in interval i, i.e., in $[\theta_{i-1}, \theta_i]$, $i = 1, ..., n$. We write $\boldsymbol{\varphi}_0(\boldsymbol{\theta}) = \theta_L$.

Given a grading standard $\boldsymbol{\theta} \in \Theta^n$ with n distinct intervals, any feasible grading rule can also be characterized by a partition $\boldsymbol{\theta}' \in \Theta^n$ because at most n distinct grades are available. The teacher's expected utility from a grading rule $\boldsymbol{\theta}'$ is given by the function $G : \Theta^n \times \Theta^n \to \mathbb{R}$ defined as

$$G(\boldsymbol{\theta}'; \boldsymbol{\theta}) = \sum_{i=1}^{n} \pi_i(\boldsymbol{\theta}') \boldsymbol{x}_i(\boldsymbol{\theta}'; \boldsymbol{\theta}), \tag{29}$$

where $\pi_i(\boldsymbol{\theta}') = F(\theta'_i) - F(\theta'_{i-1})$ is the probability weight on the score

$$\boldsymbol{x}_i(\boldsymbol{\theta}'; \boldsymbol{\theta}) = C^{-1} \left(\sum_{j=1}^{i} \theta'_{j-1} (\boldsymbol{\varphi}_j(\boldsymbol{\theta}) - \boldsymbol{\varphi}_{j-1}(\boldsymbol{\theta})) \right). \tag{30}$$

A grading standard $\boldsymbol{\theta}$ is consistent if

$$\boldsymbol{\theta} = \arg \max_{\boldsymbol{\theta}' \in \Theta^n} G(\boldsymbol{\theta}'; \boldsymbol{\theta}). \tag{31}$$

In particular, a consistent grading standard $\boldsymbol{\theta}$ with n distinct pooling intervals exists only if it is an interior solution to problem (31). In other words, it solves the system of $n-1$ equations, derived from the first-order conditions with respect to θ'_i, $i = 1, ..., n-1$,

$$\frac{\boldsymbol{\varphi}_{i+1}(\boldsymbol{\theta}) - \boldsymbol{\varphi}_i(\boldsymbol{\theta})}{f(\theta_i)} \sum_{j=i+1}^{n} \frac{\pi_j(\boldsymbol{\theta})}{C'(\boldsymbol{x}_j(\boldsymbol{\theta}; \boldsymbol{\theta}))} = \boldsymbol{x}_{i+1}(\boldsymbol{\theta}; \boldsymbol{\theta}) - \boldsymbol{x}_i(\boldsymbol{\theta}; \boldsymbol{\theta}) \tag{32}$$

with (27) satisfied with strict inequalities.

4.3 Numerical analysis

Here, we present a numerical analysis of the dynamics of grading rules done by extending the two examples from the previous section. The main purpose is to show how our model can produce grade inflation.[8]

Figure 2 illustrates the case when types are uniformly distributed over $[0.5, 1.5]$. The optimal grading rule in period t is shown as a partition of the type space. The rightmost interval stands for the pooling interval at the top, with its score allocation x indicated above it. The screening interval is marked with "Screen." As we can see, the screening interval disappears already in period $t = 3$, with only two pooling intervals left, $[0.5, 0.62]$ and $(0.62, 1.5]$. Over time, the latter expands at the expense of the former yielding grade inflation: In every period, the highest grade is given to more students and for a lower score x. Ultimately, the consistent grading standard we get is the one that pools the whole type space.

[Insert Figure 2 here]

Figure 3 illustrates the case when the distribution of types has a linear density function, $f(\theta) = \theta$. The dynamics of optimal grading rules is shown for six periods, all of which follow the same pattern. The screening interval shrinks, an additional pooling

[8] To demonstrate grade inflation, we resort to numerical methods rather than analytical ones. Analytical analysis is hardly tractable in our model because of, among other things, non-convex and variable choice set $\overline{\mathcal{R}(\varphi)}$. Drawing on part (i) of Result (3), one analytical way to proceed would be to specify the conditions when a teacher's expected utility function $G(\boldsymbol{\theta}'; \boldsymbol{\theta})$ in (29) is supermodular in $\boldsymbol{\theta}'$ and has increasing differences in all θ'_i and θ_j. This would imply monotone comparative statics, which is the expansion of pooling intervals. However, for $G(\boldsymbol{\theta}'; \boldsymbol{\theta})$ to have the desired properties, which are, equivalently, non-negative cross derivatives, it requires imposing endogenous restrictions on the cost function $C(.)$ and distribution $F(.)$.

interval appears, and the pooling interval at the top expands with its score x decreasing. For this example, the derivation of a consistent grading standard through iteration is not attempted because of the increasing complexity of the problem. Instead, we use (32) to calculate consistent grading standards. Several are possible, e.g., one with the partition $(0.5, 0.568, 0.649, 0.689, 0.792, 1.5)$ has five distinct pooling intervals.[9]

[Insert Figure 3 here]

5 Observable grading rules

This section deals with the situation when the job market observes the grading rule applied and can correctly infer a student's expected ability. In particular, the utility of a grade is given by the average ability of students who receive the grade. Thus, unlike in the case with unobservable grading rules, there is no independence of utilities among students anymore. Here, a more lenient grading rule translates into lower values of grades and, as a result, adversely affects students' choice of learning effort. Below, we analyze when this downside effect of lenient grading outweighs the upside effect – more effort extraction from lower types. The model with observable grading rules closely resembles the benchmark model of Rayo (2013) on the monopolistic signal provision of conspicuous consumption, upon which we closely draw whenever possible.

Next, we start with defining the teacher's problem of optimal grading under observable grading rules and, then, discuss conditions for separation of students' abilities.

[9] A consistent standard with more pooling intervals could not be found.

5.1 Teacher's problem

Given a grading rule $\{x(\theta), r(\theta)\}$, define a function $\gamma : [\theta_L, \theta_H] \to [\theta_L, \theta_H]$ by

$$\gamma(\theta) = E\left[\widehat{\theta} : r(\widehat{\theta}) = r(\theta)\right]. \tag{33}$$

The value $\gamma(\theta)$ is equal to the mean of the types $\widehat{\theta}$ for which the grading rule prescribes the grade $r(\theta)$. When all types report truthfully, $\gamma(\theta)$ determines the expected ability inferred by the market and, thus, a student's utility from the grade $r(\theta)$. In definition (33), the grade allocation $r(.)$ can be equivalently replaced by the corresponding score allocation $x(.)$

$$\gamma(\theta) = E\left[\theta' : x(\theta') = x(\theta)\right]. \tag{34}$$

Therefore, it is sufficient to analyze grading rules in a reduced form, i.e., as $\{x(\theta)\}$ only.

The teacher's problem is to find the score schedule $\{x(\theta)\}$ that maximizes

$$\max_{\{x(\theta)\}} \quad \int_{\theta_L}^{\theta_H} x(\theta) dF(\theta) \tag{35}$$

s.t.

$$\theta\gamma(\theta) - C(x(\theta)) \geq \theta\gamma(\widehat{\theta}) - C(x(\widehat{\theta})), \tag{36}$$

$$\theta(\gamma(\theta) - \theta_L) - C(x(\theta)) \geq 0, \text{ for all } \theta \text{ and } \widehat{\theta} \text{ in } [\theta_L, \theta_H], \tag{37}$$

where (36) and (37) are the student's incentive compatibility and individual rationality constraints, respectively, pre-multiplied by θ. Next, we start with transforming the teacher's problem (35) into the equivalent problem that is maximized with respect to ability signal schedule $\{\gamma(\theta)\}$. The set of feasible ability signal schedules under incentive compatibility, which we denote by Γ, is characterized by the following condition

$$\gamma(\theta) = E[\widehat{\theta} : \gamma(\widehat{\theta}) = \gamma(\theta)], \tag{38}$$

which is similar in meaning to (33).

As before, we replace all the incentive compatibility and individual rationality constraints with

$$x(\theta) = C^{-1}\left(\theta\gamma(\theta) - \int_{\theta_L}^{\theta}\gamma(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right) \tag{39}$$

$$x(\theta) \text{ is non-decreasing.} \tag{40}$$

From (39) the score allocation $x(\theta)$ is non-decreasing in θ if and only if $\gamma(\theta)$ is non-decreasing. Thus, the teacher's problem can be expressed as finding the signal schedule $\{\gamma(\theta)\}$ that maximizes

$$\max_{\{\gamma(\theta)\}\in\Gamma} \int_{\theta_L}^{\theta_H} C^{-1}\left(\theta\gamma(\theta) - \int_{\theta_L}^{\theta}\gamma(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right) dF(\theta) \tag{41}$$

$$\text{s.t.} \quad \gamma(\theta) \text{ is non-decreasing.}$$

5.2 Solution properties

The teacher's problem (41) can be interpreted as determining the intervals of pooling and separation of types. Similarly to Rayo (2013), the teacher needs to take into account three effects pooling has on the expected score relative to separation. First, student types at the low end of a pooling interval can be made to exert a higher effort because their type inferred is higher than that under separation. Conversely, the second effect is that student types at the high end of the pooling interval will exert a lower effort than that under separation. Finally, the third effect comes from changes in score allocations for the supramarginal types (to the right from the pooling interval). All in all, the teacher chooses to pool the types in a given interval if the joint outcome of the three effects is positive.

Problem (41) with a linear cost function $C(.)$ is analyzed in Rayo (2013). He provides a necessary and sufficient condition for the optimality of full separation of types, i.e., $\gamma(\theta) = \theta$. For our specification of the model with linear $C(.)$, this condition requires $\theta - (1 - F(\theta))/f(\theta)$ be non-decreasing for $\gamma(\theta) = \theta$ to be the solution. However, with a nonlinear cost function $C(.)$ there is no succinct characterization of the full separation condition. One of the factors determining the results for our model is whether the score schedule in (39) is convex under separation. In particular, the convexity is a sufficient condition for no pooling at the top.

Proposition 2 *If $\{x(\theta)\}$ in (39) with $\gamma(\theta) = \theta$ is convex over $[\underline{\theta}, \theta_H]$, $\theta_L \leq \underline{\theta} < \theta_H$, then there is $\theta' \in [\underline{\theta}, \theta_H)$ such that $\gamma(\theta) = \theta$ is the solution to (41) for all $\theta \in [\theta', \theta_H]$.*

The idea of the proof is that, once the convexity condition is met, a grading rule containing a uniform allocation at the top can be improved upon by separating those types with distinct allocations. In terms of the three effects discussed above, with convexity the second effect dominates the first one, whereas the third is absent (because there are no supramarginal types). The convexity condition is equivalent to that the marginal disutility from learning decreases in ability

$$\frac{d}{d\theta}\left(\frac{C_x(x(\theta))}{\theta}\right) < 0 \tag{42}$$

or, after differentiating and dropping the arguments,

$$C_{xx} \leq \left(\frac{C_x}{\theta}\right)^2. \tag{43}$$

In other words, the score schedule $\{x(\theta)\}$ in (39) is convex under separation if the cost function $C(.)$ is not "too convex." (For instance, this condition is met for a quadratic cost function.)

The convexity condition in (43) is, however, not sufficient for $\gamma(\theta) = \theta$ to be the solution to the teacher's problem. As in Rayo (2013), pooling may be, nevertheless, optimal on the interior of the type space. The decision to pool the types in a given interval depends on the joint outcome of the three effects of pooling. The convexity condition ensures that the sum of the first two effects is negative for pooling, but the third effect is ambiguous and depends on the distribution $F(.)$. The next proposition states a necessary condition for the third effect to be negative, which together with the convexity condition determines a sufficient condition for the optimality of full separation.

Proposition 3 *If (43) holds and for any interval* $[\theta_1, \theta_2] \subseteq [\theta_L, \theta_H]$ *we have*

$$\frac{\int_{\theta_1}^{\theta_2} \theta dF(\theta)}{F(\theta_2) - F(\theta_1)} \geq \frac{\int_{\theta_1}^{\theta_2} \theta d\theta}{\theta_2 - \theta_1}, \tag{44}$$

then $\gamma(\theta) = \theta$ *is the solution to (41).*

Literally, condition (44) states that the distribution $F(.)$ stochastically dominates the uniform distribution in the likelihood ratio order.[10] Economically, this condition requires that the sum of information rents obtained by the types pooled in $[\theta_1, \theta_2]$ be larger than the information rents obtained under separation. This leads to the third effect, i.e., changes in allocations for supramarginal types, being negative for pooling as compared to separation. For the numerical examples studied above, conditions (43) and (44) both hold implying the optimality of perfect screening with observable grading rules.

All in all, the observability of grading rules makes an enormous impact on the form of the optimal grading rule. The compression of grades among the most efficient types

[10] The uniform distribution arises as a consequence of the assumption about the functional form of the cost function: $C(x)/\theta$. With a more general specification, $C(x)/v(\theta)$ with $v(.) > 0$ and $v'(.) > 0$, the condition equivalent to (44) is

$$\frac{\int_{\theta_1}^{\theta_2} \theta dF(\theta)}{F(\theta_2) - F(\theta_1)} \geq \frac{\int_{\theta_1}^{\theta_2} \theta dv(\theta)}{v(\theta_2) - v(\theta_1)}.$$

no longer holds with observability unless the cost function is very steep. In general, the resultant interdependence of student utilities induces the teacher to differentiate student types at a greater rate than in the case with unobservable grading rules. Gains from lenient grading in the form of more effort extraction from lower types do not outweigh losses arising from the degrading of grade values by the job market.

6 Discussion

First, we demonstrate that the findings of our models with unobservable grading rules closely match the ubiquitously observed grading patterns of 1) grade compression, 2) grade inflation, and 3) mismatch between students' grades and their abilities. We also relate our findings to the discretization of grades and the reduction of study time. Second, we discuss policy applications.

6.1 Main findings

Grade compression

Grade compression stands for the practice of coarse grading when teachers tend to rely on a few marks to assess students' performance. Our model with unobservable grading rules offers a rational explanation for this practice. The argument is based on the costless nature of grades as incentives. First, the teacher finds it average knowledge maximizing to compress the most efficient types for the highest reward as argued when deriving Result 1. Then, perfect screening is never optimal in the long run as we show by Result 3. With unobservable grading rules, teachers face the commitment problem of not exploiting the costlessness of grades even when grading standards are subsequently adjusted. Eventually, teachers are bound to expand the range of ability types that are subject to better grades because of the diminished incentive effects of lower grades. This argument can also be taken as an explanation for the usage of discrete grading scales.

Compressed rewards pertain not only to the teacher-student relationship. Consider, for instance, the well-documented compression of ratings phenomenon at job performance appraisals (see Murphy and Cleveland 1995; Prendergast 1999). Our static model with costless rewards can also be related to this phenomenon. In firms practicing job performance appraisal systems, a line manager rates the performance of her employees, whose pay is determined by their ratings received. However, it is not a line manager that bears the resultant payroll cost but the firm (or rather its residual claimants). Hence, if a line manager has an interest in eliciting more effort from her employee then they may find themselves in an agency relationship that features costless rewards. Along the same lines of our argument for grade compression, the compression of ratings can arise in the line manager's optimum.

Grade inflation

The model with unobservable grading rules also addresses grade inflation, an increasing frequency of high grades without an accompanying rise in student ability or effort.[11] From our analysis we can distinguish two contributing factors toward the increasing degree of leniency in grading. The first factor pertains to shifts in the distribution of student abilities toward the lower end of the ability space (Result 2). The second factor is the aggravation of teachers' commitment problem (Result 3).

Both factors can originate from the same source, namely, the expanding availability of

[11]See, e.g., Sabot and Wakeman-Linn (1991); Kuh and Hu (1999); Johnson (2003).

education.[12] In recent decades, due to an increasing number of educational institutions and study programs, a larger number of study places has been offered with the implication that a larger number of low-ability applicants is enrolled.[13] Subsequently, this can lead to the emergence of the first factor discussed. Similarly, with more issuers of educational certificates, individual grading rules carry a lesser weight on the perception of the job market about grading standards applied, thus, resulting in the second factor – the commitment problem.

Finally, the numerical analysis of the dynamic model with unobservable grading rules shows a reduction in students' test scores and, accordingly, learning effort levels over time. This result is consistent with the empirical study of Babcock and Marks (2011), where they report that over the period 1961–2003 time spent for study by full-time students in the US gradually decreased from 40 to 27 hours.

Mismatch between grades and abilities

The model with unobservable grading rules is also consistent with the empirical observation that teachers of classes with less able students apply more lenient grading rules and *vice versa*, see Goldman and Widawski (1976); Strenta and Elliott (1987); Johnson (2003).[14] Drawing on Result 2, this observation can be the outcome of the optimal design

[12]See footnote 5 for empirical evidence and McKenzie and Tullock (1981, Ch. 17) for other discussion.

[13]Wilson (1999) reports declining college entrance test scores. There is also a steadily increasing number of students enrolled in US degree-granting institutions both in absolute and relative (as a percentage of high school graduates) numbers: see Table 200 of 2009 Digest of Education Statistics, http://nces.ed.gov/programs/digest/d09/tables/dt09_188.asp.

[14]Goldman and Widawski (1976) report a negative correlation between students' Scholastic Aptitude Test scores (a proxy measure of students' abilities) and the grading standards in the classes the students were majoring in. According to this study (conducted at the University of California, Riverside), the

of grading rules and not necessarily the outcome of teachers' rent-seeking behavior, as sometimes is suggested (e.g., Johnson 1997). Our model predicts that in weak classes the optimal grading rule makes high grades more easily attainable in order to elicit more effort from more numerous lower-ability students. This, correspondingly, results in a mismatch and low correlation between students' grades and their abilities.

Concerning the normative side of the differential grading standards discussed, there is a number of papers proposing grade adjustment mechanisms (see, e.g., Johnson 1997) aimed at making grades more informative of students' actual abilities. Without going into the details of this literature, it is worth noting that it is typically assumed that the reason for differential grading standards lies with some personal features of the instructor (e.g., the adaptation level, unwillingness to spend office hours on dealing with students' complaints about low grades, etc.). Therefore, proposed grade adjustment mechanisms would attempt to correct for presumed instructor-specific factors failing to recognize the possible endogeneity of those factors, which could lead mechanisms astray from the projected goals.

Interestingly, Result 2 also points to the presence of indirect peer effects. A student's allocation (grade and score) depends on the teacher's grading rule applied. At the same time, the form of the grading rule depends on the student's peers or rather their abilities. Hence, peers can also matter indirectly – through the grading rule designed by the teacher. An interesting research question would be the disentanglement of direct (typically studied) and indirect (suggested here) peer effects.

---

negative correlation observed is the consequence of that professors in fields with more high-ability students tend to grade more stringently. These empirical findings were confirmed by similar studies conducted at Dartmouth College (Strenta and Elliott 1987) and at Duke University (Johnson 2003).

6.2 Policy applications

Merit-pay programs for high school teachers have been recently introduced in different countries to foster teachers' incentives to elicit more effort from students, see Lavy (2002, 2009); Atkinson, et al. (2004); Lazear (2003). Typically, these programs offer monetary bonuses to teachers if their students improve upon their previous performance as measured by their scores on standardized tests. The goals pursued by the developers of such programs range from improving average performance (in most cases) to reducing the gap between bad and good performers as, e.g., in the "No Child Left Behind" initiative in the US.

In terms of our modeling framework, incentives set by merit-pay programs can have a direct effect on the form of the teacher's utility function $V(.)$. In the case of "No Child Left Behind" the utility function $V(.)$ can turn concave because the teacher would start putting relatively more weight on the performance of lower-ability students. Then, with a concave utility function $V(.)$ the model with costless grades predicts that the gap between low- and high-ability students would diminish as compared to the case of a linear utility. However, this reduction would come from two directions: more effort from low-ability students and less effort from those with high abilities.[15] Hence, a negative externality

[15] Denote the optimal score schedule of the solution to the teacher's problem (5) with a strictly concave utility function $V(.)$ by $\{x^V(\theta)\}$. With a reference to Proposition 1 and its proof, the optimal score allocations $x^V(\theta)$ for θ in $[\theta_L, \theta^*)$ are equal to

$$x^V(\theta) = C'^{-1}\left(\frac{f(\theta)\theta^2 V'(x^V(\theta))C'(x^V(\theta^*))}{f(\theta^*)\theta^{*2}V'(x^V(\theta^*))}\right).$$

Using the revealed preference argument and straightforward analytical manipulations, we can show that $x^V(\theta^*) < x(\theta^*)$ and $x^V(\theta_L) > x(\theta_L)$, where $x(\theta^*)$ and $x(\theta_L)$ are the optimal allocations from the linear case. Hence, with a strictly concave utility $V(.)$, the range of scores contracts. And this contraction comes from both directions: optimal allocations become lower for able students and higher for less able can arise from the introduction of such programs. In contrast, the gap between bad and good performers can increase if the teacher's utility function turns convex – as a result, e.g., of a merit-pay program that rewards for excellence.

With the help of our models, we can also address the problem whether university administration should restrict teachers' choice of grading rules by imposing relative grading such as grading on a curve. In the case of observable grading rules, grading on a curve is superfluous if not harmful when there is a homogenous group of students; a similar prediction is also made by Dubey and Geanakoplos (2010). However, under a more realistic scenario of unobservable grading rules, grading on a curve could possibly fix teachers' commitment problem. But then, the question is what goal university administration pursues. If it is the maximization of students' knowledge, then grading on a curve or the disclosure of grade distributions would be a desirable policy. But if administration aims for a higher average job placement, then, as argued by Chan, et al. (2007) and Ostrovsky and Schwarz (2010), compressed grades are a sought-after outcome. In this case, administration would not provide incentives for teachers to differentiate among their students, thus, refraining from grading on a curve and disclosing any information about class grade distribution.[16]

7 Conclusion

In this paper, we approach the teacher-student relationship from the perspectives of ability screening and knowledge maximization. We show that many empirical grading

students.

[16]For more discussion and empirical implications of making academic transcripts more informative, see Bar, et al. (2009).

patterns such as grade compression, grade inflation, or mismatch between students' grades and abilities are consistent with the predictions of our agency model that has grades as costless incentives. When the job market makes a judgment about students' grades based on a standard other than the actual grading rule applied, a teacher chooses a lenient grading rule to induce more effort from less able students, thus, yielding coarse grading. This effect turns more important in less able classes, and *vice versa*. When the job market adaptively adjusts its perception of grading standards, the practice of coarse grading can unravel over time until the point when no type is screened. In their attempts to motivate students, teachers are bound to increase leniency in grading because of depreciating values of grades arising from the commitment problem, which has ultimately implications for grade inflation. Finally, we suggest that the modeling framework presented here can serve as microfoundations for the purpose of policy applications.

Department of Economics, University of Zurich, Switzerland

Appendix

Proof of Proposition 1

Omitting monotonicity constraint (13) and the constraint $\widehat{\varphi}(\theta) \geq \theta_L$, the Hamiltonian of the reduced problem is

$$H(\widehat{\varphi}, U, \lambda, \mu, \theta) = C^{-1}\left(\theta\widehat{\varphi} - U\right) f(\theta) + \lambda\widehat{\varphi} + \mu(\theta_H - \widehat{\varphi}), \tag{45}$$

where $\lambda(\theta)$ is the co-state variable associated with (12) and $\mu(\theta)$ is the Lagrange multiplier for $\widehat{\varphi}(\theta) \leq \theta_H$. The necessary conditions for the optimal control problem, which are also sufficient because $C^{-1}(.)$ is concave, imply that

$$\frac{\theta f(\theta)}{C'(x(\theta))} + \lambda(\theta) - \mu(\theta) = 0, \tag{46}$$

$$\frac{f(\theta)}{C'(x(\theta))} = \lambda'(\theta), \tag{47}$$

$$U'(\theta) = \widehat{\varphi}(\theta), \tag{48}$$

$$\lambda(\theta_H) = 0, \tag{49}$$

where $x(\theta) = C^{-1}(\theta\widehat{\varphi}(\theta) - U(\theta))$. The condition (47) follows from Pontryagin's maximum principle and (49) is the transversality condition.

From (46) and (49), we get $\mu(\theta_H) > 0$ implying that $\widehat{\varphi}(\theta_H) = \theta_H$. The state variable $U(\theta)$, which is the student's net utility, is strictly increasing in θ and also continuous as (upward) jumps cannot be optimal for the teacher (she wants to keep the net utility as low as possible). Then, by the continuity of $U(\theta)$ and the assumed piecewise continuity of the control variable, the co-state variable $\lambda(\theta)$ is continuous. This implies that the multiplier $\mu(\theta)$ remains positive for an interval of types that are pooled for the highest reward. Denote this interval by $(\theta^*, \theta_H]$ and the corresponding uniform allocation by (x^*, θ_H). From (47) and (49), we obtain that for $\theta > \theta^*$

$$\lambda(\theta) = -\frac{1 - F(\theta)}{C'(x^*)}, \tag{50}$$

which renders (46) to satisfy

$$\frac{\theta f(\theta)}{C'(x^*)} - \frac{1 - F(\theta)}{C'(x^*)} > 0. \tag{51}$$

The last expression implies that the starting point of the pooling interval θ^* is given by

$$\theta^* = \min\{\theta : \theta f(\theta) - (1 - F(\theta)) \geq 0, \theta \in [\theta_L, \theta_H]\}. \tag{52}$$

As a consequence, we can establish that $\mu(\theta) = 0$ for $\theta \leq \theta^*$.

Suppose that $\theta^* > \theta_L$ holds (if $\theta^* = \theta_L$, all types are pooled). For $\theta \leq \theta^*$ (46) becomes

$$\frac{\theta f(\theta)}{C'(x(\theta))} - \left(\int_{\theta}^{\theta^*} \frac{f(\tilde{\theta})}{C'(x(\tilde{\theta}))} d\tilde{\theta} - \frac{1 - F(\theta^*)}{C'(x^*)} \right) = 0, \tag{53}$$

where the expression in the brackets stands for $\lambda(\theta)$ obtained from (47). Differentiating with respect to θ renders

$$2\frac{f(\theta)}{C'(x(\theta))} + \theta \frac{d}{d\theta} \left(\frac{f(\theta)}{C'(x(\theta))} \right) = 0. \tag{54}$$

Solving for $x(\theta)$, $\theta \in [\theta_L, \theta^*]$, we get

$$x(\theta) = C'^{-1} \left(f(\theta)\theta^2 \frac{C'(x(\theta^*))}{f(\theta^*)\theta^{*2}} \right). \tag{55}$$

To find the allocation $x(\theta^*)$, consider (46) at $\theta = \theta^*$

$$\lambda(\theta^*) = -\frac{\theta^* f(\theta^*)}{C'(x(\theta^*))}. \tag{56}$$

From (50) the right-hand limit of $\lambda(\theta)$ is

$$\lim_{\theta \to \theta^*} \lambda(\theta) = -\frac{1 - F(\theta^*)}{C'(x^*)}. \tag{57}$$

By the continuity of the co-state variable we obtain that $x(\theta^*) = x^*$, which implies that the optimal score schedule is continuous.

The optimal signal allocations $\widehat{\varphi}(\theta)$ for $\theta \in [\theta_L, \theta^*)$ are most easily determined from the integral form envelope condition applied to initial incentive compatibility constraint (2):

$$\widehat{\varphi}(\theta) = \theta_L + \frac{C(x(\theta))}{\theta} + \int_{\theta_L}^{\theta} \frac{C(x(\tilde{\theta}))}{\tilde{\theta}^2} d\tilde{\theta}. \tag{58}$$

Given $\widehat{\varphi}(\theta^*) = \theta_H$, we can from (58) at $\theta = \theta^*$ integrate out the score allocation x^*.

Next, we establish that the omitted monotonicity constraint holds. For this purpose it suffices to show that the score schedule $x(\theta)$ in (55) is non-decreasing. Because $C'(.)$

is increasing, so is its inverse $C'^{-1}(.)$. Hence, for $x(\theta)$ to be non-decreasing we need that the argument of $C'^{-1}(.)$ in (55) is non-decreasing in θ, i.e., $d\left[f(\theta)\theta^2\right]/d\theta \geq 0$. From (52) it follows that $f(\theta)\theta/(1-F(\theta)) < 1$ for θ in $[\underline{\theta}, \theta^*)$, and from the monotone hazard rate: $f'(\theta) > -f^2(\theta)/(1-F(\theta))$. The two properties together ensure that $d\left[f(\theta)\theta^2\right]/d\theta \geq 0$ holds, implying that $x'(\theta) \geq 0$. The omitted constraint $\widehat{\varphi}(\theta) \geq \theta_L$ holds from (58).■

Proof of Result 2

Let θ_i^* denote the starting point of the pooling-at-the-top interval in class i, $i = A, B$, defined by (16) in Proposition 1. If $\theta_B^* = \theta_L$, then the proposition holds trivially. Suppose $\theta_B^* > \theta_L$. Because the likelihood-ratio stochastic order implies the hazard-rate order (see Shaked and Shanthikumar 1994), which is $f_A(\theta)/(1-F_A(\theta)) \leq f_B(\theta)(1-F_B(\theta))$ for every θ, it immediately follows that $\theta_A^* \geq \theta_B^*$. In the less able class B, there is more pooling at the top. Thus, we have $\widehat{\varphi}_B(\theta) = \theta_H \geq \widehat{\varphi}_A(\theta)$ for $\theta \in [\theta_B^*, \theta_H]$.

Consider the optimal score schedules $\{x_A(\theta)\}$ and $\{x_B(\theta)\}$ and suppose that they are not identical (otherwise, the proposition holds trivially). Because the schedules are continuous, they must cross at least once. (Otherwise, the lower score schedule would not be optimal.) Let $\underline{\theta}$ be the left-most crossing point. We separately consider two cases: 1) $\underline{\theta} \geq \theta_B^*$, and 2) $\underline{\theta} < \theta_B^*$.

In case 1, $\underline{\theta} \geq \theta_B^*$, it must be that either $x_B(\theta) > x_A(\theta)$ or $x_B(\theta) < x_A(\theta)$ for $\theta < \underline{\theta}$. The latter is impossible because then we never have $x_B(\theta) > x_A(\theta)$ required for the optimality of $\{x_B(\theta)\}$ as $x_B(\theta)$ stays constant for $\theta \geq \underline{\theta}$ while $x_A(\theta)$ is non-decreasing. Thus, it must be that $x_B(\theta) > x_A(\theta)$ for $\theta < \underline{\theta}$. From (15) in Proposition 1 it follows that $\widehat{\varphi}_A(\theta) < \widehat{\varphi}_B(\theta)$ for $\theta < \theta_B^* \leq \underline{\theta}$, proving the proposition for case 1.

In case 2, $\underline{\theta} < \theta_B^*$, by (14) in Proposition 1 we have

$$C'^{-1}\left(f_A(\underline{\theta})\underline{\theta}^2 M_A\right) = C'^{-1}\left(f_B(\underline{\theta})\underline{\theta}^2 M_B\right), \tag{59}$$

where $M_i = C'(x_i^*)/(f_i(\theta_i^*)\theta_i^{*2})$, $i = A, B$. It yields

$$\frac{f_A(\underline{\theta})}{f_B(\underline{\theta})} = \frac{M_B}{M_A}. \tag{60}$$

Due to the assumed likelihood-ratio stochastic order, we get $f_A(\theta)/f_B(\theta) \geq M_B/M_A$ for all $\theta > \underline{\theta}$ implying that $x_A(\theta) \geq x_B(\theta)$ for all $\theta \geq \underline{\theta}$ and $x_A(\theta) < x_B(\theta)$ for all $\theta < \underline{\theta}$. Hence, by (15) in Proposition 1, it follows that $\widehat{\varphi}_A(\theta) < \widehat{\varphi}_B(\theta)$ for $\theta \leq \underline{\theta}$.

What remains to be proved is $\widehat{\varphi}_A(\theta) \leq \widehat{\varphi}_B(\theta)$ for $\theta \in (\underline{\theta}, \theta_B^*)$. To invoke a contradiction, suppose that for some $\widetilde{\theta} \in (\underline{\theta}, \theta_B^*)$ we have $\widehat{\varphi}_A(\widetilde{\theta}) > \widehat{\varphi}_B(\widetilde{\theta})$. Because $x_A(\theta) \geq x_B(\theta)$ for $\theta \geq \tilde{\theta} > \underline{\theta}$, then due to (15) in Proposition 1 we must have $\widehat{\varphi}_A(\theta) > \widehat{\varphi}_B(\theta)$ for all $\theta \geq \tilde{\theta}$. But it is a contradiction because $\widehat{\varphi}_A(\theta) \leq \widehat{\varphi}_B(\theta)$ for $\theta \in [\theta_B^*, \theta_H]$ as shown above.■

Proof of Result 3

First, we show that a consistent grading standard $\varphi_t(.)$ with (some) perfect screening of types is impossible (Part (i)). To invoke a contradiction, suppose there exists such a consistent grading standard $\varphi_t(.)$ that screens the types in an interval $[\theta_L, \theta_{t,1}]$ and pools types in intervals $(\theta_{t,1}, \theta_{t,2}]$, $(\theta_{t,2}, \theta_{t,3}]$, ..., $(\theta_{t,n}, \theta_H]$. (There must be some pooling at the top as we have previously shown.) The closure of its range is $\overline{\mathcal{R}(\varphi_t)} = [\theta_L, \theta_{t,1}] \cup \overline{\theta_{t,1}} \cup ... \cup \overline{\theta_{t,n}}$, where $\overline{\theta_{t,i}} = E(\theta \in (\theta_{t,i}, \theta_{t,i+1}])$, $i = 1, ..., n$ and $\theta_{t,n+1} = \theta_H$. Setting the starting point of the screening interval at a point other than θ_L has no impact on the argument that follows.

Let $x_t(\theta)$ denote the score allocation for θ in $[\theta_L, \theta_{t,1}]$ and $\overline{x}_{t,i}$ be the uniform score allocation for types in $(\theta_{t,i}, \theta_{t,i+1}]$, $i = 1, ..., n$. As before, the score allocations are determined by the integral form envelope theorem given signal allocations (as in (20) and (21)). The teacher's maximization problem can be expressed by the Lagrangian

$$(61)\ \mathcal{L}(\widehat{\varphi}_t(\theta), \{\widehat{\theta}_{t,i}\}_{i=1}^n, \lambda(\theta)) = \int_{\theta_L}^{\widehat{\theta}_{t,1}} x_t(\theta) f(\theta) d\theta + \lambda(\theta)(\theta_{t,1} - \widehat{\varphi}_t(\theta)) + + \sum_{i=1}^n \overline{x}_{t,i}(F(\widehat{\theta}_{t,i}) - F(\widehat{\theta}_{t,i+1})) + \sum_{i=1}^n \mu_i(\widehat{\theta}_{t,i+1} - \widehat{\theta}_{t,i}),$$

which has to be maximized with respect to the starting points of pooling intervals, $\{\widehat{\theta}_{t,i}\}_{i=1}^n$, signal allocations $\widehat{\varphi}_t(\theta)$ for $\theta \leq \widehat{\theta}_{t,1}$. As before, $\lambda(\theta)$ is the Lagrange multiplier on the constraint $\theta_{t,1} \geq \widehat{\varphi}_t(\theta)$ for $\theta \leq \widehat{\theta}_{t,1}$; μ_i is the Lagrange multiplier on the constraint $\theta_{t,i+1} \leq \theta_{t,i}$, $i = 1, ..., n$ (i.e., pooling intervals must not overlap).

Because $\varphi_t(.)$ is consistent, $\{\theta_{t,i}\}_{i=1}^n$ and $\widehat{\varphi}_t(\theta)$ need to satisfy the first-order conditions, which with respect to $\widehat{\theta}_{t,1}$ and $\widehat{\varphi}_t(\theta)$ are, respectively,

$$(62)\quad \frac{\overline{\theta_{t,1}} - \theta_{t,1}}{f(\theta_{t,1})} \sum_{i=1}^n \frac{F(\theta_{t,i+1}) - F(\theta_{t,i})}{C'(\overline{x}_{t,i})} = \overline{x}_{t,1} - x_t(\theta_{t,1}),$$

$$(63)\quad \frac{\theta f(\theta)}{C'(x_t(\theta))} - \int_{\theta}^{\theta_{t,1}} \frac{f(\tilde{\theta})}{C'\left(x_t(\tilde{\theta})\right)} d\tilde{\theta} - \sum_{i=1}^n \frac{F(\theta_{t,i+1}) - F(\theta_{t,i})}{C'(\overline{x}_{t,i})} = \lambda(\theta).$$

The summation term appears in these conditions because $\widehat{\theta}_{t,1}$ and $\widehat{\varphi}_t(\theta)$ enter the expressions for the optimal score allocations for all $\theta \geq \theta_{t,1}$. As $\theta_{t,1} < \theta_{t,2}$ by construction, it must be that $\mu_1 = 0$. Based on the same convexity argument applied for the case of $t = 2$ when deriving (24), we can show that

$$(64)\quad \overline{x}_{t,1} - x_t(\theta_{t,1}) < \frac{\theta_{t,1}\left(\overline{\theta_{t,1}} - \theta_{t,1}\right)}{C'(x_t(\theta_{t,1}))}.$$

Then, the summation term in (62) is strictly smaller than $\theta_{t,1} f(\theta_{t,1})/C'(x_t(\theta_{t,1}))$, which implies that the left-hand side of (63) is strictly positive for some $\theta < \theta_{t,1}$. As a result, the teacher cannot find it optimal to screen all the types in $[\theta_L, \theta_{t,1}]$, which contradicts that the grading standard $\varphi_t(.)$ is consistent.

To prove Part (ii), consider the first-order condition with respect to $\theta_{t,n}$, the starting point of the pooling-at-the-top interval,

$$\frac{\overline{\theta_{t,n}} - \theta_{t,n}}{f(\theta_{t,n})} \frac{1 - F(\theta_{t,n})}{C'(\overline{x}_{t,n})} = \overline{x}_{t,n} - \overline{x}_{t,n-1}. \tag{65}$$

We have that the difference

$$\overline{x}_{t,n} - \overline{x}_{t,n-1} < \frac{\theta_{t,n}\left(\overline{\theta_{t,n}} - \theta_{t,n}\right)}{C'\left(\overline{x}_{t,n}\right)}, \tag{66}$$

which from (65) implies $(1 - F(\theta_{t,n}))/(f(\theta_{t,n})\theta_{t,n}) < 1$. Thus, $\theta_{t,n} < \theta_1^*$.■

Proof of Proposition 2

Suppose that the score schedule in (39) with $\gamma(\theta) = \theta$ is convex over some non-empty interval $[\underline{\theta}, \theta_H]$. Let $\{\gamma^*(\theta)\}$ solve (41) and $\{x^*(\theta)\}$ be the optimal score schedule. Suppose that $\{x^*(\theta)\}$ pools an interval $[\theta', \theta_H]$, $\theta' \in [\underline{\theta}, \theta_H)$, for the score $\overline{x^*}$ equal to

$$\overline{x^*} = C^{-1}\left(\theta'\overline{\theta} - \int_{\theta_L}^{\theta'} \gamma^*(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right), \tag{67}$$

where $\overline{\theta} = \int_{\theta'}^{\theta_H} \theta dF(\theta)/(1 - F(\theta'))$. Now consider the following signal schedule $\{\widehat{\gamma}(\theta)\}$: $\widehat{\gamma}(\theta) = \gamma^*(\theta)$ for $\theta < \theta'$ and $\widehat{\gamma}(\theta) = \theta$ for θ in $[\theta', \theta_H]$. Denote the score schedule defined by (39) with $\{\widehat{\gamma}(\theta)\}$ by $\{\widehat{x}(\theta)\}$.

Noting that $\widehat{x}(\theta) = x^*(\theta)$ for $\theta < \theta'$ the difference in the teacher's expected utility levels from implementing $\{\widehat{x}(\theta)\}$ and $\{x^*(\theta)\}$ is equal to

$$\int_{\theta'}^{\theta_H} \widehat{x}(\theta)dF(\theta) - \overline{x^*}(1 - F(\theta')). \tag{68}$$

By the hypothesis the score schedule $\{\widehat{x}(\theta)\}$ is convex over $[\theta', \theta_H]$, thus, by the Jensen's inequality we have

$$
\begin{aligned}
(69)\int_{\theta'}^{\theta_H} \widehat{x}(\theta)dF(\theta) &= \int_{\theta'}^{\theta_H} C^{-1}\left(\theta\widehat{\gamma}(\theta) - \int_{\theta_L}^{\theta}\widehat{\gamma}(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right)dF(\theta) \\
(70) \qquad &\geq C^{-1}\left(\overline{\theta}\widehat{\gamma}(\overline{\theta}) - \int_{\theta_L}^{\overline{\theta}}\widehat{\gamma}(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right)(1-F(\theta')) \\
(71) \qquad &= C^{-1}\left(\overline{\theta}\widehat{\gamma}(\overline{\theta}) - \int_{\theta_L}^{\theta'}\gamma^*(\tilde{\theta})d\tilde{\theta} - \int_{\theta'}^{\overline{\theta}}\widehat{\gamma}(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right)(1-F(\theta')) \\
(72) \qquad &> C^{-1}\left(\overline{\theta}\widehat{\gamma}(\overline{\theta}) - \int_{\theta_L}^{\theta'}\gamma^*(\tilde{\theta})d\tilde{\theta} - \widehat{\gamma}(\overline{\theta})(\overline{\theta}-\theta') - \theta_L^2\right)(1-F(\theta')) \\
(73) \qquad &= C^{-1}\left(\theta'\widehat{\gamma}(\overline{\theta}) - \int_{\theta_L}^{\theta'}\gamma^*(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right)(1-F(\theta')) \\
(74) \qquad &= \overline{x^*}(1-F(\theta')),
\end{aligned}
$$

where the last inequality follows from $\widehat{\gamma}(.)$ and $C^{-1}(.)$ being increasing. Thus, the difference in (68) is positive, which implies that pooling at the top cannot be optimal if the score schedule under separation is convex.■

Proof of Proposition 3

Let $\{\gamma^*(\theta)\}$ solve (41) and $\{x^*(\theta)\}$ be the optimal score schedule. Suppose that $\{x^*(\theta)\}$ contains pooling intervals and denote the right-most pooling interval by $[\theta_1, \theta_2]$. By Proposition 2 it must be that $\theta_2 < \theta_H$. Now consider the following signal schedule $\{\widehat{\gamma}(\theta)\}$: $\widehat{\gamma}(\theta) = \gamma^*(\theta)$ for $\theta < \theta_1$ and $\widehat{\gamma}(\theta) = \theta$ for θ in $[\theta_1, \theta_H]$. Denote the score schedule defined by (39) with $\{\widehat{\gamma}(\theta)\}$ by $\{\widehat{x}(\theta)\}$.

The difference in the teacher's expected utility levels from implementing the score schedules $\{\widehat{x}(\theta)\}$ and $\{x^*(\theta)\}$ is equal to

$$
(75) \qquad \int_{\theta_1}^{\theta_2}(\widehat{x}(\theta) - x^*(\theta))\,dF(\theta) + \int_{\theta_2}^{\theta_H}(\widehat{x}(\theta) - x^*(\theta))\,dF(\theta).
$$

The first term of the difference in (75) is positive, which follows from the Jensen's inequality applied to the score schedule $\{\widehat{x}(\theta)\}$, which is convex over $[\theta_1, \theta_2]$. For $\theta \in (\theta_2, \theta_H]$, the score allocations $\widehat{x}(\theta)$ and $x^*(\theta)$ are, respectively, given by

$$\widehat{x}(\theta) = C^{-1}\left(\theta\widehat{\gamma}(\theta) - \int_{\theta_L}^{\theta}\widehat{\gamma}(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right) \tag{76}$$

$$= C^{-1}\left(\theta\widehat{\gamma}(\theta) - \int_{\theta_L}^{\theta_1}\gamma^*(\tilde{\theta})d\tilde{\theta} - \int_{\theta_1}^{\theta_2}\widehat{\gamma}(\tilde{\theta})d\tilde{\theta} - \int_{\theta_2}^{\theta}\widehat{\gamma}(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right) \tag{77}$$

and

$$x^*(\theta) = C^{-1}\left(\theta\gamma^*(\theta) - \int_{\theta_L}^{\theta_1}\gamma^*(\tilde{\theta})d\tilde{\theta} - \int_{\theta_1}^{\theta_2}\gamma^*(\tilde{\theta})d\tilde{\theta} - \int_{\theta_2}^{\theta}\gamma^*(\tilde{\theta})d\tilde{\theta} - \theta_L^2\right). \tag{78}$$

Since $\widehat{\gamma}(\theta) = \gamma^*(\theta)$ for $\theta \in [\theta_L, \theta_1) \cup (\theta_2, \theta_H]$, we get that $\widehat{x}(\theta) \geq x^*(\theta)$ if and only if

$$\int_{\theta_1}^{\theta_2}\widehat{\gamma}(\tilde{\theta})d\tilde{\theta} \leq \int_{\theta_1}^{\theta_2}\gamma^*(\tilde{\theta})d\tilde{\theta}. \tag{79}$$

As $\gamma^*(\theta) = E(\theta \in [\theta_1, \theta_2]) = \int_{\theta_1}^{\theta_2}\theta dF(\theta)/(F(\theta_2) - F(\theta_1))$ for θ in $[\theta_1, \theta_2]$, the latter condition becomes

$$\frac{\int_{\theta_1}^{\theta_2}\theta d\theta}{\theta_2 - \theta_1} \leq \frac{\int_{\theta_1}^{\theta_2}\theta dF(\theta)}{F(\theta_2) - F(\theta_1)}, \tag{80}$$

which is (44) in the proposition. Thus, when (44) holds together with convexity condition (43), the difference in (75) is positive. This implies that the optimal score schedule $\{x^*(\theta)\}$ cannot pool the types in $[\theta_1, \theta_2]$ nor, analogously, in any other interval.■

References

Arcidiacono, P., "Ability sorting and the returns to college major," Journal of Econometrics 121 (2004), 343-375.

Atkinson, A., S. Burgess, B. Croxson, P. Gregg, C. Propper, H. Slater and D. Wilson, "Evaluating the Impact of Performance-related Pay for Teachers in England," CMPO working paper No. 04/113, 2004.

Babcock, P. S., and M. Marks, "The Falling Time Cost of College: Evidence from Half a Century of Time Use Data," Review of Economics and Statistics 93 (2011), 468-478.

Bar, T., V. Kadiyali and A. Zussman, "Grade Information and Grade Ination: The Cornell Experiment," Journal of Economic Perspectives 23 (2009), 93-108.

Becker, G. S., Human capital: A theoretical and empirical analysis, with special reference to education, 3rd ed. (Chicago: University of Chicago Press, 1993).

Becker, W. E., and S. Rosen, "The learning effect of assessment and evaluation in high school," Economics of Education Review 11 (1992), 107-118.

Bedard, K., "Human Capital versus Signaling Models: University Access and High School Dropouts," Journal of Political Economy 109 (2001), 749-775.

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

Chan, W., L. Hao and W. Suen, "A Signaling Theory of Grade Ination," International Economic Review 48 (2007), 1065-1090.

Costrell, R. M., “A Simple Model of Educational Standards,” American Economic Review 84 (1994), 956-971.

Costrell, R. M., “Can centralized educational standards raise welfare?” Journal of Public Economics 65 (1997), 271-293.

Dubey, P., and J. Geanakoplos, “Grading Exams: 100, 99, 98,... or A, B, C?” Games and Economic Behavior 69 (2010), 72-94.

Fang, H., and G. Moscarini, “Morale hazard,” Journal of Monetary Economics 52 (2005), 749-777.

Goldman, R. D., and M. H. Widawski, “A Within-Subjects Technique for Comparing College Grading Standards: Implications in the Validity of the Evaluation of College Achievement,” Educational and Psychological Measurement 36 (1976), 355-359.

Guesnerie, R., and J.-J. Laffont, “A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm,” Journal of Public Economics 25 (1984), 329-369.

Johnson, V. E., “An Alternative to Traditional GPA for Evaluating Student Performance,” Statistical Science 12 (1997), 251-269.

Johnson, V. E., Grade Ination: A Crisis in College Education (New York: Springer-Verlag, 2003).

Kuh, G., and S. Hu, “Unraveling the Complexity of the Increase in College Grades from the Mid-1980s to the Mid-1990s,” Educational Evaluation and Policy Analysis 20 (1999), 297-320.

Lavy, V., "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement," Journal of Political Economy 110 (2002), 1286-1317.

Lavy, V., "Performance Pay and Teachers' Effort, Productivity, and Grading Ethics," American Economic Review 99 (2009), 1979-2011.

Lazear, E., "Teacher Incentives," Swedish Economic Policy Review 10 (2003), 179-203.

Levin, J., "Relational Incentive Contracts," American Economic Review 93 (2003), 835-857.

Lizzeri, A., "Information Revelation and Certication Intermediaries," RAND Journal of Economics 30 (1999), 214-231.

Mathis, J., J. McAndrews and J.-C. Rochet, "Rating the raters: Are reputation concerns powerful enough to discipline rating agencies?" Journal of Monetary Economics 56 (2009), 657-674.

McKenzie, R. B., and G. Tullock, The New World of Economics (Homewood, Illinois: Richard D. Irwin, 1981).

Milgrom, P., and I. Segal, "Envelope Theorems for Arbitrary Choice Sets," Econometrica 70 (2002), 583-601.

Moldovanu, B., A. Sela and X. Shi, "Contests for Status," Journal of Political Economy 115 (2007), 338-363.

Murphy, K. R., and J. Cleveland, Understanding Performance Appraisal: Social, Organizational, and Goal-based Perspectives, (Thousand Oaks, California: SAGE, 1995).

Mussa, M., and S. Rosen, "Monopoly and product quality," Journal of Economic Theory 18 (1978), 301-317.

Ostrovsky, M., and M. A. Schwarz, "Information Disclosure and Unraveling in Matching Markets," American Economic Journal: Microeconomics 2 (2010), 34-63.

Prendergast, C., "The Provision of Incentives in Firms," Journal of Economic Literature 37 (1999), 7-63.

Rayo, L., "Monopolistic Signal Provision," B.E. Journals in Theoretical Economics 13 (2013), 27-58.

Riley, J. G., "Testing the Educational Screening Hypothesis," Journal of Political Economy 87 (1979), S227-S252.

Sabot, R., and J. Wakeman-Linn, "Grade Ination and Course Choice," Journal of Economic Perspectives 5 (1991), 159-170.

Shaked, M., and J. G. Shanthikumar, Stochastic Orders and Their Applications (San Diego: Academic Press, Inc., 1994).

Spence, M., "Job Market Signaling," Quarterly Journal of Economics 87 (1973), 355-374.

Strenta, A. C., and R. Elliott, "Differential Grading Standards Revisited," Journal of Educational Measurement 24 (1987), 281-291.

Wilson, B. P., "The Phenomenon of Grade Ination in Higher Education," National Forum Fall 1999, 38-41.

Yang, H., and C. S. Yip, "An Economic Theory of Grade Inflation," Working paper, Department of Economics, University of Pennsylvania, 2003.

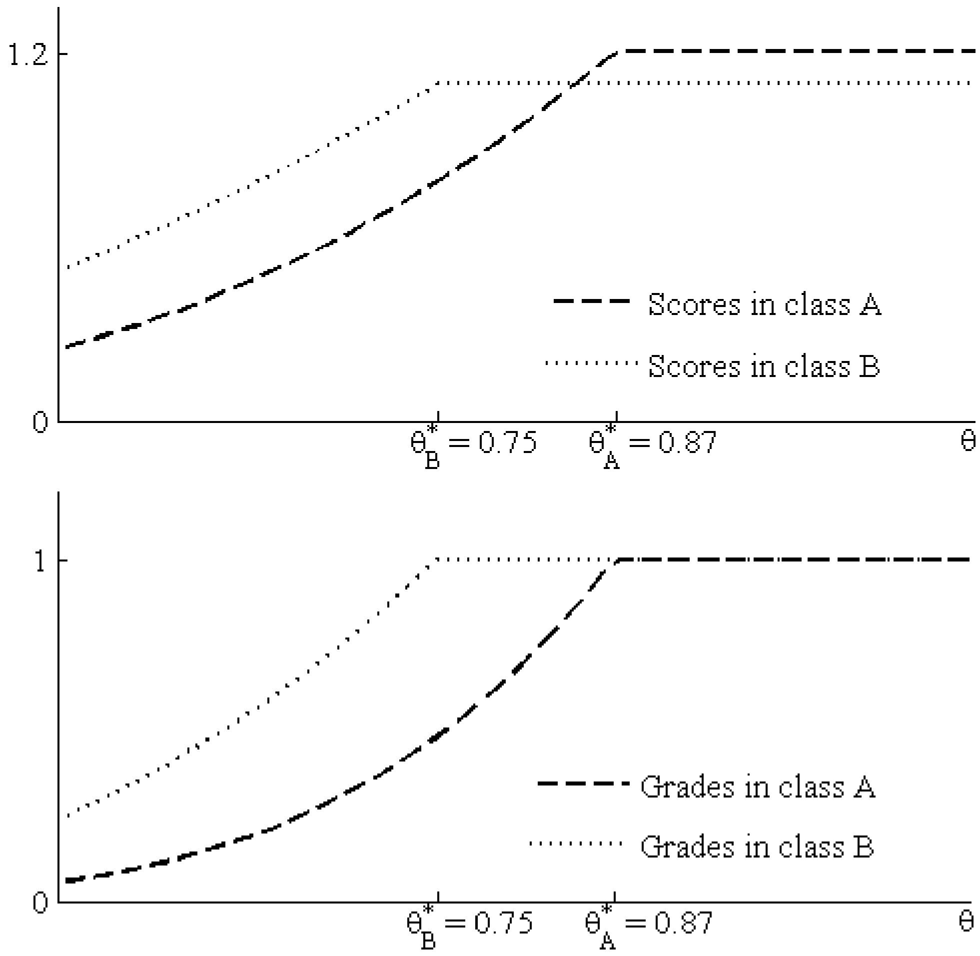

Figure 1: Optimal grading rules

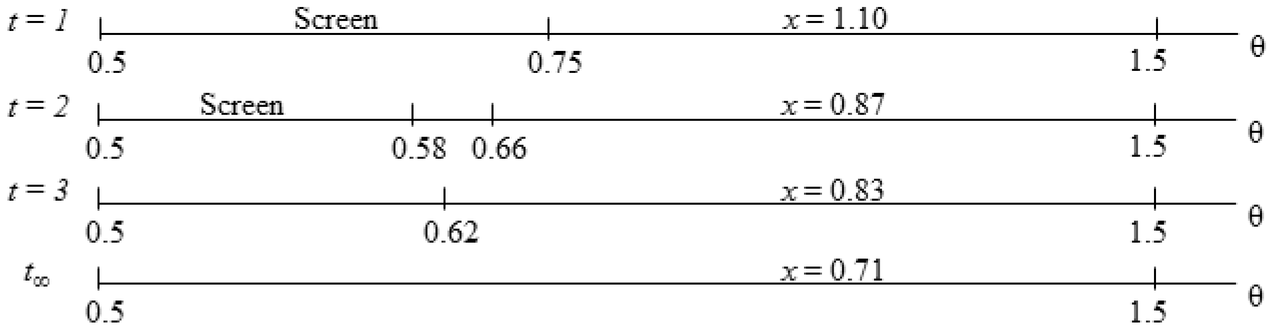

Figure 2: Dynamics of grading rules with $f(\theta) = 1$

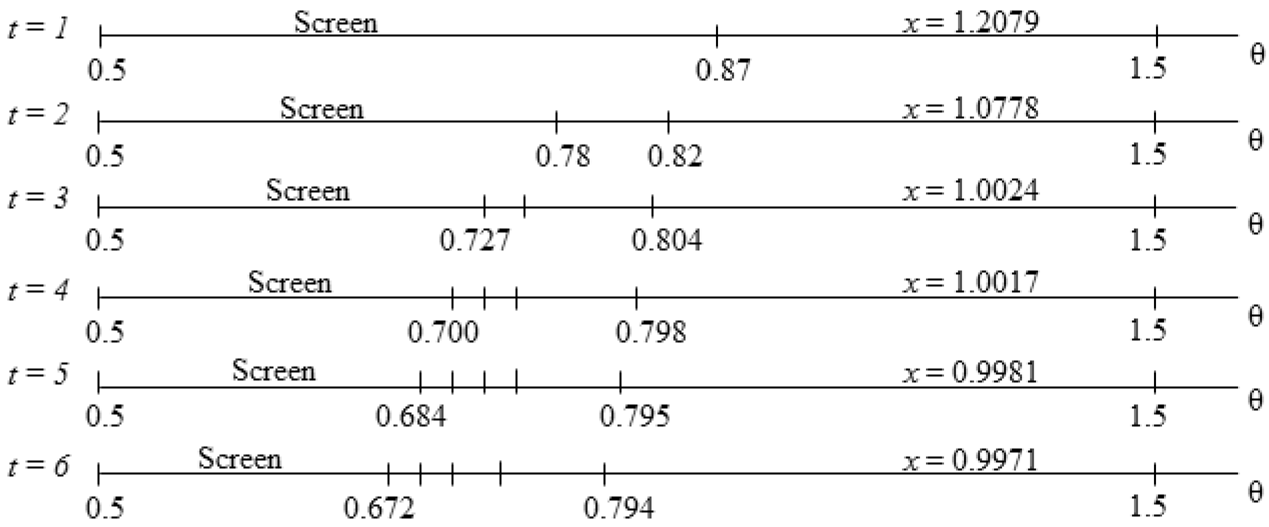

Figure 3: Dynamics of grading rules with $f(\theta) = \theta$