Submitted to *Operations Research*
manuscript OPRE-2018-05-288.R3

Economic Behavior of Information Acquisition: Impact on Peer Grading in MOOCs

Onesun Steve Yoo
UCL School of Management, University College London, 1 Canada Square, London E14 5AB, United Kingdom, onesun.yoo@ucl.ac.uk

Dongyuan Zhan
UCL School of Management, University College London, 1 Canada Square, London E14 5AB, United Kingdom, d.zhan@ucl.ac.uk

A critical issue in operating massive open online courses (MOOCs) is the scalability of providing feedback. Because it is not feasible for instructors to grade a large number of students' assignments, MOOCs employ peer grading systems. This study investigates the efficacy of that practice when student graders are rational economic agents. We characterize grading as a process of (a) acquiring information to assess an assignment's quality and (b) reporting a score. This process entails a trade-off between the cost of acquiring information and the benefits of accurate grading. Since the true quality is not observable, any measure of inaccuracy must reference the actions of other graders, which motivates student graders to behave strategically. We present the unique equilibrium information level and reporting strategy of a homogeneous group of student graders and then examine the outcome of peer grading. We show how both the peer grading structure and the nature of MOOC courses affect peer grading accuracy, and we identify conditions under which the process fails. There is a systematic grading bias toward the mean, which discourages students from learning. To improve current practice, we introduce a scale-shift grading scheme, theoretically examine how it can improve grading accuracy and adjust grading bias, and discuss how it can be practically implemented.

Key words: MOOCs, peer grading, information acquisition, equilibrium grading behavior

1. Introduction

Adoption of massive open online courses (MOOCs) has been spreading at a rapid pace, enabling a large number of students to enroll as long as they can access the course materials online. MOOCs deliver various courses (from highly quantitative to qualitative) through web-based technologies such as video presentations, online communication forums, simulation games, and computer-based assessments. Many platforms (e.g., Coursera, edX, Udacity) have demonstrated the potential to "scale" higher education, with some courses attracting more than 150,000 students (Balfour 2013, Brinton et al. 2014). As educational institutions migrate towards online education during the COVID-19 pandemic, and with increasing awareness of MOOCs, some experts predict that MOOCs will play a more prominent role in post covid era education (Young 2020).

One of the main challenges of operating MOOCs is the scalability of providing feedback.[1] Because so many students enroll in these courses, it is physically impossible for an instructor (or even a team of graders) to grade all of the assignments, and the most meaningful assignments in many courses do not lend themselves to automated grading by computer (Khalil and Ebner 2014). Hence peer grading is an essential tool for MOOC operators that need to grade complex, open-ended assignments in courses whose students number in the tens or hundreds of thousands (Piech et al. 2013).

In a typical peer review process, students are provided with a rubric that specifies the evaluation criteria to be followed when assessing the answers given to exam questions (Khalil and Ebner 2014, Luaces et al. 2015), and sometimes the reviewers are trained and then tested for competency before making any formal evaluations (Balfour 2013). The final score given to an assignment is usually determined as the average of the corresponding peer grades given by the evaluators (Balfour 2013, Luaces et al. 2015, Coursera 2018).

However, it is unclear whether peer grading systems deliver accurate assessments (Li et al. 2016). Because student graders typically lack expertise in the subject matter of an assignment, their assessment of its quality may be unreliable. Moreover, the heterogeneity of the graders' backgrounds and competence leads to subjectivity and biases (Luo et al. 2014). Most previous approaches to improving the reliability of peer grading have involved creating (complex) algorithms to identify and match graders and students in a way that minimizes predicted biases (Piech et al. 2013). Yet

[1] We focus on the peer grading mechanism and its effect on administering a MOOC. Thus we ignore other issues relevant to MOOCs, such as students' intrinsic motivation to learn, content design, and dropout rates.

little attention has been devoted to the behavior of peer graders, who may well act as rational *economic* agents.[2]

We characterize grading as a process of assessing the quality of an assignment and reporting a score. To assess quality, student graders must acquire information about the quality of the assignment. The more information students acquire, the more accurate—but also more costly—their assessment will be (Arrow 1996). Student graders are motivated to have their reported score reflect the true quality of the assignment (absolute accuracy) as well as to have it be close to the MOOC's final reported grade (relative accuracy), which depends on the scores of other peer graders. As a result, student graders will behave rationally when acquiring information. Also, "strategic" graders (unlike "honest" graders) may find it in their interest to report scores that differ from their assessed quality of an assignment. As economic agents, graders carefully balance costs and benefits before determining their strategies for assessment and reporting (Falk and Ichino 2006).

Using our model, we first examine the current peer grading practice. We find the unique symmetric equilibrium information acquisition level and reporting strategy for a homogeneous group of student graders. We then assess the resulting peer grading for inaccuracy, as measured by the mean squared error (MSE), and the expected grade that an assignment would receive based on peer grading. We observe a systematic grading *bias toward the mean*, which tends to lower students' learning effort during the course. Both inaccuracy and bias are exacerbated when the student graders are strategic rather than honest in their score reporting. We examine how course characteristics and the peer grading structure affect the outcomes of peer grading.

With the aim of improving current practice in terms of grading inaccuracy (which calls for a reduced MSE) and learning (to be addressed by de-biasing), we propose a *scale-shift* grading scheme that generalizes the current peer grade aggregation practice (namely, a simple average). We assess how much the scale-shift factor affects (a) the equilibrium level of information acquired by student

[2] In MOOCs, ensuring that grades are submitted takes precedence over ensuring that graders do a good job. Thus, students do not see their own grades until they have already graded the other students' assignments (Coursera 2018).

graders and (b) the reporting behavior of those graders. We find that a scale-shift grading scheme can improve learning because it can eliminate not only strategic score reporting but also the bias toward the mean. We examine the scale-shift factors that can minimize the MSE or eliminate the bias, discuss how these can be practically implemented, and report on improvements to grading accuracy and student learning based on simulation outcomes.

The rest of our paper proceeds as follows. Section 2 discusses how our model treats information as a nonstatic variable instead of static variable as is often the case in the field of economics, and reviews the related literature. Section 3 introduces the basic model, characterizes grading as a process of information acquisition and reporting, and describes the economic trade-offs faced by student graders. Section 4 analyzes the equilibrium information acquisition level and reporting strategy of student graders as well as their effects on grading accuracy and expected grades, which in turn impact students' learning effort during the course. Section 5 presents the scale-shift grading scheme and identifies the scale-shift factors that minimize the MSE or eliminate bias. We discuss how these can be implemented in practice and their impact on peer grading accuracy and student learning. Section 6 concludes with a brief summary of our work and some suggestions concerning related research opportunities. All proofs are given in the Appendix.

2. Related Literature

Information has long been considered an important variable governing the behavior of economic agents (Arrow 1984). Viewed as a signal that reduces uncertainty (about the state of the world), renowned scholars have framed it as an economically interesting category of goods worthy of careful enquiry. Arrow (1973) discusses information's essential characteristics and Arrow and Radner (1979) studies its role in a team setting.

Reflecting on the role played by information in the field of economics, Arrow (1996) points out that information has largely been treated as a static variable, employed predominantly in the agency setting with asymmetric information among economic agents. More specifically, information is usually considered to be either available to an agent (who observes a random variable) or unavailable to an agent (who knows only its distribution). Arrow (1996) argues that information has become much more "transient" in the sense that agents can now easily expend resources to gain more information. He therefore calls for treating information as a *continuous* variable—much as it is in such disciplines as mathematical statistics, decision theory, and communication engineering. Thus, in lieu of focusing on the consequences of information, as in information value theory (Howard 1966), Arrow argues for examining the cost of acquiring information and what that cost implies for the equilibrium economic behavior of agents.

To aid in this effort, Arrow (1996) presents a useful cost function for acquiring information; this function is based on Shannon's measure of information from communication engineering, and it is consistent with the Bayesian normal sampling used in mathematical statistics and decision analysis. We respond to Arrow's call with the present study of peer grading in MOOCs, where "grading" requires acquiring information about (and reporting) assignment quality, and a peer grader is an *economic* agent who can determine how much information to acquire for the purpose of reporting scores that are close to those reported by other graders.

Thus, a student grader's grading accuracy can be evaluated in terms of how closely that student's grades resemble other submitted grades. This approach is related in spirit to the Keynesian beauty contest (Keynes 1936), where each rational agent relies not only on their own opinion but also on (inferred) public perception before making their decision. This concept is applied formally by Morris and Shin (2002), who study the effect of public information when each rational agent strives to choose an action that mimics the actions of others. In addition to reporting a grade, the student graders in our model incur the cost of acquiring information used to assess the assignment quality; hence, our model generalizes the model presented by Morris and Shin. Shleifer (1985) and Savva et al. (2019) exploit the presumed rational behavior of economic agents (e.g., hospitals) to examine how a regulator can incorporate the simple average of others' actions (i.e., a yardstick) in a mechanism designed to improve the odds of achieving the regulator's objective—which usually involves improved social welfare. In the present paper's setting, we explain how the simple average grading scheme for an MOOC can be replaced with a scale-shift scheme to better meet the primary objectives of peer grading: accurate and unbiased grades.

The scale-shift grading scheme bears some similarity to the common notion of curved grading in that the initial evaluation of an assignment is transformed into a different final score in an order-preserving manner. However, the critical difference is that while curved grading alters grades to fit a predetermined distribution (e.g., the top 10% receive As, etc.), scale-shift grading does not do this. Moreover, curved grading often takes place in a non-peer grading setting and negatively impacts students' learning by decreasing their motivation to learn (Michaels 1976), creating unhealthy competitive classroom environments (Grant 2016), and/or harming students' confidence and long-term educational outcomes (Calsamiglia and Loviglio 2019). In contrast, we show that the scale-shift factor improves student learning outcomes in a peer grading setting by improving the accuracy and fairness of the peer grading outcome.

The efficacy of MOOC peer grading has been widely examined in the education literature. For example, Falchikov and Goldfinch (2000) find that peer assessments better resemble faculty assessments in academic courses than in professional courses, Cho et al. (2006) suggest that the aggregate ratings of at least four peers on a piece of writing are as valid as instructor ratings, and Price et al. (2016) show that using a web-based tool that facilitates peer review can produce accurate peer-reviewed scores and motivate student learning. Recently, the computer science literature has paid considerable attention to how to improve MOOC peer grading. For example, Piech et al. (2013) develop an algorithm that matches graders to assignments based on student backgrounds and then computes a composite assessment based on peer scores. Walsh (2014) introduces a method that weights grades by the reliability scores of the grading agents, and Luaces et al. (2015) propose a factorization model. All of these studies show how the proposed methods can improve the accuracy of grading when using large data sets from courses offered by MOOCs (e.g., Coursera). Luo et al. (2014) investigate the reliability of grades assessed by peers using data collected from Coursera; these authors suggest that the reliability of peer grading can be improved by training students to be better graders. Given how difficult it is for student graders to assign precise *cardinal* scores, some studies have examined *ordinal* peer evaluation mechanisms as well as their benefits and aggregation methods (see, e.g., Shah et al. 2013, Raman and Joachims 2014). These studies focus on vetting graders based on predicted skill level or bias, but they ignore the costs of grading and the graders' balancing of how much effort to exert depending on those costs and the behavior of other graders. Our framework complements these studies by addressing graders' economic behavior, rather than the effects of the heterogeneity of grader types.

A broad stream of literature in economics and education examines how the presence of peers affects student learning. Researchers have shown, for example, that the presence of peers and social interaction among peer learners can have a positive effect on their test scores (Sacerdote 2001, Zimmerman 2003, Calvo-Armengol et al. 2009). More recently, Zhang et al. (2017) find that social interaction among students on an MOOC discussion board increases not only their test participation rates but also their test scores. Our study presents a similar result in the context of peer grading: an improved peer grading mechanism motivates learning.

Aggregation of peer scores in the MOOC setting is similar to the aggregation of expert opinions, concerning which (as mentioned by Arrow 1996) there is extensive research in the fields of mathematical statistics and decision theory. Axiomatic treatments and rules for aggregating expert forecasts or opinions exist (Bordley 1982, Clemen 1986, Clemen and Winkler 1990, Satopää 2014), as well as studies of aggregating distributions that include the aggregation of point estimates (DeGroot 1974, DeGroot and Mortera 1991, Abbas 2009). Some aggregation methods (e.g., a simple average) have been criticized as inefficient (Granger 1989); more recently, Satopää (2019) generalizes this view and shows that an aggregator that never strays beyond the smallest and largest predictions is never efficient in practice. In the present study we consider student graders who are not experts, so we focus on their equilibrium information acquisition level when they forecast an assignment's quality *before* aggregation. In addition, we study the scale-shift grading scheme (a variation of the arithmetic mean), which allows the aggregation of peer grades to fall outside the graders' range, thereby encouraging student graders to acquire more information and produce more accurate assessments.

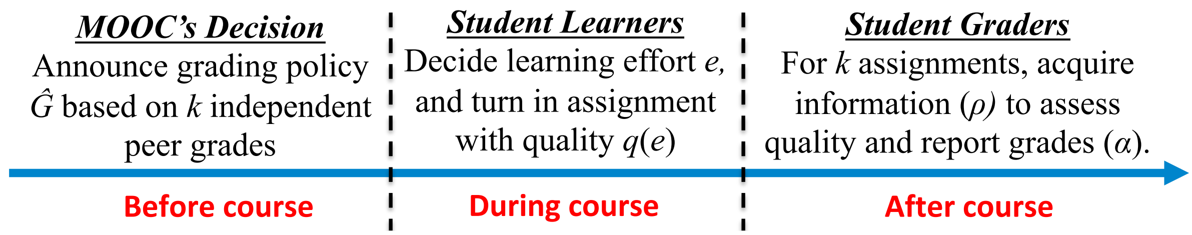
Finally, there is a related literature about providing incentives for agents to be truthful in equilibrium. Prelec (2004) presents a scoring method for eliciting truthful subjective data in situations where objective truth is unknowable. Miller et al. (2005) devise a scoring system intended to induce honest feedback in the evaluation of products and vendors. To better understand large-scale evaluations of product or service reviews on online platforms, Dasgupta and Ghosh (2013) examine the issue of incentivizing the crowd to exert the effort needed to provide good and truthful evaluations, Radanovic and Faltings (2015) discuss the solicitation of crowd-sourced opinions in the context of product reviews on the web, Shnayder et al. (2016) introduce a "correlated agreement" mechanism aimed at incentivizing participants to reveal truthful information, and Kamble et al. (2018) describe a novel reward mechanism for eliciting honest responses from agents in the absence of verifiability. However, research in this "truth serum" vein assumes that agents know the correct answers and focuses on how agents can be induced to reveal those answers. In contrast, the agents (student graders) in our study do *not* know the correct answers and so must acquire information in order to assess the quality of assignments. We show that, despite the incentives of student graders to report scores that differ from their respective true assessments, our proposed scale-shift grading scheme can motivate them to report their scores truthfully.

3. Model

We consider an MOOC with a large number of students, each of whom must submit a final assignment at the end of the course. We model this setting as a chronologically staged process; see Figure 1. Students are learners during the course and graders after the course. To better focus on how the economic behavior of graders affects learning, we assume that the students—as both graders and learners—are homogeneous.

Before the course. The MOOC announces its peer grading policy. Each student is asked to grade k randomly selected assignments, excluding the grader's own assignment, so that each submitted

Figure 1 Sequence of decisions



assignment will be independently evaluated by k students. The final grade for a given assignment of true quality q, or $\hat{G}(G_1,\ldots,G_k)$, depends on the k independent submitted grades $G_1,\ldots,G_k$. For example, if the final grade is simply an average of the grades received, then

$$\hat{G}(G_1,\ldots,G_k)=\frac{1}{k}\sum_{i=1}^{k}G_i(q).$$

The submitted peer grades $G_i(q)$ are stochastically increasing in quality, that is, a higher-quality assignment is expected to receive a higher grade. Each grade depends on (a) the level of information that each student grader decides to acquire and (b) the scores they decide to report.

During the course. As a learner, each student exerts learning effort e and submits an assignment of quality q that depends on e as follows:

$$q(e)=\Big(\kappa\log(e)+\theta\Big)+\varepsilon. \tag{1}$$

This quality is concave increasing in the effort e. The more effort is exerted in learning, the higher the expected quality of the assignment—although there are diminishing marginal quality returns for the learning effort. The parameter κ represents the ease with which quality can be achieved from new learning, and the parameter θ represents the quality derived from prerequisite knowledge. (Any other concave increasing function can replace the logarithmic function that we use.) The term $\varepsilon\sim\mathcal{N}(0,\sigma^2)$ signifies the inherent randomness in producing the assignment's ultimate quality—that is to say, more (resp. less) effort need not lead to a higher (resp. lower) quality assignment, and the variance σ^2 in that quality may reflect course characteristics. The assumption of normality reflects the "bell curve" often observed in the grades of large classes.

Student learners seek to maximize their expected grades, $\hat{G}$, after accounting for effort:

$$\max_{e} \mathbb{E}[\hat{G}|q(e)] - e.$$

If $e^\star$ denotes the optimal learning effort, the mean quality of the assignment is

$$\mu \triangleq \mathbb{E}[q(e^\star)] = \kappa \log(e^\star) + \theta.$$

After the course. As a grader, each student receives k randomly selected assignments to grade. For each assignment, the student grader i assesses its quality and reports a score G_i. The student who submitted the assignment will receive a final grade of $\hat{G}(G_1, \ldots, G_k)$. We are interested in the economics of peer grading during this stage, which we formalize next.

3.1. Model of Grading: Quality Assessment and Score Reporting

Suppose that a given assignment's true quality is q and that if the instructor were to grade the assignment, it would receive a grade identical to its quality. However, if student i grades the assignment, then i's lack of subject-matter expertise means that the assignment's student-assessed quality, $\mathcal{A}_i$, may not reflect its true quality q. In this scenario, the instructor facilitates the grading process by preparing a detailed evaluation rubric that indicates (or directly states) that the final grade distribution should be normal with mean μ and variance σ^2.

In our setting, in the spirit of the rational expectation equilibrium, we assume that the students determine their learning effort and grading decisions based on their belief about the mean quality μ of the assignments and that the realized mean quality is consistent with their belief. Moreover, the MOOC can accurately predict the sample mean μ. This is a reasonable assumption in the recurring MOOC setting we consider, for three reasons: students can observe past MOOC average scores (which do not change much with large class sizes), the grading rubric can encourage or mandate a mean, or the instructor can sample the grades after the course and announce the sample mean. Thus, as we shall see, μ does not play an important role in our model of peer grading.

Assessment of Quality. Assessing the quality of a given assignment amounts to an information acquisition process for student graders, the accuracy of which depends on the *information level*, $\rho_i \in [0,1]$, that they decide to acquire. Given an assignment of quality q, the grade $\mathcal{A}_i(\rho_i)|q$ assessed by student i satisfies

$$\mathcal{A}_i(\rho_i)|q \sim \mathcal{N}((1-\rho_i)\mu + \rho_i q, (1-\rho_i^2)\sigma^2). \tag{2}$$

That is, the quality assessment resembles an "anchor and adjustment" process. The more information that is acquired by the student, the closer the assessed quality $\mathcal{A}_i$ will be to the assignment's true quality q, and the less variance there will be in submitted grades. If $\rho_i = 0$, then the student acquires no information and assesses the quality as a random draw from $\mathcal{N}(\mu, \sigma^2)$, while at the other extreme, if $\rho_i = 1$, then full information is acquired and so the assessed quality $\mathcal{A}_i$ is equal to the true quality q.

The quality of the assignment is a random variable represented by upper case Q; after the grader receives a specific assignment, its realized quality is represented by lower case q. (For simplicity, we will write "$|q$" instead of "$|Q = q$"). Expression (2) corresponds to the marginal distribution of the following bivariate normal distribution between the quality assessed by a student grader i, $\mathcal{A}_i(\rho_i)$, and quality of assignment Q,[3]

$$(\mathcal{A}_i(\rho_i), Q) \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho_i\sigma^2 \\ \rho_i\sigma^2 & \sigma^2 \end{pmatrix} \right). \tag{3}$$

Here, the information level $\rho_i \in [0,1]$ corresponds to the correlation coefficient between the student-assessed quality and the true quality. Empirical studies have found that the correlation between peer and teacher grades can vary from 0.29 to 0.69 (Li et al. 2016).

For the purposes of grade aggregation, we also assume conditional independence between peer graders—that is: given an assignment's quality q, any two assessed qualities $\mathcal{A}_i|q$ and $\mathcal{A}_j|q$ provided by student graders are uncorrelated with each other. Yet *prior* to knowing an assignment's quality,

[3] The variance of Q and $\mathcal{A}_i$ need not be identical. For simplicity, we assume that both are equal to σ^2.

$\mathcal{A}_i$ and $\mathcal{A}_j$ are correlated (by $\rho_i\rho_j$) because both ρ_i and ρ_j are correlated with the underlying quality q.[4] Thus, $(\mathcal{A}_1(\rho_1), \mathcal{A}_2(\rho_2), \ldots, \mathcal{A}_k(\rho_k), Q) \sim \mathcal{N}(\vec{\mu}, \Sigma)$, where $\vec{\mu} \in \mathbb{R}^{k+1}$ and $\Sigma \in \mathbb{R}^{(k+1)\times(k+1)}$:

$$\vec{\mu} = \begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \\ \mu \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 1 & \rho_1\rho_2 & \ldots & \rho_1\rho_k & \rho_1 \\ \rho_2\rho_1 & 1 & \ldots & \rho_2\rho_k & \rho_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_k\rho_1 & \rho_k\rho_2 & \ldots & 1 & \rho_k \\ \rho_1 & \rho_2 & \ldots & \rho_k & 1 \end{pmatrix} \sigma^2.$$

Reporting the Score. Together with acquiring information, the student graders also decide whether to make an accurate report about their assessed quality. A similar problem is examined by Morris and Shin (2002), where agents employ a reporting strategy that takes the convex combination of public and private signals in equilibrium. In much the same spirit, we consider a linear reporting strategy that combines public information μ and private information $\mathcal{A}_i(\rho_i)$:

$$G_i = \alpha_i \mathcal{A}_i(\rho_i) + (1 - \alpha_i)\mu.$$

In other words, given an assignment to grade, a student grader has an additional lever $\alpha_i \in [0, 1]$ that can substitute for information acquisition ρ_i. If $\alpha_i = 1$, then the reported score equals the outcome of the assessment (i.e., we have truthful reporting); but if $\alpha_i < 1$, then the reported score has been adjusted toward the sample mean μ. Hence we say that the higher α_i is, the more *truthful* the reporting. We consider two types of student graders, as described next.

Honest type. Graders of the honest type always report truthfully; that is, $\alpha_i = 1$ and $G_i = \mathcal{A}_i$ for all i. In other words, these graders behave strategically only when determining their level ρ_i of information to acquire about the assignment's quality.

Strategic type. Graders of the strategic type behave strategically regarding both information acquisition and reporting. That is, they behave strategically when deciding (ρ, α) and, in contrast to the case of honest graders, the level $\alpha_i \in [0, 1]$ of reporting is determined by an equilibrium.

[4] $\mathrm{Cov}(\mathcal{A}_i, \mathcal{A}_j) = \mathbb{E}[(\mathcal{A}_i - \mu)(\mathcal{A}_j - \mu)] = \mathbb{E}\big[\big(\rho_i(Q-\mu) + \varepsilon_i\sqrt{1-\rho_i^2}\sigma\big)\big(\rho_j(Q-\mu) + \varepsilon_j\sqrt{1-\rho_j^2}\sigma\big)\big] = \mathbb{E}[\rho_i(Q-\mu)\rho_j(Q-\mu)] = \rho_i\rho_j\mathbb{E}[(Q-\mu)^2] = \rho_i\rho_j\sigma^2$, where ε_i and ε_j are independent standard normal variables.

In what follows, we examine the effects of strategic grading behavior by analyzing homogeneously honest graders and homogeneously strategic graders separately. To make the reference to honest or strategic graders explicit, we will use the subscript h or s, respectively, wherever appropriate. However, for general usage we will omit the subscript.

3.2. Economics of Assessing and Reporting Grades

In this section we describe the economic trade-off facing each rational grader.

Cost of Information Acquisition. Information acquisition is costly, and how that cost is modeled plays a significant role in determining the economic behavior of peer graders. Arrow (1996) proposes two functional forms for the information acquisition cost $\mathcal{C}(\rho)$. One employs the Shannon entropy measure of information,[5] and the other relies on Bayesian normal sampling. Specifically, Arrow (1996) suggests that the cost of acquiring information ρ, that is, $\mathcal{C}(\rho)$, should be *proportional* either to the decrease in expected entropy $(H(X) - \mathbb{E}[H(X|\rho)])$ or to the increase in precision τ ($\triangleq 1/\sigma^2$). This economic logic leads to two functional forms for $\mathcal{C}(\rho)$ that result in identical insights. For the sake of brevity, we shall focus on the first illustrated form:[6]

Lemma 1 **(Information Acquisition Cost (Arrow 1996))**.

$\mathcal{C}(\rho) = -c\log(1-\rho^2)$, *for some constant* c.

The proportionality constant c represents the *marginal cost* of acquiring information. We next examine this cost relative to the benefit of information acquisition.

Benefits of Information Acquisition and Reporting. The student graders have incentives to acquire information. First, they have an intrinsic motivation to have their submitted score G_i reflect the true quality q. We remark, however, that the true quality q of an assignment is not

[5] The Shannon entropy of random variable X is defined as $H(X) \triangleq \mathbb{E}[-\log(f(X))]$, where $f(\cdot)$ denotes the probability density function.

[6] The other form leads to $\mathcal{C}(\rho) = -c\rho^2/(1-\rho^2)$. Other more intuitive cost functions, such as a linear or quadratic cost of information as well as $\frac{\rho}{1-\rho}$ or $-\log(1-\rho)$, all lead to expressions that are not amenable to analysis and insights.

observable—only the reported peer grade $\hat{G}$ can be observed. It follows that student graders also have an extrinsic motivation for their submitted score G_i to be close to the MOOC's reported grade $\hat{G}(G_1, \ldots, G_k)$. Thus, we have described two different forms of grading inaccuracies that student graders seek to minimize, which we now define formally. We use the expected squared distance for tractability.

Definition 1 (Absolute Error and Relative Error).

(i) The *absolute* error is $\mathbb{E}[(G_i - Q)^2]$.

(ii) The *relative* error is $\mathbb{E}[(G_i - \hat{G})^2]$.

Decisions of Student Graders. Observe that the relative error depends on other student graders' decisions—a dependence that induces student graders to behave strategically when determining (ρ_i, α_i). Let $\vec{\alpha} \triangleq (\alpha_1, \ldots, \alpha_k)$ and $\vec{\rho} \triangleq (\rho_1, \ldots, \rho_k)$. Each student grader determines their best-response grade assessment (ρ_i, α_i) by solving

$$\min_{\rho_i, \alpha_i} \; \mathcal{C}(\rho_i) + w_1 \mathbb{E}\big(\alpha_i \mathcal{A}_i(\rho_i) + (1-\alpha_i)\mu - Q\big)^2 + w_2 \mathbb{E}\big(\alpha_i \mathcal{A}_i(\rho_i) + (1-\alpha_i)\mu - \hat{G}(\vec{\alpha}, \vec{\rho})\big)^2. \tag{4}$$

The weights w_1 and w_2 respectively represent student graders' intrinsic motivation to reduce the absolute error and their extrinsic motivation to reduce the relative error.

If ρ_i were a constant, then (4) would be similar to the problem proposed by Morris and Shin (2002). Those authors show that, when deciding on the costless reporting strategy α_i, there is a unique equilibrium linear grading strategy $\alpha^\star$, and this equilibrium represents the unique equilibrium for all strategies. Our model generalizes Morris and Shin (2002)'s model by incorporating the information acquisition decision ρ_i and its cost $\mathcal{C}(\rho_i)$.

4. Analysis

In this section we examine MOOCs' current peer grading practice, which uses the simple average as the aggregation method, and we discuss how the strategic (economic) behavior of student graders affects the outcomes of that practice.

An assignment's final grade is the average of the grades assigned by k independent graders. By (2), the assessed quality of an assignment with actual quality q is characterized by the following random variable:

$$\bar{G}(\rho_1,\ldots,\rho_k) \triangleq \frac{\sum_{i=1}^k G_i(\rho_i)}{k} \sim \mathcal{N}\left(\mu + (q-\mu)\frac{\sum_{i=1}^k \alpha_i\rho_i}{k},\ \frac{\sigma^2}{k^2}\left(\sum_{i=1}^k \alpha_i^2(1-\rho_i^2)\right)\right).$$

Recall, from (4), that graders have an incentive to submit grades that are close to those submitted by their peers. As a consequence, the strategies regarding both information acquisition ρ_i and the score reporting α_i depend on those of other students. Our first proposition establishes the existence of a unique (and symmetric) equilibrium.

Proposition 1 **(Unique Symmetric Equilibrium)**. *Let $\hat{G} = \bar{G}$.*

(i) *If student graders are homogeneously of the* honest *type, then the following statements hold:*
- *If $w_1 > 0$, then there exists a unique symmetric equilibrium information acquisition level $\rho_h^\star \in (0,1)$.*
- *If $w_1 = 0$, then one of the equilibria is the failure of peer grading: $\rho_h^\star = 0$.*[7]

(ii) *If student graders are homogeneously of the* strategic *type, then the following statements hold:*
- *If $c < \sigma^2 w_1^2 k^2/(w_1 k^2 + w_2(k-1)^2)$, then there exist a unique symmetric equilibrium information acquisition level and reporting strategy $(\rho_s^\star, \alpha^\star) \in (0,1)\times(0,1)$.*
- *If $c \geq \sigma^2 w_1^2 k^2/(w_1 k^2 + w_2(k-1)^2)$, then the failure of peer grading ($\rho_s^\star = \alpha^\star = 0$) is one of the equilibria.*

(iii) *Furthermore, $\rho_s^\star \leq \rho_h^\star < 1$ and $\alpha^\star \leq \rho_s^\star$.*

When student graders are honest in reporting scores, we observe that there is always a nonfailing unique equilibrium *provided* that those graders have intrinsic motivation ($w_1 > 0$). Absent intrinsic motivation, peer grading fails—in which case, the MOOC's peer grade $\hat{G}_h$ is uncorrelated with the assignment's true quality q. When student graders care about their submitted scores' proximity to the true quality, they seek to minimize the absolute inaccuracy and hence will acquire positive

[7] To identify a nonfailing equilibrium, we treat as a failure any multiple equilibria involving failure.

information; however, when they only care about their scores' proximity to other graders, they are not inclined to acquire costly information, and thus they submit scores that are uncorrelated to the true qualities. That peer grading does *not* fail in practice (see, e.g., Li et al. 2016) suggests the existence of intrinsic motivation—that is, of $w_1 > 0$.

Similarly, when student graders are strategic in reporting their assessment, there exists a unique nonfailing equilibrium if the marginal cost c is below a certain threshold. However, if that cost is above the threshold, then peer grading fails and degenerates to a pure-strategy equilibrium in which everyone reports the average grade, μ. In that event, $\hat{G}_s = \mu$, and there will be no variance in the assessed grades.

Finally, we observe that student graders who are strategic about reporting acquire less information than do honest graders, and they always report untruthfully.

The following corollary provides explicit expressions for the equilibrium when $w_2 = 0$.

COROLLARY 1 **(Effect of w_1 and w_2 on the Failure Regime)**. *Suppose $w_1 > 0$ and $w_2 = 0$. If $c \le w_1\sigma^2$, then*

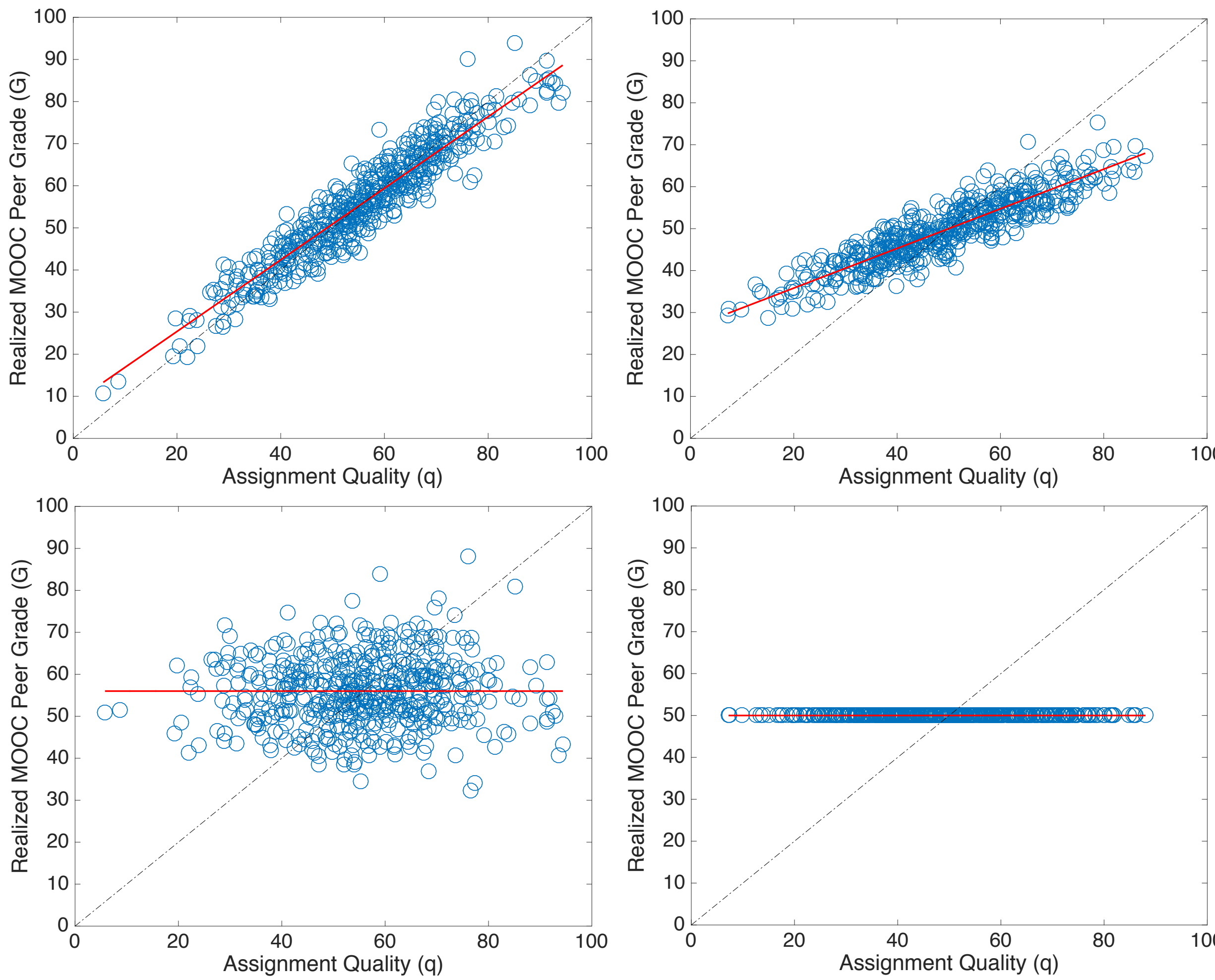
$$\rho_h^\star = \frac{\sqrt{c^2 + 4w_1^2\sigma^4} - c}{2w_1\sigma^2}, \qquad \rho_s^\star = \alpha^\star = \left(1 - \frac{c}{w_1\sigma^2}\right)^{1/2};$$

otherwise, peer grading fails, and $\rho_h^\star = \rho_s^\star = \alpha^\star = 0$.

We observe that, in contrast to the case of no intrinsic motivation ($w_1 = 0$), if there is no extrinsic motivation ($w_2 = 0$), then peer grading never fails for honest graders; it fails only when $c \ge w_1\sigma^2$ for strategic graders (in line with the failure condition of Proposition 1(ii)). Moreover, equilibrium information acquisition levels do not depend on the number k of peer graders, because relative inaccuracy is costless ($w_2 = 0$).

The realized peer grades as a function of quality are plotted in Figure 2, with the left and right columns illustrating the peer grading outcome for homogeneously honest and homogeneously strategic graders, respectively. The slope of the solid (red) lines represents the correlation between an assignment's quality Q and the MOOC's peer grade $\hat{G}$, and also corresponds to the equilibrium information acquisition level $\rho_h^\star$ if graders are honest or to $\alpha^\star\rho_s^\star$ if they are strategic. Moreover, the slope is less than 1 (45 degrees) and less for strategic graders (upper right panel) than for honest graders (upper left panel). The figure's lower panels correspond to a setting in which $w_1 = 0$—that is, when peer grading fails. In this setting, the reported MOOC peer grades are uncorrelated with the actual quality even when the graders are honest reporters (lower left), whereas they are uniformly μ when the graders are strategic reporters (lower right).

Figure 2 The realized MOOC peer grade ($\hat{G}$) as a function of the true assignment quality (q).



Note. The solid (red) line represents the expected MOOC peer grade $\mathbb{E}[\hat{G}|q]$ as a function of the true assignment quality q. The panels in the left and right columns plot results for graders who are respectively honest and strategic. The slope corresponds to the correlation between $\hat{G}$ and Q: $\rho_h^\star = 0.85$ (upper left panel); $\alpha^\star \rho_s^\star = (0.63)(0.75)$ (upper right panel); $\rho_h^\star = 0$ (lower left panel); and $\rho_s^\star = \alpha^\star = 0$ (lower right panel). Parameters: $\mu_h = 56$, $\mu_s = 50$ ($\mathbb{E}[q(e^\star)] = 10\log(e^\star) + 35$), $w_1 = 1$, $w_2 = 1$ (upper row panels), $w_1 = 0$ (lower row panels), $k = 3$, $\sigma = 15$, and $c = 100$.

From the MOOC's perspective, reducing the difference between peer grades and the assignment's true quality is necessary in order to sustain the scalable peer grading practice. If the instructor could replace peer grading, then the points on these graphs would be uniformly scattered along the 45-degree line. However, we can see that the peer grading average (solid red line) deviates from the 45-degree line and that there are additional variations around the peer grading average. These two factors contribute to the inaccuracy of grading, which is measured by the MSE. From the students' perspective, prior to taking the course, there is a need for assurance that the final grade $\hat{G}$ based on peer grading will accurately reflect a high level of quality q. These considerations lead to our next proposition.

PROPOSITION 2 **(Effect of Graders' Strategic Behavior on the MOOC's Reported Peer Grades)**.

(i) *The inaccuracies in the reported grades as measured by the MSE are as follows:*

$$\mathbb{E}[(\hat{G}_h - Q)^2] = \sigma^2\left(\frac{1-\rho_h^{\star 2}}{k} + (1-\rho_h^\star)^2\right), \quad \mathbb{E}[(\hat{G}_s - Q)^2] = \sigma^2\left(\frac{\alpha^{\star 2}(1-\rho_s^{\star 2})}{k} + (1-\alpha^\star\rho_s^\star)^2\right).$$

In addition, $\mathbb{E}[(\hat{G}_h - Q)^2]$ *is decreasing in* $\rho_h^\star$ *and* $\mathbb{E}[(\hat{G}_s - Q)^2]$ *is decreasing in both* $\rho_s^\star$ *and* $\alpha^\star$*.*

(ii) *The expected peer grades for an assignment of quality* q *are as follows:*

$$\mathbb{E}[\hat{G}_h|q] = (1-\rho_h^\star)\mu + \rho_h^\star q, \quad \mathbb{E}[\hat{G}_s|q] = (1-\alpha^\star\rho_s^\star)\mu + \alpha^\star\rho_s^\star q.$$

(iii) *We have* $\mathbb{E}[(\hat{G}_s - Q)^2] \geq \mathbb{E}[(\hat{G}_h - Q)^2]$ *and* $\mathbb{E}|[\hat{G}_s|q] - q| \geq \mathbb{E}|[\hat{G}_h|q] - q|$*.*

First, we observe that a lower equilibrium information acquisition level $\rho^\star$ or less truthful reporting behavior $\alpha^\star$ leads to grading that is less accurate. Second, peer grading exhibits a systematic bias toward the mean. Namely, $q > \mu$ implies $\mathbb{E}[\hat{G}] < q$, and $q < \mu$ implies $\mathbb{E}[\hat{G}] > q$. The implication of this latter observation is that a high-quality (resp. low-quality) assignment is expected *not* to receive the high (resp. low) grade it deserves. That outcome can be clearly seen in the upper right panel of Figure 2, where the variation in the assignment quality q (ranging mostly between 10 and 90 on the horizontal axis) is *greater* than the variation in the realized MOOC peer grades $\hat{G}$ (mainly between 30 and 70 on the vertical axis). Finally, we observe again that, when student graders are strategic in reporting their scores, peer grading suffers in terms of both accuracy and bias.

From a student's perspective, the peer grading's bias toward the mean implies that a high-quality assignment's grade ($\mathbb{E}[\hat{G}_h|q]$ or $\mathbb{E}[\hat{G}_s|q]$) would be *discounted* by the factor $\rho_h^\star$ (with honest graders) or the factor $\alpha^\star\rho_s^\star$ (with strategic graders). The knowledge of that bias may affect students' motivation to exert learning effort during the course, as formalized next.

COROLLARY 2 **(Effect of Bias toward the Mean on Learning)**.

(i) *If graders are* honest, *then a student's learning effort during the course is* $e_h^\star = \kappa\rho_h^\star$.

(ii) *If graders are* strategic, *then a student's learning effort during the course is* $e_s^\star = \kappa\alpha^\star\rho_s^\star$.

Lower values of $\rho_h^\star$ (and $\rho_s^\star\alpha^\star$) reflect a *higher* grading bias towards the mean, in which case students exert less learning effort. Also, based on Proposition 1(i), in environments with honest graders, students exert more effort to learn, i.e., $e_h^\star > e_s^\star$. This can be observed in Figure 2. The average assignment quality is higher when graders are honest ($\mu_h = 56$, left column) compared to when graders are strategic ($\mu_s = 50$, right column).

We have shown that a higher equilibrium information acquisition level $\rho^\star$ and truthful reporting $\alpha^\star$ improve grading accuracy, reduce grading bias, and motivate student learning. In which situations, then, is peer grading most effective? What can an MOOC do to improve peer grading? We develop some insight regarding these concerns by examining the properties of $\rho_h^\star$, $\rho_s^\star$, and $\alpha^\star$.

COROLLARY 3 **(Effects of k, w_1, w_2, c, and σ)**.

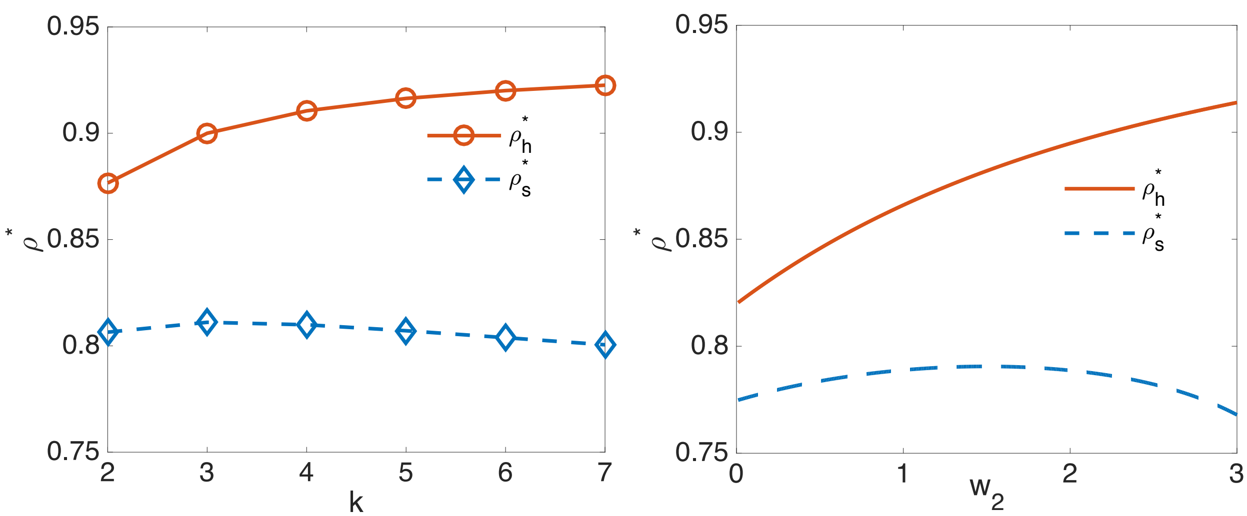
(i) $\rho_h^\star$ *is increasing in* k, σ, w_1/c, *and* w_2/c;

(ii) $\rho_s^\star$ *and* $\alpha^\star$ *are increasing in* σ *and* w_1/c.

We observe that higher intrinsic motivation (w_1) or lower marginal cost (c) leads students to increase their level of information acquisition and also makes it more likely that they will report truthfully. Moreover, a course exhibiting a higher standard deviation in quality (σ) also motivates students to acquire more information, because higher variation in assignment quality requires student graders to adjust from the anchor (mean) to a greater degree if they are to improve both relative and absolute accuracy.

It is interesting that the number k of peer graders and the extrinsic motivation w_2 can each have different effects on the equilibrium outcome depending on whether student graders are honest or strategic in reporting their assessed scores. When student graders are honest, increasing these parameters motivates them to acquire more information (part (i)). However, the same cannot always be said when graders are strategic (part (ii)). The difference reflects these graders' access to an additional lever α, by which an increase in w_2 or k encourages strategic reporting behavior while reducing the cost of acquiring information. This dynamic is illustrated in Figure 3.

Figure 3 Effects of k (left panel) and w_2 (right panel) as a function of grader type: honest or strategic.



Note. Parameters: $\sigma = 1$, $c = 0.38$, $w_1 = 1$, $w_2 = 2$ (left panel), and $k = 3$ (right panel).

From the MOOC's perspective, it is easy to alter the number k of graders. Moreover, the observability of $\hat{G}$ implies that an MOOC can affect the extrinsic motivation parameter w_2 by keeping track of graders' accuracy and by rewarding accurate grading—either financially (e.g., giving discounts on future course enrollments) or nonfinancially (e.g., offering recognition; see Gallus 2017). The other factors w_1, c, and σ are more difficult to influence. So if the student graders are honest, then the MOOC can improve its peer grading by influencing k or w_2. Any increase in k, however, results in each student i being required to grade more assignments. Because an MOOC will not want to overburden student graders (note that the grading burden $k \cdot \mathcal{C}(\rho_i)$ increases with k when graders are honest), k should be not be increased beyond an acceptable grading burden. In contrast, if the student graders are strategic, the equilibrium information acquisition level $\rho_s^\star$ is nonmonotonic in k and w_2, making such a strategy less straightforward. Next, we present a grading scheme that bypasses such challenges by improving peer grading dynamics more fundamentally.

5. The Scale-Shift Grading Scheme $\hat{G}$

In the previous section, we showed that the peer grading policy based on the simple average $\bar{G}$ is easy to implement, but suffers from inaccuracy and systematic bias. In this section, we propose a practically implementable scale-shift grading scheme that can improve these shortcomings.

As motivation for the scale-shift grading scheme, suppose that an MOOC seeks to improve grading accuracy so that the MSE $\mathbb{E}[(Q-\hat{G})^2]$ is minimized or to de-bias grading so that $\mathbb{E}[\hat{G}|q]=q$. Each of these objectives can be achieved by applying a simple linear transformation of the simple average $\bar{G}$ as follows:

Lemma 2 (**$\hat{G}$ That Minimizes the MSE or Eliminates Bias Given ρ**).

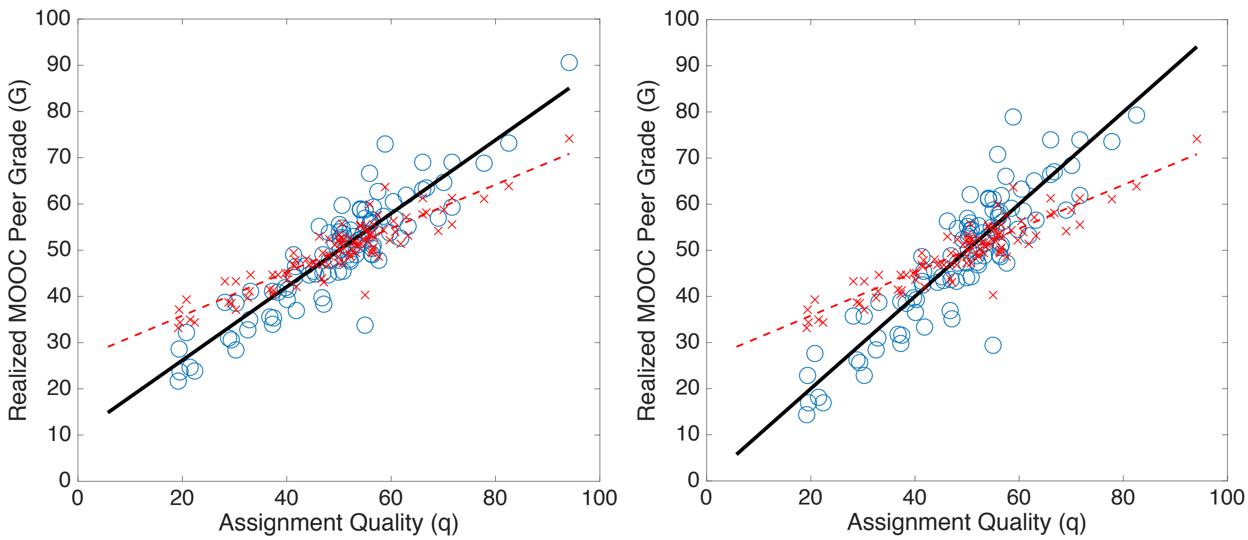
(i) Minimize the MSE: $$\hat{G}=\mu+\frac{k\rho}{\alpha(1+(k-1)\rho^2)}(\bar{G}-\mu).$$

(ii) Eliminate bias: $$\hat{G}=\mu+\frac{1}{\alpha\rho}(\bar{G}-\mu).$$

Observe that both cases take the functional form $\hat{G}=\mu+\xi\cdot(\bar{G}-\mu)$. It is important to bear in mind that an MOOC's announcement of the scale-shift factor ξ will affect each student's trade-off (4) and the resulting equilibrium values of $\rho_h^\star$ and $(\rho_s^\star,\alpha^\star)$. Hence, the values $\frac{k\rho}{\alpha(1+(k-1)\rho^2)}$ and $\frac{1}{\alpha\rho}$ do not achieve the respective aims. However, if the appropriate "scale-shift factor" ξ is applied, the desired effects can be achieved, as illustrated in Figure 4. (The small red $\times$ marks and the dotted red line represent the MOOC peer-grading outcome based on $\bar{G}$; the blue circles and the solid black line illustrate the adjusted grade that each sample point would have received if the scale-shift factor ξ had been employed.) In the left panel, the moderate adjustment represented by $\xi = 1.32$ minimizes MSE; in the right panel, the larger shift in $\xi = 1.38$ helps re-adjust the grade toward the 45-degree line, on which $\mathbb{E}[\hat{G}|q] = q$.

In what follows, we study the impact of announcing the scale-shift factor ξ on equilibrium grading behavior, identify the appropriate scale-shift factors that minimize grading inaccuracy or eliminate bias, and discuss the practical implementation challenges.

Figure 4 Effects of adjusting ξ.



Note. The dashed (red) line represents the expected MOOC peer grade $\bar{G}$ as a function of the true assignment quality q. The left panel illustrates the shift that minimizes grading inaccuracy, represented by the solid (black) line, and the right panel, the de-biasing function, represented by the solid (black) line. Parameters: fully strategic graders with $\rho_s = 0.75$, $\alpha^\star = 0.63$, $\mu = 50$, $w_1 = 1$, $w_2 = 1$, $k = 3$, $\sigma = 15$, and $c = 100$.

The next proposition summarizes how ξ affects a grader's equilibrium level for information acquisition and the reporting strategy.

Proposition 3 **(Impact of ξ on Unique Symmetric Equilibrium)**.

(i) *Suppose that student graders are* honest. *Then there exists a unique equilibrium information acquisition level $\rho_h^\star(\xi)$, $\xi \in (0, k)$. Moreover, $\rho_h^\star(\xi)$ is increasing-decreasing in ξ, and its maximum value is achieved at $\xi = \frac{k}{2}$.*

(ii) *Suppose that student graders are* strategic. *Then there exist* $\bar{c}(\xi), \underline{c}(\xi)$, *and* $\bar{c}(\xi) \geq \underline{c}(\xi)$ [8] *such that:*

- *if* $c < \underline{c}(\xi)$, *then there exists a unique symmetric equilibrium* $(\rho_s^\star(\xi), \alpha^\star(\xi)) = (\rho_h^\star(\xi), 1)$;
- *if* $c \in [\underline{c}(\xi), \bar{c}(\xi)]$, *then there exists a unique symmetric equilibrium* $(\rho_s^\star(\xi), \alpha^\star(\xi))$, *where* $\rho_s^\star(\xi) < \rho_h^\star(\xi)$ *and* $\alpha^\star(\xi) < 1$;
- *if* $c > \bar{c}(\xi)$, *then one equilibrium is the failure of peer grading* $(\rho_s^\star = \alpha^\star = 0)$.

Moreover, $\rho_s^\star(\xi)$ *and* $\alpha^\star(\xi)$ *are increasing at* $\xi \leq 1$ *and decreasing at* $\xi \geq k$.

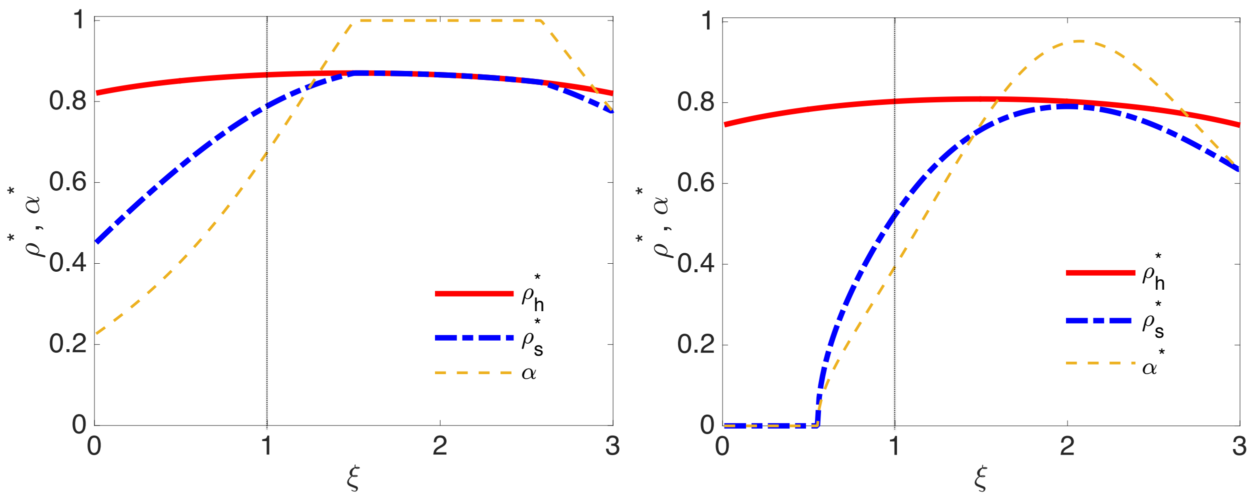
Comparing this proposition to Proposition 1, there are two key observations that demonstrate the benefit of having $\xi > 1$. First, we find that if the grading cost c is small enough, an appropriate scale-shift factor can motivate strategic graders to report their scores truthfully, like their honest counterparts. This outcome indicates that scale-shift grading can altogether eliminate the detrimental effects of graders' strategic reporting behavior. Second, implementing a scale-shift factor $\xi > 1$ can move peer grading out of its failure region (because the value of $\bar{c}(\xi)$ is increasing in ξ), thus enabling peer grading to be applicable to a wider range of MOOC courses.

The impact of a scale-shift factor $\xi > 1$ on the information acquisition levels $\rho_h^\star(\xi)$ and $\rho_s^\star(\xi)$ and on the reporting behavior $\alpha^\star(\xi)$ of student graders are illustrated in Figure 5. At $\xi = 1$, we have $\alpha^\star < \rho_s^\star < \rho_h^\star < 1$, relations that are consistent with Proposition 1. The equilibrium information level $\rho_h^\star(\xi)$ for the honest type of grader is unimodal and symmetric, and its maximum is reached at $\xi = k/2$; in contrast, the corresponding equilibrium information level $\rho_s^\star(\xi)$ for graders of the strategic type is maximized either at $\xi = k/2$, when strategic reporting behavior disappears (left panel of Figure 5), or at $\xi > k/2$ (right panel). Observe that $\rho_h^\star(\xi) \geq \rho_s^\star(\xi)$ for all ξ. The initial increase occurs because the scale-shifted grade $\hat{G}$ places greater weight on each student grader's reported grade, which increases the variance of the aggregation from those grades. Minimizing relative inaccuracy requires that the grader acquire more information (much as when the standard deviation σ in grades is high). The eventual decrease in these values as ξ further increases is caused

[8] Explicit expressions for $\bar{c}$ and $\underline{c}$ are given in the proof in the Appendix.

by the high cost of acquiring a large amount of information, which prompts the graders to reverse their strategy. Observe that the decrease occurs for $\xi < k$, suggesting that it suffices to consider $\xi \in [1, k]$ for the remainder of this paper.

Figure 5 Comparison of honest and strategic graders under low (left panel) and high (right panel) grading costs.



Note. When c is small, an increase in ξ induces strategic graders to report honestly. Parameters: $k = 3$, $w_1 = 1$, $w_2 = 1$, $\sigma = 1$, $c = 0.1$ (left panel), $c = 0.6$ (right panel), and $\xi \in (0, k)$.

The following result reports the impact of ξ on the grading accuracy and bias.

Proposition 4 **(Effect of the Scale-Shift Factor ξ).**

(i) *The MSE depends on ξ as follows:*

$$\mathbb{E}[(\hat{G}_h(\xi) - Q)^2] = \sigma^2 \left((\rho^\star_h(\xi)\xi - 1)^2 + \frac{(1 - \rho^\star_h(\xi)^2)\xi^2}{k} \right),$$
$$\mathbb{E}[(\hat{G}_s(\xi) - Q)^2] = \sigma^2 \left((\alpha^\star(\xi)\rho^\star_s(\xi)\xi - 1)^2 + \frac{(1 - \rho^\star_s(\xi)^2)\alpha^\star(\xi)^2\xi^2}{k} \right).$$

(ii) *The expected peer grade given to an assignment of quality q depends on ξ as follows:*

$$\mathbb{E}[\hat{G}_h(\xi)|q] = (1 - \rho^\star_h(\xi)\xi)\mu + \xi\rho^\star_h(\xi)q,$$
$$\mathbb{E}[\hat{G}_s(\xi)|q] = (1 - \alpha^\star(\xi)\rho^\star_s(\xi)\xi)\mu + \xi\alpha^\star(\xi)\rho^\star_s(\xi)q.$$

The next corollary reveals how the dynamics of ξ increases students' incentive to exert learning effort during the course.

Corollary 4 **(Effect of ξ on Learning $e^\star$).**

(i) *When graders are* honest, *the level of student learning is* $e_h^\star(\xi) = \kappa\xi\rho_h^\star(\xi)$.

(ii) *When graders are* strategic, *the level of student learning is* $e_s^\star(\xi) = \kappa\xi\alpha^\star\rho_s^\star(\xi)$.

Using the expressions of Proposition 4, the scale-shift factors ξ_h^{mse} and ξ_s^{mse} that minimize the MSE for honest and strategic types of graders, respectively, can be represented as

$$\xi_h^{\text{mse}} \triangleq \arg\min_{\xi} \mathbb{E}[(\hat{G}_h(\xi) - Q)^2] \text{ and } \xi_s^{\text{mse}} \triangleq \arg\min_{\xi} \mathbb{E}[(\hat{G}_s(\xi) - Q)^2].$$

Similarly, the scale factors ξ_h^{deb} and ξ_s^{deb} that eliminate bias for honest and strategic types of graders, respectively, exist, as the next lemma shows.

Lemma 3 **(Existence of ξ^{deb}).** *The scale-shift factors* $\xi_h^{\text{deb}} \triangleq \min\left\{\xi \mid \xi = \frac{1}{\rho_h^\star(\xi)}\right\}$ *and* $\xi_s^{\text{deb}} \triangleq \min\left\{\xi \mid \xi = \frac{1}{\alpha^\star(\xi)\rho_s^\star(\xi)}\right\}$ *that de-bias the grades always exist, provided that peer grading does not fail (i.e.,* $c < \bar{c}$*). Moreover,* $\xi_s^{\text{deb}} > \xi_h^{\text{deb}} > 1$.

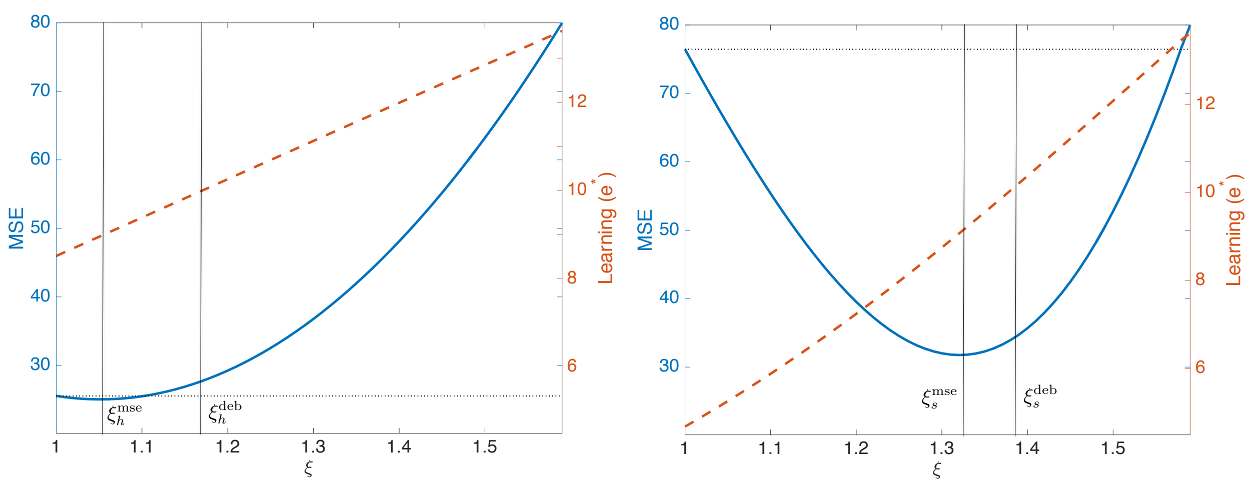
We also remark that as long as peer grading does not fail, we can numerically observe that the scale-shift factor that minimizes the MSE is smaller than the one that eliminates bias, i.e., $\xi_h^{\text{mse}} < \xi_h^{\text{deb}}$ and $\xi_s^{\text{mse}} < \xi_s^{\text{deb}}$.

For any $\xi < \xi^{\text{deb}}$, as in the current peer grading scheme ($\xi = 1$), there is a bias toward the mean, which implies that students will receive similar scores regardless of differences in the quality of their respective assignments (or in their efforts). For example, if $\mu = 50$, an assignment of quality 60 is expected to receive a score of 55 and one of quality 40 should receive a score of 45. However, when $\xi > \xi^{\text{deb}}$, there is a bias *away from* the mean that disproportionately rewards or penalizes the students for (respectively) higher and lower quality assignments. In this case, an assignment of quality 60 is expected to receive a score of 80 and one of quality 40 should receive a score of 20.

Learning $e^\star$ increases with ξ and continues to do so until it reaches $\xi_h^{e\star} \triangleq \arg\max_\xi e_h^\star(\xi)$ and $\xi_s^{e\star} \triangleq \arg\max_\xi e_s^\star(\xi)$. However, these tend to be significantly larger (close to $k = 3$) than their respective de-biasing points ξ_h^{deb} and ξ_s^{deb}. In other words, high learning effort can be achieved by creating a large bias away from the mean.[9] However, implementing $\xi^{e\star}$ results in both highly inaccurate and highly biased peer grading, making this an undesirable candidate for being a practically viable scale-shift factor.

Thus, we will focus on the impact of implementing ξ^{mse} and ξ^{deb}. Figure 6 illustrates the MSE and the learning effort as a function of ξ. In both panels, the prevailing MSE and level of student learning $e^\star$ are represented at $\xi = 1$. As ξ increases, the MSE falls until $\xi = \xi^{\mathrm{mse}}$, after which the MSE increases, and the bias towards the mean decreases until it disappears at $\xi = \xi^{\mathrm{deb}}$, after which bias away from the mean increases. Thus, the level of bias is smaller when ξ^{mse} is employed than when $\xi = 1$. For honest graders (left panel), relative to the current practice ($\xi = 1$), implementing ξ^{mse} improves both the MSE and learning, i.e., it achieves Pareto-efficient improvement, whereas implementing ξ^{deb} improves learning but hurts the MSE. For strategic graders (right panel), both ξ^{mse} and ξ^{deb} achieve Pareto-efficient improvements over the current practice.

Figure 6 The MSE (solid curve) and learning effort (dashed curve) when graders are honest (left panel) versus strategic (right panel).



Note. Parameters: $w_1 = 1$, $w_2 = 1$, $k = 3$, $\sigma = 15$, and $c = 100$.

Comparing the two settings, we observe that the values of ξ_h^{mse} and ξ_h^{deb} are smaller than those of ξ_s^{mse} and ξ_s^{deb}, and their effects on the MSE and learning are smaller for honest graders. Recall

[9] In fact, it can be shown that $\xi^{e\star} > \xi_h^{\mathrm{deb}}$. Otherwise, because $\xi\rho_h^\star(\xi)$ is continuous in ξ and is also less than 1 at $\xi = 1$, there must exist $\tilde{\xi} \le \xi_h^{e\star} < \xi_h^{\mathrm{deb}}$ such that $\xi_h^{e\star}\rho_h^\star(\xi_h^{e\star}) \ge 1$—however, this is contradicted by the definition of ξ_h^{deb}.

from Corollary 3 that if the graders are honest types, then the equilibrium effort $\rho_h^\star$ increases with both k and w_2. Thus, for honest graders, increasing k or w_2 to a certain extent should be an effective alternative to implementing the scale-shift grading. For strategic graders, the effect of ξ is stronger than for honest graders, but the effects of k and w_2 are unclear. We conclude that a scale-shift grading scheme should be considered if an MOOC suspects that its graders are reporting their assessment strategically—which may be the case, for example, when the correlation between assignment quality and assigned grades is low.

Next we discuss, within the bounds of our model assumptions, how the scale-shift grading policy can be implemented in practice and also investigate the performance using a simulation study. Implementation beyond our model assumptions (e.g., involving heterogeneous types of graders) lies beyond the scope of this paper.

Practical Implementation

A key practical challenge of implementing the scale-shift grading policy is determining an appropriate ξ. This is because the parameters for the peer graders' tradeoff in information acquisition, i.e., the cost of acquiring information (c) and the intrinsic and external rewards for accurate grading (w_1, w_2)–or more simply, w_1/c and w_2/c–are unobservable by the MOOC. (For notational simplicity, we normalize $c = 1$ and use w_1 and w_2, respectively.) To be able to implement the scale-shift grading, the MOOC must be able to estimate these.

Thus, the platform operator must first identify the MOOCs for which simple controlled experiments can be conducted. These are the ones that run frequently (e.g., every quarter) with similar student profiles and that produce similar assignment qualities. For the given courses, over subsequent running of the same course, the MOOC can run peer grading using different values of k. For example, if $k = 3$ is currently being employed, then the MOOC can run a peer grading scheme with $k = 2$ or $k = 4$, or both. Next, for each course, assuming that the instructor can recover the true quality of the assignments, the MOOC can ask the instructor to grade a sample of the assignments and submit the scores. The MOOC can then evaluate the correlations between the MOOC's peer grades and the true assignment qualities that result from different values of k, or $\hat{\rho}(k)$. The MOOC can then infer the most likely $\hat{w}_1$ and $\hat{w}_2$ via the maximum likelihood estimate, or

$$(\hat{w}_1, \hat{w}_2) = \underset{(w_1, w_2)}{\arg\min} \sum_{k \in \mathcal{K}} \Big(\hat{\rho}(k) - \alpha^\star(w_1, w_2|k)\rho^\star(w_1, w_2|k) \Big)^2,$$

where $\mathcal{K}$ represents the set of values of k that were used. Observe that $|\mathcal{K}|$ should be no less than 2 in order to estimate two parameters. Finally, using the estimated values of $\hat{w}_1$ and $\hat{w}_2$, the MOOC can compute $\hat{\xi}^{\text{mse}}$ or $\hat{\xi}^{\text{deb}}$ and employ this in the next quarter.

In practice, it can be prohibitively costly or infeasible to run too many experiments. Given the small data sample (2 or 3), the estimates $(\hat{w}_1, \hat{w}_2)$ can result in a large variation and could lead to shift-scale factors $\hat{\xi}^{\text{mse}}$ or $\hat{\xi}^{\text{deb}}$ that are different than their theoretical values. To gauge the performance of the scale-shift policy in practice, we conduct a simulation study. We consider an MOOC that runs every quarter, where $N = 3,000$ assignments are submitted at the end of the course to be graded. We assume student graders with $(w_1, w_2) = \{0.01, 0.01\}$ and that the MOOC seeks to implement the appropriate scale shift policy at $k = 3$.

We consider eight $(2 \times 2 \times 2)$ different combinations of scenarios. The MOOC runs

- either one experiment (with $k = 2$) over one quarter or two experiments (with $k = 2$ and with $k = 4$) over two consecutive quarters for each course;
- out of the $3,000$ assignments, the instructor samples either $n = 30$ or $n = 100$ assignments and submits their grades (reflecting their true qualities);
- the MOOC then employs either $\hat{\xi}^{\text{mse}}$ or $\hat{\xi}^{\text{deb}}$.

For all scenarios, we conduct 1,000 iterations of the scale-shift policy implementation and report the average values and the standard deviations (shown in brackets) for their estimates in Table 1.

First, we observe that the averages of the estimates of $(\hat{w}_1, \hat{w}_2)$ are relatively close to the true values $(0.01, 0.01)$. However, while the standard deviation of $\hat{w}_1$ is relatively low (10% of the mean), that of $\hat{w}_2$ is high (100% of the mean). This is because the theoretical equilibrium values of $\alpha^\star$ and $\rho^\star$ are monotonic in w_1/c, whereas they are not in w_2/c (Corollary 3), which makes it easier to estimate the former. The variation in the estimation leads to variations in the scale-shift factors ξ^{mse}

Table 1 Scale-shift policy implementation results.

setup		estimates		minimize grading inaccuracy (MSE)					eliminate systematic grading bias				
$\mathcal{K}$	n	$\hat{w}_1$	$\hat{w}_2$	$\hat{\xi}^{\text{mse}}$	MSE	$e^\star$	μ	%P	$\hat{\xi}^{\text{deb}}$	MSE	$e^\star$	μ	%P
$\{2,3\}$	30	.010(.002)	.015(.016)	1.35(.18)	51.4(19.6)	9.7(2.9)	57.2	.88	1.59(.41)	117.9(90.6)	12.3(5.1)	59.2	.53
	100	.010(.001)	.014(.013)	1.33(.14)	43.7(10.8)	9.3(2.3)	57.0	.98	1.48(.32)	82.5(68.7)	11.1(4.2)	58.4	.68
$\{2,3,4\}$	30	.010(.002)	.015(.015)	1.34(.17)	49.2(16.7)	9.5(2.7)	57.1	.89	1.55(.39)	104.5(87.9)	11.8(4.8)	58.8	.61
	100	.010(.001)	.013(.011)	1.33(.13)	42.0(10.7)	9.4(2.1)	57.1	.97	1.46(.28)	72.7(60.2)	11.0(3.7)	58.4	.71
Theoretical ξ implementation:				1.32	31.8	6.9	54.3	-	1.38	33.8	7.3	54.8	-
Current practice ($\xi = 1$):				1	76.4	4.7	50	-	1	76.4	4.7	50	-

Note. $\mu = \mathbb{E}[q(e^*)] = 10\log(e^*) + 35$.

and ξ^{deb}. We observe that the average estimates of ξ^{mse} and ξ^{deb} are greater than the theoretical values. The variation in the estimates for ξ^{mse} is less than that for ξ^{deb} because, by Lemma 2, the former (which mimics the form $\frac{k\rho^\star}{\alpha^\star(1+(k-1)\rho^{\star 2})}$) is less sensitive to deviations in $\alpha^\star$ and $\rho^\star$ than the latter (which mimics the form $\frac{1}{\alpha^\star \rho_s^\star}$ and is therefore highly sensitive). We observe that as the number of experiments $|\mathcal{K}|$ or the number of instructor grades n increases, the variability in the estimates decreases, and the average values of parameters approach the theoretical values.

Due to the overestimation of ξ, we observe that both the average MSE and the average learning ($e^\star$) are greater than their theoretical values. Due to the convexity of the MSE with respect to ξ (Figure 6), the variability of $\hat{\xi}$ also contributes significantly to the increase in the expected MSE. Compared to the base case of $\xi = 1$, when the scale-shift factor $\hat{\xi}^{\text{mse}}$ is employed, improvement in both the MSE and learning (i.e., Pareto-efficient improvement) is achieved 88%–98% of the time. The MSE, while higher than the theoretical value (31.8), is significantly lower on average than when a scale-shift policy is not employed (76.4); the increase in the learning effort results in an increase in the average assignment quality (μ) of around 7 points. In contrast, when a scale-shift factor $\hat{\xi}^{\text{deb}}$ is employed, a Pareto-efficient improvement between 53% and 71% is achieved. The average MSE is higher than the current practice with a significant standard deviation; the expected improvement in learning effort results in an improvement of the average assignment quality by 8 to 9 points.

In summary, for the assumptions of this section and resulting Table 1, we can recommend practitioners to announce a peer grading scheme where a student's grade would be determined *not* by the simple average of peer grades, $\bar{G}$, but a scale shifted grade $\mu + 1.33 \cdot (\bar{G} - \mu)$. Here, μ is the predicted average grade (e.g., based on previous course outcomes), and the scale shift factor of 1.33 reflects ξ^{mse} (for any values of $\mathcal{K}$ and n considered). Doing so would induce more grading effort from the student graders and result in greater learning in MOOCs. A more sophisticated methods for eliciting or even influencing w_1 and w_2 can be further researched to make scale-shift grading scheme more effective for MOOCs to improve the performance of peer grading. Nevertheless, relative to the current peer grading practice, with moderate involvement of the instructors and experimentations run by MOOCs, we believe that the scale-shift policy can be reliably implemented to achieve simultaneous improvements in both peer grading accuracy and student learning outcomes.

6. Summary and Discussion

This paper treats information as a nonstatic variable (Arrow 1996), and examines how the cost of acquiring information affects economic behavior in the context of online education. We present a model of massive open online courses and develop insights into how peer grading performs under the assumption that student graders are rational economic agents. To the best of our knowledge, this is the first study to investigate how the accuracy of an MOOC's peer grading scheme is affected by graders' economic behavior with regard to information acquisition.

Investigating the current peer grading practice allowed us to identify a unique equilibrium information acquisition level and the reporting strategy of student graders. We present results regarding how course characteristics and the peer grading scheme affect not only this equilibrium, but also the inaccuracy of the peer grading (as measured by the mean squared error). We also find that peer grading exhibits a systematic grading bias toward the mean, a trend that discourages the students who observe it from exerting a learning effort during the course. Both grading inaccuracy and bias increase—and so learning is compromised—when student graders are strategic in reporting their assessments.

To improve the current peer grading practice, we introduce the scale-shift grading scheme. We demonstrate that this scheme can discourage strategic score reporting behavior, improve grading accuracy, and be used to eliminate bias (and perhaps even to create bias away from the mean)—improvements that will motivate student learning. We identify scale-shift factors that improve on the current practice along both dimensions and in a Pareto-efficient manner. Finally, our results establish that the scale-shift grading scheme is more effective when student graders are strategic (rather than honest) in reporting their assessments. We discuss how the grading scheme can be implemented and offer insights concerning its practical reliability.

Our model and insights offer some directions for future research. First, our results are based on the assumption that students are homogeneous in terms of their unfamiliarity with the subject of the assignments they are grading. In reality, however, some students might know the material well and thus be able to grade accurately and with less costly effort. Although we believe that such heterogeneity would not alter our main insights, understanding its effects is crucial for the fine-tuning of any peer grading policy. A second direction involves accurately estimating the parameters (including additional parameters for the case of heterogeneous graders). There is a rich literature in the field of computer science that addresses the effects of grader heterogeneity and how those effects can be estimated to improve the accuracy of peer grading mechanisms. It would be worthwhile to explore combining such methods with our economic insights about rational peer grading behavior in order to generate a deeper understanding of peer grading and how best to design a (scale-shift) peer grading scheme that can be implemented.

Acknowledgments

The second author is the corresponding author of this paper. We would like to thank the area editor, associate editor(s), and two anonymous reviewers whose thoughtful comments and suggestions helped improve the paper significantly during the review process.

References

Abbas, A.E. 2009. A Kullback-Leibler view of linear and log-linear pools. *Decision Analysis*, 6(1): 25–37.

Arrow, K.J. 1973. Information and economic behavior. *National Technical Information Service, US Department of Commerce*, Technical Report No. 14.

Arrow, K.J. 1984. *Collected Papers of Kenneth J. Arrow, Volume 4: The Economics of Information.* Harvard University Press, Cambridge, MA.

Arrow, K.J. 1996. The economics of information: An exposition. *Empirica*, 23: 119–128.

Arrow, K.J., R. Radner. 1979. Allocation of resources in large teams. *Econometrica*, 47(2): 361–385.

Balfour, S.P. 2013. Assessing writing in MOOCs: automated essay scoring and calibrated peer review. *Research & Practice in Assessment*, 8(1): 40–48.

Brinton, C.G., M. Chiang, S. Jain, H. Lam, Z. Liu, F.M.F. Wong. 2014. Learning about social learning in MOOCs: From statistical analysis to generative model. *IEEE Trans. on Learning Technologies*, 7(4): 346–359.

Bordley, R.F. 1982. A multiplicative formula for aggregating probability assessments. *Management Science*, 28(10): 1137–1148.

Calvo-Armengol A., E. Patacchini, Y. Zenou. 2009. Peer effects and social networks in education. *Review of Economics Studies*, 76(4): 1239–1267.

Calsamiglia, C., A. Loviglio. 2019. Grading on a curve: When having good peers is not good. *Economics of Education Review,* 73: 101916.

Cho, K., C.D. Schunn, R.W. Wilson. 2006. Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, 98(4): 891–901.

Clemen, R.T. 1986. Calibration and the aggregation of probabilities. *Management Science*, 32(3): 312–314.

Clemen, R.T., R.L. Winkler. 1990. Unanimity and compromise among probability forecasters. *Management Science*, 36(7): 767–779.

Coursera Help Center. 2018. https://learner.coursera.help/hc/en-us/articles/208279946-See-feedback-and-grades-for-peer-reviewed-assignments.

Dasgupta, A., A. Ghosh. 2013. Crowdsourced judgement elicitation with endogenous proficiency. *Proceedings of the 22nd international conference on World Wide Web*, 319–330. Rio de Janeiro, Brazil, ACM.

DeGroot, M.H. 1974. Reaching a consensus. 1974. *Journal of the American Statistical Association*, 69(345): 118–121.

DeGroot, M.H., J. Mortera. 1991. Optimal linear opinion pools. *Management Science*, 37(5): 546–558.

Eliashberg, J., R.L. Winkler. 1981. Risk sharing and group decision making. *Management Science*, 27(11): 1221–1235.

Falk A, A. Ichino. 2006. Clean evidence on peer effects. *Journal of Labor Economics*. 24(1): 39–57.

Falchikov, N., J. Goldfinch. 2000. Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Review of Educational Research*. 24(1): 287–322.

Gallus, J. 2017. Fostering public good contributions with symbolic awards: a large-scale natural field experiment at Wikipedia. *Management Science*, 63(12): 3999–4015.

Granger, C.W. 1989. Invited review: Combining forecasts – twenty years later. *Journal of Forecasting*, 8(3): 167–173.

Grant, A. 2016. Why we should stop grading students on a curve. NY Times, September 10, 2016. Sunday Review. http://nyti.ms/2cCFFfO.

Howard, R.A. 1966. Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, 2(1): 22–26.

Kamble, V., D. Marn, N. Shah, A. Parekh, K. Ramchandran. 2018. A truth serum for large-scale evaluations. Working paper, arXiv:1507.07045v4.

Keynes, J.M. 1936. *The General Theory of Employment, Interest and Money.* New York: Harcourt Brace and Co.

Khalil, H., M. Ebner. 2014. MOOCs completion rates and possible methods to improve retention - a literature review. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 1236–1244. Chesapeake, VA: AACE.

Li, H.L., Y. Xiong, X.J. Zang, M. Kornbaher, Y.S. Lyu, K.S. Chung, H.K. Suen. 2016. Peer assessment in a digital age: A meta-analysis comparing peer and teacher ratings. *Assessment & Evaluation in Higher Education*, 41(2): 245–264.

Luaces, O., J. Diez, A. Alonso-Betanzos, A. Troncoso, A. Bahamonde. 2015. A factorization approach to evaluate open-response assignments in MOOCs using preference learning on peer assessments. *Knowledge-Based Systems*, 85, 322–328.

Luo, H., A.C. Robinson, J.Y. Park. 2014. Peer grading in a MOOC: reliability, validity, and perceived effects. *Online Learning: Official Journal of the Online Learning Consortium*, 18(2): 1–14.

Michaels, J.W. 1976. A Simple view of the grading issue. *Teaching Sociology,* 3(2): 198–203.

Miller, N., P. Resnick, R. Zeckhauser. 2005. Eliciting informative feedback: the peer-prediction method. *Management Science*, 51(9): 1359–1373.

Morris, S., H.S. Shin. 2002. Social value of public information. *American Economic Review*, 92(5): 1521–1534.

Piech, C., J. Huang, Z. Chen, C. Do, A. Ng, D. Koller. 2013. Tuned models of peer assessment in MOOCs, *The 6th International Conference on Educational Data Mining*, Memphis, Tennessee, USA.

Prelec, D. 2004. A Bayesian truth serum for subjective data. *Science*, 306(5695): 462-466.

Price, E., F. Goldberg, S. Robinson, M. McKean. 2016. Validity of peer grading using Calibrated Peer Review in a guided-inquiry, conceptual physics course *Physical Review Physics Education Research*, 12(2): 020145.

Radanovic, G., B. Faltings. 2015. Incentives for subjective evaluations with private beliefs. *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, 1014–1020. AAAI.

Raman, K., T. Joachims. 2014. Methods for ordinal peer grading. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1037-1046. ACM.

Sacerdote, B. 2001. Peer effects with random assignment: results for Dartmouth roommates. *Quarterly Journal of Economics*, 116(2): 681–704.

Satopää, V. A. 2019. Collecting information from multiple forecasters: inefficiency of central tendency. *Working paper.*

Satopää, V. A., J. Baron, D.P. Foster, B.A. Mellers, P.E. Tetlock, L.H. Ungar. 2014 Combining multiple probability predictions using a simple logit model. *International Journal of Forecasting*, 30(2): 344–356.

Savva, N., T. Tezcan, O. Yildiz. 2019. Can yardstick competition reduce waiting times? *Management Science*, 65(7): 3196–3215.

Shah, N.B., J.K. Bradley, A. Parekh, M. Wainwright, K. Ramchandran. 2013. A case for ordinal peer-evaluation in MOOCs, *NIPS Workshop on Data Driven Education*, Dec. 2013.

Shleifer, A. 1985. A theory of yardstick competition. *RAND Journal of Economics*, 16(3): 319–327.

Shnayder, V., A. Agarwal, F. Frongillo, D. Parkes. 2016. Informed truthfulness in multi-task peer prediction. *Proceedings of the 2016 ACM Conference on Economics and Computation*, 179–196. Maastricht, The Netherlands.

Walsh, T. 2014. The PeerRank method for peer assessment. *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI)*, 909–914.

Young, Jeffery, R. 2020. Will COVID-19 lead to another MOOC moment? *Ed Surge, Higher Education*, March 25, 2020. https://www.edsurge.com/news/2020-03-25-will-covid-19-lead-to-another-mooc-moment

Zhang, D., G. Allon, J.A. Van Mieghem. 2017. Does social interaction improve learning outcomes? Evidence from field experiments on massive open online courses. *Manufacturing & Services Operations Management.* 19(3): 347–367.

Zimmerman, D.J. 2003. Peer effects in academic outcomes: evidence from a natural experiment. *Review of Economic and Statistics*, 85(1): 9–23.

Appendix

Proof of Lemma 1. We first focus on deriving the first cost function. Without any information, assessment is $\mathcal{A} \sim \mathcal{N}(\mu, \sigma^2)$, and the Shannon entropy is

$$\begin{aligned}\mathbb{E}[-\log(f(\mathcal{A}))] &= \int_{-\infty}^{\infty} -\log\left(\frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right)\right) f(x)\,dx \\ &= \log\left(\sqrt{2\pi}\sigma\right)\int_{-\infty}^{\infty} f(x)\,dx - \int_{-\infty}^{\infty}\frac{(x-\mu)^2}{2\sigma^2} f(x)\,dx = \log(\sigma) + \frac{\log(2\pi)-1}{2}.\end{aligned}$$

With information level ρ, we have $\mathcal{A}|\rho \sim \mathcal{N}(\mu + \rho(q-\mu), (1-\rho^2)\sigma^2)$, and the Shannon entropy is

$$\mathbb{E}[-\log(f(\mathcal{A}|\rho))] = \log\left(\sqrt{1-\rho^2}\sigma\right) + \frac{\log(2\pi)-1}{2}.$$

Applying Arrow's cost function now yields

$$\mathcal{C}(\rho) = \bar{c}(\mathbb{E}[-\log(f(\mathcal{A}))] - \mathbb{E}[-\log(f(\mathcal{A}|\rho))]) = \bar{c}\left(\log(\sigma) - \log\left(\sqrt{1-\rho^2}\sigma\right)\right) = c\log(1-\rho^2),$$

where $c \triangleq \bar{c}/2$.

Proof of Proposition 1 (and Corollary 1). We prove this proposition for general ξ, or for $\hat{G} = \mu + \xi(\bar{G} - \mu)$. All results hold when $\xi = 1$. We first consider the honest type, then the strategic type of graders.

(i) Honest Reporting. First we analyze the setting in which $\alpha = 1$ and $G_i = \mathcal{A}_i$ for all i. With regard to *absolute* inaccuracy: since $G_i|q$ is Normally distributed with mean $(1-\rho_i)\mu + \rho_i q$ and variance $(1-\rho_i^2)\sigma^2$, it follows that

$$\mathbb{E}[(Q-G_i)^2] = \mathbb{E}\left\{\mathbb{E}\left[(Q-G_i)^2|Q\right]\right\} = \mathbb{E}[(1-\rho_i)^2(\mu-Q)^2 + (1-\rho_i^2)\sigma^2] = 2(1-\rho_i)\sigma^2.$$

As for *relative* inaccuracy, we have

$$\xi\bar{G} - G_i = \frac{\xi}{k}\sum_{\forall\ell} G_\ell(\rho_\ell) - G_i(\rho_i) = \frac{\xi}{k}\sum_{\ell\neq i} G_\ell(\rho_\ell) + \left(\frac{\xi}{k} - 1\right) G_i(\rho_i)$$

Conditional on $Q = q$, the $G_i|q, i = 1, 2, \ldots, k$ are independent Normal distributions $\mathcal{N}((1-\rho_i)\mu + \rho_i q, (1-\rho_i)\sigma^2)$. Therefore,

$$\hat{G} - G_i|q \sim \mathcal{N}\left((q-\mu)\left(\frac{\xi\sum_{\forall\ell}\rho_\ell}{k} - \rho_i\right),\ \frac{\xi^2\sigma^2}{k^2}\sum_{\ell\neq i}(1-\rho_\ell^2) + \left(\frac{\xi-k}{k}\right)^2\sigma^2(1-\rho_i^2)\right).$$

and

$$\begin{aligned}\mathbb{E}\left[(\hat{G}-G_i)^2|q\right] &= Var[\hat{G}-G_i|q] + \left(\mathbb{E}[\hat{G}-G_i|q]\right)^2 \\ &= \frac{\xi^2\sigma^2}{k^2}\sum_{\ell\neq i}(1-\rho_\ell^2) + \left(\frac{k-\xi}{k}\right)^2\sigma^2(1-\rho_i^2) + (q-\mu)^2\left(\frac{\xi\sum_{\forall\ell}\rho_\ell}{k} - \rho_i\right)^2.\end{aligned}$$

A priori, student i does not know the assignment's quality. Hence,

$$\begin{aligned}\mathbb{E}[(\hat{G}-G_i)^2] &= \mathbb{E}\left\{\mathbb{E}\left[(\hat{G}-G_i)^2|Q\right]\right\} \\ &= \mathbb{E}\left\{\frac{\xi\sigma^2}{k^2}\sum_{\ell\neq i}(1-\rho_\ell^2)+\left(\frac{k-\xi}{k}\right)^2\sigma^2(1-\rho_i^2)+(Q-\mu)^2\left(\frac{\xi\sum_{\forall\ell}\rho_\ell}{k}-\rho_i\right)^2\right\} \\ &= \frac{\xi^2\sigma^2}{k^2}\sum_{\ell\neq i}(1-\rho_\ell^2)+\left(\frac{k-\xi}{k}\right)^2\sigma^2(1-\rho_i^2)+\mathbb{E}\Big\{(Q-\mu)^2\Big\}\left(\frac{\xi\sum_{\forall\ell}\rho_\ell}{k}-\rho_i\right)^2 \\ &= \sigma^2\left(\frac{\xi^2}{k^2}\sum_{\ell\neq i}(1-\rho_\ell^2)+\left(\frac{k-\xi}{k}\right)^2(1-\rho_i^2)+\left(\frac{\xi\sum_{\forall\ell}\rho_\ell}{k}-\rho_i\right)^2\right).\end{aligned}$$

For a given assignment, grader i (of k peer graders) acquires the information level ρ_i that minimizes $\mathcal{C}(\rho_i)+w_1\mathbb{E}[(G_i-Q)^2]+w_2\mathbb{E}\left[(G_i-\hat{G})^2\right]$, or

$$\begin{aligned}\min_{\rho_i}\; & -c\log(1-\rho_i^2)+w_1(1-\rho_i)\sigma^2 \\ & \qquad +w_2\sigma^2\left[\frac{\xi^2}{k^2}\sum_{\ell\neq i}(1-\rho_\ell^2)+\left(\frac{k-\xi}{k}\right)^2(1-\rho_i^2)+\left(\frac{\xi\sum_{\forall\ell}\rho_\ell}{k}-\rho_i\right)^2\right].\end{aligned}$$

The second-order derivative with respect to ρ_i is $2c(1+\rho_i^2)/(1-\rho_i^2)^2>0$; the objective is therefore a convex function in ρ_i. The first-order condition (FOC) is

$$\frac{2c\rho_i}{1-\rho_i^2}-w_1\sigma^2+2w_2\sigma^2\frac{\xi\sum_{l\neq i}\rho_l}{k}\frac{\xi-k}{k}=0. \tag{A-1}$$

We search for symmetric equilibrium such that

$$\frac{2c\rho_i}{1-\rho_i^2}=w_1\sigma^2+2w_2\sigma^2\rho_i\frac{\xi(k-\xi)(k-1)}{k^2}.$$

This equilibrium is the solution on $[0,1]$ of a cubic function

$$F(\rho)\triangleq ck^2\rho-(1-\rho^2)\sigma^2(k^2w_1+(k-1)(k-\xi)\xi\rho w_2)=0. \tag{A-2}$$

When $w_1=0$, one of the equilibria is $\rho_h^\star=0$. If $w_2>ck^2/(\xi(k-1)(k-\xi)\sigma^2)$, then there is another symmetric equilibrium $\rho_h^\star=\left(1-\frac{ck^2}{\xi(k-1)(k-\xi)w_2\sigma^2}\right)^{1/2}$. If $w_1>0$, then $F(0)=-w_1k^2\sigma^2<0$ and $F(1)=ck^2>0$, which guarantees the existence of at least one solution on $(0,1)$.

$$\begin{aligned}\frac{\partial F}{\partial\rho}&=ck^2+\sigma^2((3\rho^2-1)(k-1)(k-\xi)\xi w_2+2k^2\rho w_1), \\ \rho\frac{\partial F}{\partial\rho}-F&=\sigma^2(2(k-1)(k-\xi)\xi\rho^3w_2+k^2(1+\rho^2)w_1)>0.\end{aligned}$$

As a result, $\frac{\partial F}{\partial\rho}>0$ provided that $F(\rho)=0$. We conclude that $F(\rho)$ cannot have more than one zero point on $[0,1]$ and that there is a unique zero point $\rho_h^\star$. At the optimal solution $\rho_h^\star$, we have

$F^{'}(\rho_h^\star) > 0$. In other words, as long as $w_1 > 0$, the symmetric equilibrium $\rho_h^\star$ is the unique solution to the cubic function (A-2).

By the implicit function theorem,

$$\frac{\partial \rho_h^\star}{\partial c} = - \left.\frac{\partial F/\partial c}{\partial F/\partial \rho}\right|_{\rho=\rho_h^\star} < 0.$$

Since

$$\frac{\partial F(\rho)}{\partial w_1} = -k^2(1-\rho^2)\sigma^2 < 0 \quad \text{and} \quad \frac{\partial F(\rho)}{\partial w_2} = -(k-1)(k-\xi)\rho(1-\rho^2)\sigma^2 < 0,$$

it follows that

$$\frac{d\rho_h^\star}{dw_1} = - \left.\frac{\partial F(\rho)/\partial w_1}{\partial F(\rho)/\partial \rho}\right|_{\rho=\rho_h^\star} > 0 \quad \text{and} \quad \frac{d\rho_h^\star}{dw_2} = - \left.\frac{\partial F(\rho)/\partial w_2}{\partial F(\rho)/\partial \rho}\right|_{\rho=\rho_h^\star} > 0.$$

That means that the optimal solution $\rho_h^\star$ increases with both w_1 and w_2. Similarly, $\frac{\partial F(\rho)}{\partial \sigma} = -2(1-\rho^2)\sigma(w_2\xi(k-1)(k-\xi)\rho + w_1k^2) < 0$, which implies that $\rho_h^\star$ increases with σ, and $\frac{\partial F(\rho)}{\partial c} = k^2\rho > 0$, which implies that $\rho_h^\star$ decreases with c. We are now in a position to check the monotonicity of $\rho_h^\star$ on k. From the equality

$$\frac{\partial F(\rho)}{\partial k} = -2(1-\rho^2)\sigma(w_2\xi(k-1)(k-\xi)\rho + w_1k^2),$$

we can derive

$$2F(\rho) - k\frac{\partial F(\rho)}{\partial k} = \rho(1-\rho^2)\xi(k+(k-2)\xi)w_2\sigma^2 > 0.$$

Since $F(\rho_h^\star) = 0$, it follows that $\frac{\partial F(\rho_h^\star)}{\partial k} < 0$. Therefore, $\rho_h^\star$ is increasing in k.

(ii) Strategic Reporting. Now we analyze the case where $G_j = (1-\alpha_j)\mu + \alpha_j\mathcal{A}_j, \alpha_j \in [0,1]$ is the score that student j reports. The reported grade is $(1-\xi)\mu + \xi\bar{G}$. Given all other graders chooses ρ, α, the objective of student 1 choosing ρ_1, α_1 is

$$\min_{\rho_1,\alpha_1} O = -c\log(1-\rho_1^2) + w_1\mathbb{E}[(G_1-Q)^2] + w_2\mathbb{E}\big[\big((1-\xi)\mu + \xi(G_2+\ldots+G_k)/k - G_1(k-\xi)/k\big)^2\big].$$

We know $\mathcal{A}_1|q \sim \mathcal{N}((1-\rho_1)\mu + \rho_1 q, (1-\rho_1^2)\sigma^2)$, which means that $G_1 - q|q \sim \mathcal{N}\big((1-\alpha_1\rho_1)\mu + \alpha_1\rho_1 q - q, \alpha_1^2(1-\rho_1^2)\sigma^2\big)$. To find the absolute inaccuracy, we have

$$\mathbb{E}[(G_1 - q|q)^2] = ((1-\alpha_1\rho_1)\mu + \alpha_1\rho_1 q - q)^2 + \alpha_1^2(1-\rho_1^2)\sigma^2.$$

Therefore,

$$\begin{aligned}\mathbb{E}[(G_1-Q)^2] = \mathbb{E}[\mathbb{E}[(G_1-Q)^2|Q]] &= ((1-\alpha_1\rho_1)\mu + \alpha_1\rho_1\mu - \mu)^2 + (1-\alpha_1\rho_1)^2\sigma^2 + \alpha_1^2(1-\rho_1^2)\sigma^2 \\ &= (1-2\alpha_1\rho_1+\alpha_1^2)\sigma^2.\end{aligned}$$

Next, to find the relative inaccuracy,

$$\mathbb{E}[\hat{G}-G_1|q] = \left(\alpha\rho\xi\frac{k-1}{k} - \alpha_1\rho_1\frac{k-\xi}{k}\right)(q-\mu),$$
$$Var[\hat{G}-G_1|q] = \left(\frac{(k-\xi)^2}{k^2}\alpha_1^2(1-\rho_1^2) + \frac{k-1}{k^2}\xi^2\alpha^2(1-\rho^2)\right)\sigma^2.$$

Therefore,

$$\begin{aligned}\mathbb{E}[(\hat{G}-G_1)^2] &= \mathbb{E}\left[(\mathbb{E}[\hat{G}-G_1|Q])^2 + Var[\hat{G}-G_1|Q]\right]\\ &= \left(\alpha_1\rho_1\frac{k-\xi}{k} - \alpha\rho\xi\frac{k-1}{k}\right)^2\mathbb{E}[(\mu-Q)^2] + \left(\frac{(k-\xi)^2}{k^2}\alpha_1^2(1-\rho_1^2) + \frac{k-1}{k^2}\xi^2\alpha^2(1-\rho^2)\right)\sigma^2\\ &= \left((\alpha_1\rho_1(k-\xi)^2 - \alpha\rho\xi(k-1))^2 + (k-\xi)^2\alpha_1^2(1-\rho_1^2) + (k-1)\xi^2\alpha^2(1-\rho^2)\right)\frac{\sigma^2}{k^2}.\end{aligned}$$

Since $\frac{\partial^2 O}{\partial\alpha_1^2} = \frac{2\sigma^2(w_2(k-\xi)^2 + w_1k^2)}{k^2} > 0$, the objective is convex. Given ρ_1, ρ, and α, from the FOC with respect to α_1, we can solve out

$$\alpha_1 = \rho_1\frac{\alpha\rho w_2\xi(k-1)(k-\xi) + w_1k^2}{w_2(k-\xi)^2 + w_1k^2}. \tag{A-3}$$

If we substitute α_1 into the objective function, then the second-order derivative with respect to ρ_1 is

$$\frac{2c(1+\rho_1^2)}{(1-\rho_1^2)^2} - \frac{2\sigma^2(w_2\rho\alpha(k-1)(k-\xi)\xi + w_1k^2)^2}{w_1k^4 + w_2k^2(k-\xi)^2};$$

this expression is positive provided that

$$c > \frac{\sigma^2(1-\rho_1^2)^2(w_2\rho\alpha(k-1)(k-\xi)\xi + w_1k^2)^2}{k^2(1+\rho_1^2)(w_1k^2 + w_2(k-\xi)^2)}. \tag{A-4}$$

Note that the positive definite Hessian matrix requires the same above condition. The FOC with respect to ρ_1 is

$$\frac{2c\rho_1}{1-\rho_1^2} - 2\alpha_1\sigma^2\left(w_1 + \frac{w_2\alpha\rho\xi(k-1)(k-\xi)}{k^2}\right) = 0. \tag{A-5}$$

Assuming symmetric equilibrium, $\rho_1 = \rho$ and $\alpha_1 = \alpha$. Using (A-3), we solve out

$$\alpha = \frac{\rho w_1k^2}{w_1k^2 + w_2(k-\xi)(k(1-\rho^2\xi) - (1-\rho^2)\xi)}.$$

Now suppose that α is an interior solution. Then, substituting α into (A-5) gives

$$\frac{2\rho}{1-\rho^2}\left(c - \frac{k^2\sigma^2w_1^2(1-\rho^2)(w_2(k-\xi)^2 + w_1k^2)}{(w_2(k-\xi)(k(1-\rho^2\xi) - (1-\rho^2)\xi) + w_1k^2)^2}\right) = 0,$$

an equation with three roots. The local minimality of $\rho = 0$ requires that (A-4) be satisfied for $\rho_1 = \rho = 0$. That is, we must have

$$c > \frac{\sigma^2w_1^2k^2}{w_2(k-\xi)^2 + w_1k^2}.$$

We now check that $\rho_1 = 0$ is indeed the minimizer under the above condition. When $\rho = \alpha = 0$, (A-3) becomes $\alpha_1 = \rho_1 \frac{w_1 k^2}{w_1 k^2 + w_2(k-\xi)^2}$. The objective becomes

$$-c\log(1-\rho_1^2) + w_1\sigma^2 \frac{w_2(k-\xi)^2 + w_1 k^2(1-\rho_1^2)}{w_2(k-1)^2 + w_1 k^2}.$$

The first-order derivative with respect to ρ_1 is

$$2\rho_1\left(\frac{c}{1-\rho_1^2} - \frac{\sigma^2 w_1^2 k^2}{w_2(k-\xi)^2 + w_1 k^2}\right),$$

which is positive on $(0,1)$ provided that $c \geq \frac{\sigma^2 w_1^2 k^2}{w_2(k-\xi)^2 + w_1 k^2}$. Therefore, the optimal $\rho_1 = 0$. Yet if $c < \frac{\sigma^2 w_1^2 k^2}{w_2(k-\xi)^2 + w_1 k^2}$, then the first-order derivative is negative at $\rho_1 = 0$, and so $\rho = 0$ is *not* an equilibrium.

We continue checking the FOC's other roots. Let $A_\xi(\rho) \triangleq \frac{k^2\sigma^2 w_1^2(1-\rho^2)(w_2(k-\xi)^2 + w_1 k^2)}{(w_2(k-\xi)(k(1-\rho^2\xi)-(1-\rho^2)\xi)+w_1 k^2)^2}$. Since $A_\xi(1) = 0$ and $A_\xi(\rho) \geq 0$ on $[0,1]$, it follows that $A_\xi(\rho) = c$ must has a solution in the range $(0,1)$ as long as

$$0 < c < A_\xi(0) = \frac{\sigma^2 w_1^2 k^2}{w_2(k-\xi)^2 + w_1 k^2}.$$

Denoting ρ^2 by ρ_2; then algebra shows that $A_\xi(\rho) - c = 0$ is a quadratic function of ρ_2 with two roots:

$$\rho_\pm = \frac{2cw_2(k-1)(k-\xi)\xi(w_2(k-\xi)^2 + w_1k^2) \pm k^{3/2} w_1\sigma\left(\sqrt{\Delta} - \sqrt{k}\sigma w_1(w_2(k-\xi)^2 + w_1 k^2)\right)}{2cw_2^2(k-1)^2(k-\xi)^2\xi^2},$$

where

$$\Delta = (w_2(k-\xi)^2 + w_1k^2)\big(k\sigma^2 w_1^2(w_2(k-\xi)^2 + w_1k^2) + 4c(k-1)(k-\xi)\xi w_2(w_2(k-\xi)(\xi-1) - kw_1)\big).$$

Using the subscripts "+" and "−" to denote (respectively) the larger and smaller root, we can write their product as

$$\rho_+\rho_- = \frac{((k-\xi)^2 w_2 + k^2 w_1)\,(c((k-\xi)^2 w_2 + k^2 w_1) - k^2\sigma^2 w_1^2)}{c(k-1)^2(k-\xi)^2\xi^2 w_2^2}.$$

Since $c < \frac{\sigma^2 w_1^2 k^2}{w_2(k-\xi)^2 + w_1 k^2}$, we have $\rho_+\rho_- < 0$. It follows that $\rho_+ > 0$ and $\rho_- < 0$ and also that ρ_+ is the unique solution in $(0,1)$. Therefore, the unique equilibrium is $\rho_s^\star = \sqrt{\rho_+}$ if

$$\alpha^\star = \frac{\rho_s^\star w_1 k^2}{w_1k^2 + w_2(k-\xi)(k(1-\rho_s^{\star 2}\xi) - (1-\rho_s^{\star 2})\xi)} \in (0,1).$$

Note that when $\xi = 1$, we have $\alpha^\star = \frac{w_1 k^2 \rho_s^\star}{w_1 k^2 + w_2(k-1)^2(1-\rho_s^{\star 2})} \leq \rho_s^\star$.

When ξ is general, $\alpha^\star \in (0,1)$ is equivalent to

$$H(\rho_s^\star) \triangleq w_2\xi(k-1)(k-\xi)\rho_s^{\star 2} + w_1k^2\rho_s^\star - w_1k^2 - w_2(k-\xi)^2 < 0. \tag{A-6}$$

If $\xi \geq k$, $H(\rho_s^\star) < 0$ when $\rho_s^\star \in [0,1]$. If $\xi < k$, $H(\rho_s^\star)$ has one positive and one negative roots. (A-6) requires $\rho_s^\star$ to be in between of the two roots, and therefore

$$0 \leq \rho_s^\star < \tilde{\rho} \triangleq \frac{\sqrt{k^4 w_1^2 + 4\xi(k-1)(k-\xi)w_2((k-\xi)^2 w_2 + k^2 w_1)} - k^2 w_1}{2\xi(k-1)(k-\xi)w_2}. \tag{A-7}$$

When $\xi < k$ (including $\xi = 1$), we have

$$\frac{\partial \alpha^\star}{\partial \rho_s^\star} = \frac{w_1 k^2((k-\xi)(k-\xi+(k-1)\xi\rho_s^{\star 2})w_2 + k^2 w_1)}{(w_1 k^2 + w_2(k-\xi)(k(1-\rho_s^{\star 2}\xi) - (1-\rho_s^{\star 2})\xi))^2} > 0.$$

If $\rho_s^\star = 0$ then $\alpha^\star = 0$, and if $\rho_s^\star = 1$ then $\alpha^\star = \frac{kw_1}{kw_1 - (\xi-1)(k-\xi)w_2}$. If $0 < \xi \leq 1$, then $0 < \alpha^\star \leq 1$, which implies that $\tilde{\rho} \geq 1$; if $1 < \xi < k$ then, as $\rho_s^\star$ increases from 0 to 1, the value of $\alpha^\star$ either increases to infinity before $\rho_s^\star$ reaches 1, or increases continuously to exceed 1 at $\rho_s^\star = 1$, which implies that $\tilde{\rho} < 1$. The condition (A-7) confirms $\alpha^\star$ is increasing in $\rho_s^\star$ on $[0, \tilde{\rho})$.

Next we prove that $\rho_s^\star$ is decreasing in c on $[0, \tilde{\rho})$, which is equivalent to showing that $A_\xi'(\rho) < 0$.

$$A_\xi'(\rho) = \frac{2k^2\rho\sigma^2 w_1^2((k-\xi)^2 w_2 + k^2 w_1)A_1}{A_2^3},$$

where $A_1 = -(k-\xi)(k + (1-\rho^2)\xi + k(\rho^2 - 2)\xi)w_2 - k^2 w_1$, and $A_2 = k^2 w_1 - (k-\xi)(\xi(1-\rho^2) + k(\rho^2\xi - 1))w_2$. We have

$$A_1 + A_2 = -2(1-\rho^2)(k-\xi)(k-1)\xi w_2 < 0.$$

Therefore, A_1 and A_2 cannot both be positive. Next, $A_1 < 0$ requires $w_1 > \frac{(k-\xi)w_2}{k^2}(k(2-\rho^2)\xi - \xi(1-\rho^2) - k)$; $A_2 < 0$ requires $w_1 < \frac{(k-\xi)w_2}{k^2}((\xi(1-\rho^2) + k(\rho^2\xi - 1))w_2)$. Since $\frac{(k-\xi)w_2}{k^2}(k(2-\rho^2)\xi - \xi(1-\rho^2) - k) - \frac{(k-\xi)w_2}{k^2}((\xi(1-\rho^2) + k(\rho^2\xi - 1))w_2) = 2\frac{(k-\xi)w_2}{k^2}\xi w_2(k-1)(1-\rho^2) > 0$. Therefore, it is impossible to have $A_1 < 0$ and $A_2 < 0$. That means, A_1 and A_2 must have opposite signs and $A_\xi'(\rho) < 0$. Condition (A-7) is therefore equivalent to $c > \tilde{c}$, where

$$\tilde{c} = A_\xi(\tilde{\rho}).$$

Similarly, we investigate the monotonicity of $\rho_s^\star$ on σ and w_1 . Because A_ξ is increasing in σ, the implicit function theorem implies that $\frac{\partial \rho_s^\star}{\partial \sigma} < 0$. When $\xi = 1$,

$$\frac{\partial A_\xi}{\partial w_1} = \frac{k^2(1-\rho^2)\sigma^2 w_1(k^4 w_1^2 + 2(k-1)^4(1-\rho^2)w_2^2 + 3(k-1)^2 k^2(1-\rho^2)w_1 w_2)}{((k-1)^2(1-\rho^2)w_2 + k^2 w_1)^3} > 0.$$

Therefore, $\frac{\partial \rho_s^\star}{\partial w_1} < 0$ when $\xi = 1$.

We further investigate the corner solution where $\alpha^\star = 1$, in which case the equilibrium would be exactly the same as that in the honest reporting setting. Here $\rho_h^\star$ is the solution on $[0,1]$ of the cubic function (A-2). Given that all others choose $\alpha = 1$, sustaining this corner solution requires that (A-3) be no smaller than 1:

$$\rho_h^\star \frac{\rho_h^\star w_2 \xi (k-1)(k-\xi) + w_1 k^2}{w_2 (k-\xi)^2 + w_1 k^2} \geq 1 \iff w_2 \xi (k-1)(k-\xi) \rho_h^{\star 2} - w_1 k^2 (1-\rho_h^\star) - (k-\xi)^2 w_2 \geq 0.$$

This statement is complementary to (A-6), and it is equivalent to

$$1 > \rho_h^\star \geq \tilde{\rho}. \tag{A-8}$$

From

$$F(\tilde{\rho}) = ck^2 \tilde{\rho} - (1-\tilde{\rho}^2)\sigma^2 (k^2 w_1 + (k-1)(k-\xi)\xi \tilde{\rho} w_2) = 0,$$

we find the corresponding c as follows:

$$\tilde{c} = \frac{(1-\tilde{\rho}^2)\sigma^2 (k^2 w_1 + (k-1)(k-\xi)\xi \tilde{\rho} w_2)}{k^2 \tilde{\rho}}$$

It is easy to verify that the above $\tilde{c}$ is exactly $A_\xi(\tilde{\rho})$, which makes sure the conditions are consistent.

We can summarize our results so far as follows.

1. If $A_\xi(0) \leq A_\xi(\tilde{\rho})$:
   - if $0 < c < A_\xi(0)$, then the equilibrium is $\alpha^\star = 1$, $\rho_s^\star = \rho_h^\star$;
   - if $c > A_\xi(0)$, then $\alpha^\star = \rho_s^\star = 0$ is one equilibrium.
2. If $A_\xi(0) > A_\xi(\tilde{\rho})$:
   - if $0 < c \leq A_\xi(\tilde{\rho})$, then the equilibrium is $\alpha^\star = 1$, $\rho_s^\star = \rho_h^\star$;
   - if $A_\xi(\tilde{\rho}) < c < A_\xi(0)$, then the equilibrium is $\alpha^\star < 1$, $\rho_s^\star = \sqrt{\rho_+}$;
   - if $c > A_\xi(0)$, then $\alpha^\star = \rho_s^\star = 0$ is one equilibrium.

Next we show that, if $\alpha^\star < 1$, then $\rho_s^\star < \rho_h^\star$. Since all students in this scenario choose $\alpha^\star$, it follows that $\rho_s^\star$ should satisfy the FOC with respect to ρ. That is:

$$F(\rho, \alpha^\star) \triangleq ck^2 \rho - \alpha^\star (1-\rho^2)\sigma^2 (\alpha^\star (k-1)(k-\xi)\xi \rho w_2 + k^2 w_1) = 0.$$

Since $F(0, \alpha^\star) = -\alpha^\star k^2 \sigma^2 w_1 < 0$ and $F(1, \alpha^\star) = ck^2 > 0$, there exists at least one zero point.

$$\begin{aligned}
\frac{\partial F}{\partial \alpha^\star} &= -(1-\rho^2)\sigma^2 (2\alpha^\star (k-1)(k-\xi)\xi \rho w_2 + k^2 w_1) < 0, \\
\frac{\partial F}{\partial \rho} &= ck^2 + \alpha^\star \sigma^2 ((3\rho^2 - 1)\alpha^\star (k-1)(k-\xi)\xi w_2 + 2k^2 \rho w_1), \\
\rho \frac{\partial F}{\partial \rho} - F &= \alpha^\star \sigma^2 (2\alpha^\star (k-1)(k-\xi)\xi \rho^3 w_2 + k^2 (1+\rho^2) w_1) > 0.
\end{aligned}$$

Hence $\frac{\partial F}{\partial \rho} > 0$ provided that $F(\rho, \alpha^\star) = 0$. There can be no more than a single zero point on $[0,1]$, and the unique zero point is $\rho_s^\star$. From the implicit function theorem, we have

$$\frac{\partial \rho_s^\star}{\partial \alpha^\star} = - \left. \frac{\partial F / \partial \alpha^\star}{\partial F / \partial \rho} \right|_{\rho = \rho_s^\star} > 0.$$

Therefore, if $\alpha^\star < 1$, then we have $\rho_s^\star < \rho_h^\star$.

Finally, we explore the boundary parameters. In the case of honest reporting, if $w_2 = 0$ and $w_1 > 0$ then solving $F(\rho) = 0$ in (A-2) gives

$$\rho = \frac{\sqrt{c^2 + 4w_1^2\sigma^4} - c}{2w_1\sigma^2}.$$

If $w_1 = 0$ and $w_2 > \frac{ck^2}{\sigma^2(k-1)^2}$, then $F(\rho) > 0$ on $[0,1]$. In this case, the equilibrium grading effort level is 0; otherwise, solving $F(\rho) = 0$ yields $\rho = 0$ and $\rho = \sqrt{1 - \frac{ck^2}{w\sigma^2\xi(k-\xi)(k-1)}}$— each of which is an equilibrium.

In the case of strategic reporting, if $w_2 = 0$ and $w_1 > 0$ then ρ_+ (as defined in proof of Proposition 1) is $\rho_+ = 1 - \frac{c}{w_1\sigma^2}$. So as long as $c < w_1\sigma^2$, the equilibrium is $\rho = \sqrt{\rho_+}$; otherwise, the equilibrium is 0. It now follows from (A-3) that $\rho = \alpha$. When $w_1 = 0$, since the failure condition $c \geq \frac{\sigma^2 w_1^2 k^2}{w_2(k-\xi)^2 + w_1 k^2} = 0$ is always valid; hence $\rho_s^\star = \alpha^\star = 0$ is one equilibrium. $\square$

Proof of Proposition 2. When the student grades choose ρ and α, each student reports $G_i \sim \mathcal{N}(\mu + \alpha\rho(q-\mu), \alpha^2(1-\rho^2)\sigma^2)$. The aggregate grade is then

$$\hat{G} \sim \mathcal{N}\left(\mu + \xi\alpha\rho(q-\mu), \frac{\xi^2\alpha^2(1-\rho^2)}{k}\sigma^2\right).$$

Therefore,

$$\mathbb{E}[(\hat{G} - Q)^2] = \sigma^2\left(\frac{\xi^2\alpha^2(1-\rho^2)}{k} + (\xi\alpha\rho - 1)^2\right).$$

Substituting for α and ρ and then setting $\xi = 1$ gives $\mathbb{E}[(\hat{G}_h - Q)^2]$ and $\mathbb{E}[(\hat{G} - f - Q)^2]$. If $\xi = 1$, then

$$\frac{\partial\mathbb{E}[(\hat{G} - Q)^2]}{\partial\rho} = \frac{2\alpha\sigma^2(\alpha\rho(k-1) - k)}{k} < 0, \quad \frac{\partial\mathbb{E}[(\hat{G} - Q)^2]}{\partial\alpha} = \frac{2\sigma^2(\alpha(\rho^2(k-1)+1) - k\rho)}{k};$$

both of these expressions are nonpositive because $\alpha^\star \leq \rho_s^\star$ (by Proposition 1). We also know that $\rho_h^\star \geq \rho_s^\star$, from which it follows that $\mathbb{E}[(\hat{G}_s - Q)^2] \geq \mathbb{E}[(\hat{G}_h - Q)^2]$.

$$\mathbb{E}[(\hat{G}|q) - q] = (\alpha\rho - 1)(q - \mu).$$

Thus, $q > \mu$ implies $\mathbb{E}[\bar{G}] < q$, and $q < \mu$ implies $\mathbb{E}[\bar{G}] > q$. Finally, $\rho_h^\star \geq \rho_s^\star$ implies that $\mathbb{E}\big|[\hat{G}_s|q] - q\big| \geq \mathbb{E}\big|[\hat{G}_h|q] - q\big|$. $\square$

Proof of Proposition 2. Given ρ and α, the expected grade is $E[\hat{G}] = \mu + \kappa\xi\rho\alpha(q-\mu)$. The utility of a student who exerts learning effort e is

$$\mu + \kappa\xi\rho\alpha(\log(e) - \mu) - e,$$

and the optimal learning effort level is $e^\star = \kappa\xi\rho\alpha$. Substituting ρ, α, and $\xi = 1$ now yields parts (i) and (ii) of the proposition. $\square$

Proof of Corollary 3. The monotonicity on these parameters is established in the proof of Proposition 1. $\square$

Proof of Lemma 2. (i) Given ρ, the MSE is

$$\mathbb{E}[(\hat{G}-Q)^2]=\sigma^2\left(\frac{\xi^2\alpha^2(1-\rho^2)}{k}+(\xi\alpha\rho-1)^2\right).$$

We can use the FOC to derive the optimal $\xi=\frac{k\rho}{\alpha(1+(k-1)\rho^2)}$. Alternatively, the conditional expectation $\mathbb{E}[Q|G_1,\ldots,G_k]$ minimizes the MSE. From Bayes' rule we deduce that

$$f_{Q|G_1(\rho_1)=g_1,G_2(\rho_2)=g_2,\ldots,G_k(\rho_k)=g_k}(q)=\frac{f_Q(q)f_{\mathcal{A}_1|Q}(a_1),\ldots,f_{\mathcal{A}_k|Q_k}(a_k)}{f(\mathcal{A}_1(\rho_1)=a_1,\mathcal{A}_2(\rho_2)=a_2,\ldots,\mathcal{A}_k(\rho_k)=a_k)},$$

where $a_j=(g_j-\mu)/\alpha+\mu$.

Since $f_Q(q)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(q-\mu)^2}{2\sigma^2}\right)$, $f_{\mathcal{A}_i|Q}(a_i)=\frac{1}{\sqrt{2\pi(1-\rho_i^2)}\sigma}\exp\left(-\frac{(a_i-\mu-\rho_i(q-\mu))^2}{2(1-\rho_i^2)\sigma^2}\right)$, it follows that

$$\begin{aligned}f_{Q|\mathcal{A}_1(\rho)=a_1,\mathcal{A}_2(\rho)=a_2,\ldots,\mathcal{A}_k(\rho)=a_k}(q)&\sim\exp\left(-\frac{(q-\mu)^2}{2\sigma^2}-\sum_{i=1}^{k}\frac{(a_i-\mu-\rho_i(q-\mu))^2}{2(1-\rho_i^2)\sigma^2}\right)\\&\sim\exp\left(\frac{\mu_X}{2\sigma_X^2}\right)\end{aligned}$$

We can find

$$\mu_X=\frac{\sum_{i=1}^k\frac{\rho_i}{1-\rho_i^2}a_i+\left(1-\sum_{i=1}^k\frac{\rho_i}{1+\rho_i}\right)\mu}{1+\sum_{i=1}^k\frac{\rho_i^2}{1-\rho_i^2}}=\mu+\frac{\sum_{i=1}^k\frac{\rho_i}{1-\rho_i^2}\frac{g_i-\mu}{\alpha_i}}{1+\sum_{i=1}^k\frac{\rho_i^2}{1-\rho_i^2}}.$$

When all the graders choose the same ρ and α, we have

$$\begin{aligned}\mathbb{E}[Q\mid G_1(\rho)=g_1,G_2(\rho)=g_2,\ldots,G_k(\rho)=g_k]&=\frac{\rho\sum_{i=1}^k a_i+(1-(k-1)\rho)(1-\rho)\mu}{1+(k-1)\rho^2}=\mu+\frac{\rho\sum_{i=1}^k\frac{g_i-\mu}{\alpha}}{1+(k-1)\rho^2}\\&=\mu+\frac{k\rho}{\alpha(1+(k-1)\rho^2)}(\bar{G}-\mu).\end{aligned}$$

(ii) Given assignment quality q, if $\rho^\star\neq 0$ (i.e., if peer grading does not fail) then eliminating bias requires that $\mathbb{E}[\hat{G}|q]=q$. That is, we must have

$$\mu+\xi((1-\alpha\rho)\mu+\alpha\rho q-\mu)=q\Rightarrow\xi=\frac{1}{\alpha\rho}.\quad\square$$

Proof of Proposition 3. This proof for the first and second parts are integrated into the proof of Proposition 1. The the monotonicity statement, we have

$$\frac{\partial F(\rho)}{\partial\xi}=-\rho(1-\rho^2)(k-1)(k-2\xi)w_2\sigma^2,$$

which is negative when $\xi<k/2$ or positive when $\xi>k/2$. Therefore, $\rho^\star$ is increasing in ξ when $\xi<k/2$ but decreasing in ξ when $\xi>k/2$. This means that $\xi=k/2$ does maximize $\rho_h^\star$.

Next we show that $\rho_s^\star(\xi)$ is increasing in ξ at $\xi \le 1$ and deceasing in ξ at $\xi \ge k$. If $\alpha_\star = 1$, then $\rho_s^\star = \rho_h^\star$ and the statement is true. If $\alpha_\star < 1$, it follows from the proof of Proposition 1 that $\rho_s^\star$ is the unique solution on $[0,1]$ to $A_\xi(\rho) = c$, and that $A_\xi(\rho)$ is decreases with ρ when $0 < c < \frac{k^2\sigma^2 w_1^2}{(k-\xi)^2 w_2 + k^2 w_1}$. We next show that $A_\xi(\rho)$ is decreasing in ξ at $\xi \le 1$ and increasing in ξ at $\xi \ge k$. Since $\alpha^\star$ is increasing in $\rho_s^\star$, we know that $\alpha^\star$ is also decreasing in ξ at $\xi \le 1$ and increasing in ξ at $\xi \ge k$.

$$\frac{\partial A_\xi}{\partial \xi} = \frac{2k^2(1-\rho^2)\sigma^2 w_1^2 w_2 A_3}{A_2^3},$$

where $A_3 = k^2(k + (k-1)k\rho^2 - \xi - 2(k-1)\rho^2\xi)w_1 + (1 + (k-1)\rho^2)(k-\xi)^3 w_2$, and recall $A_2 = k^2 w_1 - (k-\xi)(\xi(1-\rho^2) + k(\rho^2\xi - 1))w_2$ is defined previously. When $\xi > k$ or $\xi < 1$, since $(k-\xi)(\xi(1-\rho^2) + k(\rho^2\xi - 1)) < 0$, we have $A_2 > 0$. When $\xi > k$, both summands of A_3 are negative, hence $A_3 < 0$, implying $\frac{\partial A_\xi}{\partial \xi} < 0$; when $0 < \xi < 1$, both summands of A_3 are positive, hence $A_3 > 0$, implying $\frac{\partial A_\xi}{\partial \xi} > 0$. $\square$

Proof of Proposition 4. The proof is integrated into the proof of Proposition 2. $\square$

Proof of Corollary 4. Parts (i) and (ii) follow directly Proposition 2. For part (iii), observe that if $\xi_h^{\text{deb}} > \xi_h^{e\star}$, then $\xi_h^{e\star}\rho_h^\star(\xi_h^{e\star}) \ge 1$. $\square$

Proof of Lemma 3. From the proof of Proposition 1 it follows that, given ξ, both $\rho_i^\star$ $(i = h, s)$ and $\alpha^\star$ are decreasing in c. Therefore, $\xi\rho_h^\star$ and $\xi\rho_s^\star\alpha^\star$ are decreasing in c. We need to show that, when c is small enough, $\max_\xi \xi\rho_h^\star > 1$ and $\max_\xi \xi\rho_s^\star\alpha^\star > 1$. It is sufficient to establish that, when c is sufficiently small, $\xi\rho_h^\star > 1$ and $\xi\rho_s^\star\alpha^\star > 1$ at $\xi = k$.

In the case of honest reporting, if $\xi = k$ then the FOC is

$$F(\rho) = ck^2\rho - (1-\rho^2)\sigma^2 k^2 w_1 = 0. \quad \text{We solve } \rho_h^\star = \frac{\sqrt{c^2 + 4w_1^2\sigma^4} - c}{2w_1\sigma^2}.$$

$\rho_h^\star$ is decreasing in c and $\lim_{c\downarrow 0} \rho_h^\star = 1$. Therefore there exists an $\tilde{c}$ such that $\rho_h^\star = 1/k$. And when $c < \tilde{c}$, $\rho_h^\star k > 1$. For strategic graders, if $\xi = k$ then the FOC is

$$c = A_\xi(\rho) = w_1\sigma^2(1-\rho^2) \implies \rho_s^\star = \sqrt{1 - \frac{c}{w_1\sigma^2}} \implies \lim_{c\downarrow 0} \rho_s^\star = 1.$$

Since

$$\alpha^\star = \min\left\{1, \frac{\rho_s^\star w_1 k^2}{w_1 k^2 + w_2(k-\xi)(k(1-\rho_s^{\star 2}\xi) - (1-\rho_s^{\star 2})\xi)}\right\},$$

we have $\lim_{c\downarrow 0} \alpha^\star = 1$. Therefore, $\lim_{c\downarrow 0} \xi\rho_s^\star(\xi)\alpha^\star|_{\xi=k} = k > 1$. From Corollary 3, both $\rho_s^\star$ and $\alpha^\star$ are decreasing in c, there exists an $\tilde{c}$ such that $\rho_s^\star\alpha^\star = 1/k$. When $c < \tilde{c}$, $k\rho_s^\star\alpha^\star > 1$.

Finally, for a given ξ we can write

$$\rho_h^\star \ge \rho_s^\star \implies \xi_s^{\text{deb}}\rho_h^\star \ge \xi_s^{\text{deb}}\rho_s^\star\alpha^\star = 1$$

Therefore, $\xi_h^{\text{deb}} \le \xi_s^{\text{deb}}$. $\square$