

# Report-Sensitive Spot-Checking in Peer-Grading Systems\*

Hedayat Zarkoob  
University of British Columbia  
Vancouver, BC  
hzarkoob@cs.ubc.ca

Hu Fu  
University of British Columbia  
Vancouver, BC  
hufu@cs.ubc.ca

Kevin Leyton-Brown  
University of British Columbia  
Vancouver, BC  
kevinlb@cs.ubc.ca

## ABSTRACT

Peer grading systems make large courses more scalable, provide students with faster and more detailed feedback, and help students to learn by thinking critically about the work of others. A key obstacle to the broader adoption of peer grading systems is motivating students to provide accurate grades. The literature has explored many different approaches to incentivizing accurate grading (which we survey in detail), but the strongest incentive guarantees have been offered by mechanisms that compare peer grades to trusted TA grades with a fixed probability. In this work, we show that less TA work is required when these probabilities are allowed to depend on the grades that students report. We prove this result in a model with two possible grades, arbitrary numbers of agents, no requirement that students grade multiple assignments, arbitrary but homogeneous noisy observation of the ground truth which students can pay a fixed cost to observe, and the possibility of misreporting grades before or after observing this signal. We give necessary and sufficient conditions for our new mechanism’s feasibility, prove its optimality under these assumptions, and characterize its improvement over the previous state of the art both analytically and empirically. Finally, we relax our homogeneity assumption, allowing each student and TA to observe the ground truth according to a different noise model.

## KEYWORDS

Peer Grading; Peer Prediction; Mechanism Design

### ACM Reference Format:

Hedayat Zarkoob, Hu Fu, and Kevin Leyton-Brown. 2020. Report-Sensitive Spot-Checking in Peer-Grading Systems. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020)*, Auckland, New Zealand, May 9–13, 2020, IFAAMAS, 9 pages.

## 1 INTRODUCTION

Peer grading has the potential to improve educational outcomes in three main ways: (i) making classes more scalable by offloading some grading work to students, (ii) providing students with faster and more detailed feedback, and (iii) improving student learning by providing opportunities to think critically about the work of others. Various recent implementations of peer grading mechanisms make such systems relatively easy to deploy in practice [2, 13, 24]. The broader adoption of such systems faces a common, critical obstacle: motivating students to provide accurate grades. A natural solution is asking multiple students to grade the same assignment

and rewarding them based on their behavior (e.g., based on the extent to which their grades agree with the grades given by other students). Such solutions have been explored in detail in a large literature on *peer prediction*, which considers how to incentivize agents to truthfully disclose unverifiable private information [4, 8–11, 15–19, 22, 23]. Miller et al. [14] were the first to introduce peer prediction mechanisms in which truthful declarations constitute a Nash equilibrium. Unfortunately, these mechanisms (and, indeed, many others that were subsequently proposed) also give rise to uninformative equilibria in which agents do not reveal their private information; e.g., all students grading an assignment favorably regardless of its quality [1, 9, 11, 19, 22]. Human experiments show that such strategic behavior does arise in practice [5].

Much subsequent work has attempted to identify peer prediction mechanisms in which either no uninformative equilibrium exists or the truthful equilibrium is always preferred by agents [9, 11, 19, 22]. Two examples are particularly notable. First, Dasgupta and Ghosh [1] considered a model in which agents make a binary decision about whether or not to invest costly effort, in the former case observing a noisy signal about the assignment’s true value. Agents are paid according to a function that rewards agreement between graders on the same assignment and penalizes correlations in the grades assigned across different assignments. Under this mechanism, truthful reporting yields payoffs that exceed those of any other equilibrium for every agent. Furthermore, if the system contains a small fraction of agents (e.g., TAs) who are always truthful, the truthful equilibrium becomes unique. Second, De Alfaro et al. [3] showed how to achieve unique, truthful equilibria by combining peer prediction with trusted reports in a hierarchical mechanism. One drawback of all such approaches is that they only achieve Nash equilibrium implementations, because agents’ payoffs depend on other agents’ actions, and so agents must reason about each other’s behavior. In a classroom setting, where some students will almost surely fail to invest effort, stronger incentives may be required.

Liu and Chen [12] showed that stronger incentives can be guaranteed by a peer prediction mechanism based on surrogate scoring rules, achieving “uniform dominant strategy incentive compatibility” in a setting where students make noisy observations about two possible grades. This result relies on there existing sufficiently many tasks per student; more critically, it also requires that students follow the same strategy for all tasks (and be entirely sure that all other students also do so), rather than making the more standard assumption that students make a separate strategic decision every time they are faced with a grading task.

Another approach for strengthening incentive guarantees relies on incorporating trusted graders (TAs) more fundamentally into the mechanism. For example, Goel and Faltings [7] proposed a second uniform dominant strategy mechanism that uses a set of “golden tasks” (for which the ground truth is known via an oracle)

\*This article supersedes an AAMAS 2019 extended abstract with the same title.

to incentivize high quality evaluation for a small set of agents. Then they use the reports provided on non-golden tasks by the agents in this set to incentivize high quality grading among other agents. Like the mechanism of Liu and Chen [12], this mechanism relies on (1) the existence of multiple tasks and (2) agents following the same strategy for all tasks; it further requires (3) the determination of ground truth for golden tasks before agents begin grading.

In this work, we follow a different thread of past work [6–8, 21, 24], obtaining (fully) dominant strategy mechanisms by guaranteeing that, with sufficiently high probability, each student report is compared to a trusted but noisy observation of ground truth (a TA grade). Because such “spot checking” is expensive (e.g., TAs need to be paid in proportion to the amount of work they do), it is natural to seek to minimize the amount of spot checking required to obtain dominant strategies. This minimization problem was first attacked by Gao et al. [6], who proposed a very simple mechanism that makes truthfulness a dominant strategy by unconditionally rewarding students when they are not spot checked and otherwise penalizing them to the extent that they disagree with the TA. They compared this mechanism with various alternatives based on peer prediction, showing that the latter require strictly more spot checking than the former, even despite the fact that peer-prediction-based mechanisms do not offer dominant strategies.

Gao et al’s model always performs spot checks with some fixed probability. It is intuitive to think that *report-sensitive spot checking*—that is, varying the spot-checking probability based on the students’ reports—could lower the expected amount of spot checking required overall. For example, imagine that an instructor already knows that a given problem set is extremely difficult. If the reported grades for a given submission are all very high, the instructor might believe that there is an increased likelihood that students have reported dishonestly, and so might want to spot check with a higher probability. It turns out to be nontrivial to confirm or refute this intuition, for two key reasons. First, more complex ways of computing spot check probabilities opens the door to new ways for students to manipulate a mechanism. Second, students’ interests become intertwined in a new way, since spot-checking probabilities now depend on other agents’ strategies.

Despite these obstacles, this paper (1) identifies the optimal dominant-strategy incentive-compatible (DSIC) report-sensitive spot-checking mechanism; (2) identifies necessary and sufficient conditions for such a mechanism’s existence; (3) shows that the new mechanism requires less TA work on expectation than the mechanism of Gao et al; and that (4) the new mechanism exists for a broader range of parameter values. One interesting property of our new mechanism is that it sometimes chooses not to spot check students even when TA assessments are available for the assignments they graded: we show that making full use of TA assessments has the perverse effect of increasing expected workload overall.

Like much other work in the literature [e.g., 1, 12, 21], our analysis is limited to the case where students are asked to report only binary (positive or negative) grades about each assignment. Our new mechanism is general in several other, important senses: it allows for arbitrary numbers of graders per assignment and nearly arbitrary prior probability distributions over both the true grades and the noise models describing the probabilities that students and TAs will observe each signal given the ground truth. In the first

part of the paper we assume that these noise models are homogeneous (identical across all students and TAs); in the last section, we generalize our results to the fully heterogeneous case (each agent can grade according to a different noise model).

Finally, we mention two final strands of recent, related work. Wang et al. [21] proposed a different approach for designing peer grading systems that also varies spot check probabilities. Their model is substantially different from ours, and hence their mechanism is not directly applicable to our setting. Like us, they study strategic students who make a binary decision about whether to invest effort; as we do in the last section of this paper, they assume that agents’ noise models are heterogeneous. However, they also assume that TAs can directly observe whether a student invested effort, making it simple to ensure that a spot-checked student who invested no effort gets no reward. In contrast, we assume that the TA noisily observes the assignment’s grade and is only able to compare this observation with the student’s own report, which is either a noisy signal or a misreport; thus, students who invest no effort cannot reliably be identified. Xiao et al. [25] studied incentive mechanisms for peer reviewing in a repeated setting. In their model, students benefit from receiving high quality reviews and decide on the amount of effort to put in when reviewing others. Their proposed mechanism incentivizes high-quality reviewing by rating students’ reviewing performance over time and using this rating to match highly ranked reviewers with each other. Key caveats are that the mechanism relies on the ability to accurately score reviewing quality and that high-quality reviewing is incentivized only in equilibrium.

In the following, we first define our model and formalize the mechanisms that we study throughout the paper (Section 2). We next prove that our proposed mechanism is optimal (Section 3) and show that it outperforms alternatives (Section 4) by demonstrating a separation analytically and then quantifying the gap via numerical experiments. Finally, we extend our results to heterogeneous student and TA noise models (Section 5) and conclude (Section 6).

## 2 MODEL

A single assignment<sup>1</sup> needs to be graded by a set  $N$  of students (with  $|N| = n$ ) and has an unobservable binary quality  $q \in Q = \{a, b\}$  drawn from a commonly known distribution  $\Pr[q]$ .

Each student  $i$ , by exerting effort at cost  $c$ , can examine the submission and observe a signal  $s_i \in Q$  that is informative about the assignment’s quality. More formally, in a way that depends on the true quality  $q$ , the signals observed by different students are independently drawn from a single,<sup>2</sup> commonly known distribution  $\Pr[s|q]$ . The *ex ante signal distribution* is then

$$\Pr[s = l] = \sum_{t \in Q} \Pr[s_i = l | q = t] \Pr[q = t].$$

We denote by  $P_{\vec{l}}$  the *ex ante* probability of each agent receiving the signal corresponding to its index in vector  $\vec{l}$ . By our assumption

<sup>1</sup>When multiple assignments must be peer graded, our mechanisms can simply be run in parallel.

<sup>2</sup>In Section 5, we relax this assumption that signal distributions are homogeneous.

of conditional independence, this is

$$\begin{aligned} P_{\vec{l}} &:= \Pr [s_1 = l_1, \dots, s_n = l_n] \\ &= \sum_{t \in Q} \Pr [q = t] \prod_{j \in N} \Pr [s_j = l_j | q = t]. \end{aligned} \quad (1)$$

Because of conditional independence, any two vectors of reports  $\vec{l}$  and  $\vec{l}'$  containing the same numbers of  $a$ 's and  $b$ 's occur with the same probability:  $P_{\vec{l}} = P_{\vec{l}'}$ . For this reason, we often drop the ordering in the subscript, writing, e.g.,  $P_{(a,b)} = P_{(b,a)} = P_{ab}$ , and similarly for longer vectors. We name the signals so that  $P_a \geq P_b$  (note that this implies  $P_a - P_{ab} = P_{aa} > P_{bb} = P_b - P_{ab}$ ). We also denote by  $P_{l|t}$ , the probability that an agent observes signal  $l \in Q$  conditioned on another agent observing signal  $t \in Q$ .

Besides the students, a teaching assistant (TA) may also receive a signal. Formally, signal  $s_{TA}$  is drawn from  $\Pr[s|q]$  independently from the students' signals.<sup>3</sup>

*Strategy space.* In our model, each student faces two strategic choices: whether to expend effort grading the assignment and what grade to report. Three actions are thus possible: the student (i) may be *truthful*, investing effort to examine the assignment, observing her signal, and reporting this signal; (ii) may invest effort but report a different signal than the one she observed; or (iii) may choose not to invest effort and report an arbitrary signal. In contrast, the TA is not a strategic agent. When asked to grade the assignment, the TA always reports an independently observed signal.

## 2.1 Spot-checking Mechanisms

A focus of our work is on minimizing the need for the TA's input via *spot-checking mechanisms*. A spot-checking mechanism takes in students' reported signals and decides both whether a TA signal is needed and how much to reward the students.

*Definition 2.1 (Spot-checking mechanism).* A *spot-checking mechanism* is defined by a tuple  $(x_a, x_b, Y)$ , where:

- (1)  $x_a : \{0\} \cup \mathbb{N} \times \mathbb{N} \rightarrow [0, 1]$  denotes the probability of spot checking an agent who reports  $a$ . Given two non negative integers  $(k, n)$  specifying the number of  $a$ 's reported by the agents and the total number of agents,  $x_a(k, n)$  returns the probability that the mechanism will spot check each agents reporting  $a$ . The probability  $x_a(0, n)$  is set to zero.
- (2)  $x_b : \{0\} \cup \mathbb{N} \times \mathbb{N} \rightarrow [0, 1]$  is an analogous function for computing the probability of spot checking an agent who reports  $b$ . The first argument remains the total number of agents who report  $a$ , not  $b$ . The probability  $x_b(n, n)$  is zero.
- (3)  $Y : Q \times Q \rightarrow \mathbb{R}^+$  denotes the reward given to a student who is spot checked.  $Y(r, s_{TA})$  is the reward of a spot-checked student who reported  $r$  when the TA reported signal  $s_{TA}$ . When a student is not spot checked, she receives no reward.

The  $x$  functions are defined for every  $n$  because we require peer grading mechanisms to work for any number of agents. However, when  $n$  is obvious from context, we will overload notation and write simply  $x_a(k)$  and  $x_b(k)$ .

Throughout this paper we focus on mechanisms where the reward function  $Y$  is the simplest identity function:

$$Y(r, s_{TA}) = \begin{cases} R & \text{if } r = s_{TA} \\ 0 & \text{otherwise,} \end{cases}$$

where  $R \in \mathbb{R}^+$  denotes the reward of matching with the TA signal. This function is called the output agreement reward function; it has been widely studied in the peer prediction literature [20, 22].

We model students as having quasilinear utility: i.e., when investing effort and being rewarded  $Y$ , a student's utility is  $Y - c$ .

*Definition 2.2 (DSIC).* A spot-checking mechanism is *dominant strategy incentive compatible* (DSIC) if, for each student  $i$  and for any strategies that the other students choose,  $i$ 's expected utility-maximizing strategy is to be truthful, i.e., to invest effort to observe her signal and to report what she observes.

The mechanism we will later show to be optimal is DSIC; however, we also define a weaker solution concept ("Incentive Compatible with Conscientious Plays", or ICCP), to allow us to compare our preferred mechanism to a broader set. A mechanism is ICCP if it is DSIC in a simplified strategy space in which students who observe their signals must report them honestly. In other words, ICCP does not allow for the possibility that students could misreport a signal that they invested effort to observe, but continues to allow for lazy students pretending to have observed arbitrary signals.

*Definition 2.3 (ICCP).* A spot-checking mechanism is *Incentive Compatible with Conscientious Plays* (ICCP) if, for each student and for any strategy that other students choose, *as long as each student that examines the assignment always reports the observed signal*, the truthful strategy, i.e., to invest effort to examine the assignment and report her observed signal, is expected utility maximizing.

*Definition 2.4 (TA workload).* For a DSIC or ICCP spot-checking mechanism, the *TA workload* (or simply the *workload*) is the probability with which the TA needs to provide a signal, assuming all students are truthful:

$$\begin{aligned} \sum_{t \in Q} \Pr [q = t] \sum_{j=0}^n \binom{n}{j} & (\Pr [s_i = a | q = t])^j \\ & \cdot (\Pr [s_i = b | q = t])^{n-j} \max\{x_a(j), x_b(j)\}. \end{aligned} \quad (2)$$

The expression  $\max\{x_a(j), x_b(j)\}$  is TA workload needed to ensure that every student reporting  $a$  is spot checked with probability  $x_a(j)$  and every student reporting  $b$  is spot checked with probability  $x_b(j)$ . We say that a mechanism is *optimal* for some class if it minimizes TA workload among all mechanisms in that class.

## 2.2 ROS, RSS, and RSUS Mechanisms

We now introduce three families of mechanisms, beginning with one studied by Gao et al. [6].

*Definition 2.5 (ROS Mechanism).* A *Report-Oblivious Spot-checking* (ROS) mechanism spot checks every student with fixed probability  $x$ , regardless of the students' reports.

The focus of this paper is on *report-sensitive* spot-checking.

*Definition 2.6 (RSS Mechanism).* A *Report-Sensitive Spot-checking* (RSS) mechanism spot checks every student with probability that can depend on all the students' reports.

<sup>3</sup>In Section 5, we relax this assumption that signal distributions are homogeneous.

In a RSS mechanism, we determine whether to spot check each student independently: i.e., we sometimes refrain from spot checking one student even when the TA signal is already available from spot checking another student. This seems wasteful: once the TA spot checks one student, it costs them no additional work to use the observed signal to spot check other students too. We now define a class of mechanisms that leverages this fact.

*Definition 2.7 (RSUS Mechanism).* A *Report-Sensitive, Uniform Spot-checking (RSUS)* mechanism ensures that whenever one student is spot checked, all are spot checked.

In RSUS mechanisms,  $\forall j \in \{0, \dots, n\}$ ,  $x_a(j) = x_b(j)$ ; of course, it still allows for  $j \neq j'$  that  $x_a(j) \neq x_a(j')$ . A main result of this paper is that RSUS mechanisms can require strictly larger spot-checking budgets than a DSIC RSS mechanism, and are hence strictly sub-optimal, not just under the DSIC solution concept but even if they only need to satisfy the weaker ICCP solution concept. In other words, paradoxically, in order to minimize the overall TA workload, it is necessary to commit sometimes *not* to use the TA's signal to spot check some students.

### 3 OPTIMAL DSIC MECHANISMS

We now characterize optimal DSIC mechanisms in both the ROS and RSS families.

#### 3.1 Optimal DSIC ROS Mechanisms

We begin by stating a result about ROS mechanisms from the literature. Given a prior distribution  $\Pr[q]$  over the assignment's quality, the conditional signal distributions  $\Pr[s|q]$ , and the cost of the effort needed to examine the assignment and the reward  $R$ , there might be a minimum spot-checking probability  $x$  (which is independent of  $n$ , the number of students) that guarantees an ROS mechanism to be DSIC.

**THEOREM 3.1.** (Consequence of Lemma 1 in [6]) For any signal  $s \in Q$ , and any  $c, R > 0$ , if  $P_{bb} - P_{ab} \geq \frac{c}{R}$ , a DSIC ROS mechanism spot checks a student with probability at least

$$x^* = \frac{\frac{c}{R}}{P_{aa} + P_{bb} - P_a} = \frac{\frac{c}{R}}{P_{bb} - P_{ab}}; \quad (3)$$

if  $P_{bb} - P_{ab} < \frac{c}{R}$ , there does not exist any DSIC ROS mechanism.

Recall that  $P_{ab}$  is the *ex ante* probability that, when two signals  $s_1$  and  $s_2$  are drawn independently from  $\Pr[s|q]$ ,  $s_1 = a$  and  $s_2 = b$ . The assumption  $P_{bb} - P_{ab} \geq \frac{c}{R}$  is necessary and sufficient to ensure that the ROS mechanism is DSIC if spot checks are performed with probability 1: a student who is always spot checked prefers to be truthful than to report either signal without effort. We refer to  $x^*$  as the optimal DSIC ROS solution and the ROS mechanism that spot checks with probability  $x^*$  as the optimal ROS mechanism.

Equation (3) demonstrates that TA workload of the optimal ROS mechanism falls as the reward–cost ratio  $R/c$  increases. However, both costs (effort required) and rewards (e.g., portion of a course grade devoted to performing peer review) are often fixed in practical peer grading scenarios. We thus take the approach of characterizing which (fixed) values of  $R/c$  give rise to feasible spot-checking mechanisms and of investigating how spot-checking policies can be tuned to minimize TA workload.

#### 3.2 Optimal DSIC RSS Mechanisms

We begin by defining a class of mechanisms and then go on to prove our first main result: that a mechanism from this class minimizes TA workload across all DSIC RSS mechanisms.

*Definition 3.2 (PRSS Mechanism).* A *Personal-Report-Sensitive Spot-checking mechanism*, or *PRSS mechanism*, spot checks each student with a probability that only depends on the student's own report.

In PRSS mechanisms, the functions  $x_a$  and  $x_b$  are constant, i.e., for  $\forall j, j' \in \{0, \dots, n\}$  and  $j \neq j'$ ,  $x_a(j) = x_a(j')$  and  $x_b(j) = x_b(j')$ . A remarkable feature of PRSS mechanisms is that, if  $x_a(k) \neq x_b(k)$  for some  $k$ , the TA's input is sometimes “wasted”, in the sense that it is not used to spot check every student. To see this, say if  $x_b(k) > x_a(k)$ , then the mechanism should consult the TA with probability  $x_b(k)$  but spot checks a student reporting  $a$  with probability only  $x_a(k)$ .

Our first result is a characterization of the optimal DSIC RSS mechanism. We show that this optimum is achieved by a PRSS mechanism with appropriate spot-checking probabilities. We also characterize the necessary and sufficient condition for the existence of DSIC RSS mechanisms. Later we show that this condition is weaker than the one for DSIC ROS mechanisms in Theorem 3.4.

**THEOREM 3.3.** For any signal  $s_i \in Q$ , and any  $c, R > 0$ , if  $P_{a|a} - P_a \geq \frac{c}{R}$ , a PRSS mechanism with the following spot-checking probabilities is the DSIC spot-checking mechanism that minimizes the TA workload among all other DSIC spot-checking mechanisms:

$$x_a(k) = \frac{\frac{c}{R}}{P_{b|b} - P_b} \quad \forall k \in \{1, \dots, n\} \quad (4)$$

$$x_b(k) = \frac{\frac{c}{R}}{P_{a|a} - P_a} \quad \forall k \in \{0, \dots, n-1\}; \quad (5)$$

if  $P_{a|a} - P_a < \frac{c}{R}$ , there does not exist any DSIC spot-checking mechanism.

This result is nontrivial. Recall that an RSS mechanism has the power to vary the spot-checking probabilities for each student based on the reports of all other students, and such side information conceivably could help reduce the TA workload while maintaining the students' incentives. This turns out not to be the case. One immediate consequence of Theorem 3.3 is that the TA workload *increases* with the number of students grading each assignment.

The proof consists of two steps. In Step 1, we reason about a convex optimization problem that minimizes the workload but relaxes all of the DSIC constraints except those that incentivize the truthful strategy when all other students make no effort. We show that the intersection point of the non-trivial constraints in this optimization problem is locally optimal, and hence, also, the global optimal of the defined optimization problem. Then, in Step 2, we show that this solution in fact gives rise to a DSIC RSS mechanism. Since it is an optimal solution with most DSIC constraints relaxed, it is also optimal when one enforces all constraints.

**PROOF. Step 1.** In a DSIC mechanism, fixing a student  $i$ , if all students other than  $i$  report a signal without expending effort,  $i$  should be incentivized to be truthful, i.e., to invest effort and to report the signal observed. Suppose  $k$  other students report  $a$ : then,

the utility of student  $i$  reporting signal  $a$  without investing effort is  $R \cdot \Pr[s_{TA} = a] \cdot x_a(k+1)$ , whereas her utility for being truthful is

$$R \cdot (\Pr[s_i = a] \cdot \Pr[s_{TA} = a \mid s_i = a] x_a(k+1) + \Pr[s_i = b] \cdot \Pr[s_{TA} = b \mid s_i = b] x_b(k)) - c.$$

We therefore should have that, for  $k \in \{0, \dots, n-1\}$ ,

$$\begin{aligned} & \Pr[s_i = a] \cdot \Pr[s_{TA} = a \mid s_i = a] x_a(k+1) \\ & + \Pr[s_i = b] \cdot \Pr[s_{TA} = b \mid s_i = b] x_b(k) \\ & - \Pr[s_{TA} = a] x_a(k+1) \geq \frac{c}{R}. \end{aligned} \quad (6)$$

Similarly we should have for  $k \in \{0, \dots, n-1\}$  that

$$\begin{aligned} & \Pr[s_i = a] \cdot \Pr[s_{TA} = a \mid s_i = a] x_a(k+1) \\ & + \Pr[s_i = b] \cdot \Pr[s_{TA} = b \mid s_i = b] x_b(k) \\ & - \Pr[s_{TA} = b] x_b(k) \geq \frac{c}{R}. \end{aligned} \quad (7)$$

Simplifying (6) and (7), we have

$$-P_{ab}x_a(k+1) + P_{bb}x_b(k) \geq \frac{c}{R}. \quad (8)$$

$$P_{aa}x_a(k+1) - P_{ab}x_b(k) \geq \frac{c}{R}. \quad (9)$$

Consider minimizing the TA workload subject to (8) and (9):

$$\begin{aligned} \min_{x_a, x_b} \sum_{t \in Q} \Pr[q = t] \sum_{j=0}^n \binom{n}{j} (\Pr[s_i = a \mid q = t])^j \\ \cdot (\Pr[s_i = b \mid q = t])^{n-j} \max\{x_a(j), x_b(j)\}. \end{aligned}$$

with  $0 \leq x_a(k), x_b(k) \leq 1$  for all  $k \in \{0, \dots, n\}$ . The value of this optimization problem is an upper bound to the workload of any DSIC mechanism, since it relaxed all DSIC constraints except (6) and (7). Note also that the objective function is convex in the variables, due to the presence of the max functions, and its feasible region is a convex polytope since all constraints are linear. There are  $2n$  variables:  $x_b(0), x_a(1), x_b(1), x_a(2), \dots, x_a(n-1), x_b(n-1), x_a(n)$ . (Note that  $x_a(0)$  and  $x_b(n)$  are zero.) These variables can be grouped into  $n$  pairs, with  $x_b(k)$  and  $x_a(k+1)$  as a pair for each  $k$ . Each pair is subject to a pair of constraints from (8) and (9) with the corresponding  $k$ , and there is no constraint governing variables from different pairs. This means the optimization problem is separated into  $n$  subproblems, one for each pair of variables. We now claim that the optimal solution for each subproblem is given by forcing (8) and (9) to be tight, which gives us the following expressions:

$$x_a(k) = \frac{c}{R} \cdot \frac{P_b}{P_{aa}P_{bb} - (P_{ab})^2} \quad \forall k \in \{1, \dots, n\} \quad (10)$$

$$x_b(k) = \frac{c}{R} \cdot \frac{P_a}{P_{aa}P_{bb} - (P_{ab})^2} \quad \forall k \in \{0, \dots, n-1\}; \quad (11)$$

Simplifying (11), we get

$$\begin{aligned} x_b(k) &= \frac{c}{R} \cdot \frac{P_a}{P_{aa}P_{bb} - (P_{ab})^2} = \frac{c}{R} \cdot \frac{P_a}{P_{aa}P_{bb} - P_{ab}(P_a - P_{aa})} \\ &= \frac{c}{R} \cdot \frac{P_a}{P_{aa}P_b - P_aP_{ab}} = \frac{c}{R} \cdot \frac{P_a}{P_{aa}(1 - P_a) - P_aP_{ab}} \\ &= \frac{\frac{c}{R}}{\frac{P_{aa}}{P_a} - P_a} = \frac{\frac{c}{R}}{P_{a|a} - P_a}, \end{aligned}$$

which is the expression in (5). We get the expression in (4) by simplifying (10), similarly.

We first check that these solutions are feasible: we have that  $P_a > P_b$  in (10) and (11), therefore, if  $P_{a|a} - P_a \geq \frac{c}{R}$ , the spot-checking probabilities are indeed between 0 and 1. To see that these solutions are optimal, we invoke the convexity of the problem, which means we only need to argue that the solutions are locally optimal. At the point where (8) and (9) are tight for a given  $k$ , increasing  $x_a(k+1)$  forces  $x_b(k)$  to increase, and vice versa, in order for the pair to keep being feasible — this is a consequence of the coefficients' signs in (8) and (9) — but it is not feasible to decrease both variables. Therefore the only local move within the feasible region is to increase both variables. However, both variables have positive coefficients in the objective function. Therefore the solution in (4) and (5) is both locally optimal and globally optimal for our convex program.

We note that if  $P_{a|a} - P_a < \frac{c}{R}$ , to make sure that Constraints (8) and (9) are all satisfied for  $k \in \{0, \dots, n-1\}$ , at least one of  $x_b(k)$  or  $x_a(k+1)$  has to be strictly greater than one. This is also a consequence of the coefficients' signs in (8) and (9). However,  $0 \leq x_a(k), x_b(k) \leq 1$  for all  $k \in \{0, \dots, n-1\}$ . Therefore, there does not exist any feasible DSIC RSS mechanism.

**Step 2.** We now show that the spot-checking probabilities given in (4) and (5) in fact give rise to a DSIC mechanism. We only need to check the validity of the DSIC constraints not included in the convex program above.

We first check the condition that a student having spent the effort to get a signal should be incentivized to report the observation faithfully. We need for  $k \in \{0, \dots, n-1\}$ ,

$$\begin{aligned} \Pr[s_{TA} = a \mid s_i = a] x_a(k+1) \\ \geq \Pr[s_{TA} = b \mid s_i = a] x_b(k), \end{aligned} \quad (12)$$

and

$$\begin{aligned} \Pr[s_{TA} = b \mid s_i = b] x_b(k) \\ \geq \Pr[s_{TA} = a \mid s_i = b] x_a(k+1), \end{aligned} \quad (13)$$

which simplify to

$$P_{aa}x_a(k+1) - P_{ab}x_b(k) \geq 0. \quad (14)$$

$$-P_{ab}x_a(k+1) + P_{bb}x_b(k) \geq 0. \quad (15)$$

Note that (14) and (15) are in fact implied by (8) and (9).

For the other DSIC constraints that concern a student's utility when some other students may spend effort to observe a signal, note that the spot-checking probabilities given by (4) and (5) are independent of  $k$ , i.e., they are independent from what the other students report, and therefore the corresponding mechanism is PRSS. In such a mechanism, a student's utility is independent from the other students' strategies. If the truthful strategy maximizes a student's utility when no other student spends any effort, it still does when other students spend effort. Therefore the mechanism obtained from Step 1 is indeed DSIC, and this completes the proof.  $\square$

The following theorem shows that for a given set of input distributions, while the optimal DSIC RSS mechanism might be feasible, there might not exist any DSIC ROS mechanism.

**THEOREM 3.4.** *For any signal  $s \in Q$ , and any  $c, R > 0$ , whenever there is a DSIC ROS mechanism, there is a DSIC RSS mechanism; however, when  $P_a > P_b$ , there exist  $c, R > 0$  for which  $P_{bb} - P_{ab} < \frac{c}{R} \leq P_{a|a} - P_a$ , i.e., the optimal DSIC RSS mechanism exists while DSIC ROS mechanisms do not.*

**PROOF.** By Theorem 3.3, we know that for any signal  $s_i \in Q$  and any  $c, R > 0$ , if  $\frac{c}{R} \geq P_{a|a} - P_a = \frac{P_{aa}P_{bb} - (P_{ab})^2}{P_a}$ , then the optimal DSIC RSS mechanism exists. Therefore, we only need to show that  $\frac{P_{aa}P_{bb} - (P_{ab})^2}{P_a} > P_{bb} - P_{ab}$ , i.e., there is a gap between these expressions. We can write

$$\begin{aligned} P_{aa} \cdot P_{bb} - P_{ab}^2 &> P_{aa} \cdot P_{bb} - P_{ab}^2 - P_{ab} \cdot (P_{aa} - P_{bb}) \\ &= (P_{aa} + P_{ab})(P_{bb} - P_{ab}) = P_a \cdot (P_{bb} - P_{ab}), \end{aligned}$$

where the first inequality is due to the fact that  $P_{aa} > P_{bb}$ .  $\square$

## 4 COMPARING MECHANISMS

In this section, we compare our optimal DSIC RSS mechanism with alternatives both analytically and experimentally.

### 4.1 Analytic Comparisons

Drawing on our characterization from Section 3, we can now compare the workloads required by the optimal DSIC ROS, RSS and RSUS mechanisms. We begin by comparing the TA workload of the optimal DSIC RSS mechanism with that of the optimal DSIC ROS mechanism.

**COROLLARY 4.1.** *When  $P_a > P_b$  and  $P_{bb} - P_{ab} > \frac{c}{R}$ , the TA workload of the optimal DSIC ROS mechanism exceeds that of the optimal DSIC RSS mechanism by at least  $\frac{c}{R} \cdot \frac{P_b}{P_{aa}P_{bb} - (P_{ab})^2}$ .*

**PROOF.** By Theorem 3.1, the optimal DSIC ROS solution is  $\frac{c}{R} \cdot \frac{1}{P_{bb} - P_{ab}}$ . Let  $\vec{a}$  be a vector of length  $n$  where all signals are  $a$ . By (10), this vector is spot checked with probability  $x_a(n)$ . By (10) and (11),  $x_b(k) > x_a(k)$  for any  $k \in \{0, \dots, n-1\}$  (recall that  $x_b(n) = x_a(0) = 0$ ). Therefore,  $\max\{x_a(j), x_b(j)\} = x_b(j)$  for any vector other than  $\vec{a}$ . Hence, the TA workload saved by the RSS mechanism is:

$$\begin{aligned} &\left( P_{\vec{a}} \left( x^* - \frac{c}{R} \cdot \frac{P_b}{P_{aa}P_{bb} - (P_{ab})^2} \right) \right. \\ &+ (1 - P_{\vec{a}}) \left( x^* - \frac{c}{R} \cdot \frac{P_a}{P_{aa}P_{bb} - (P_{ab})^2} \right) \Big) \\ &\geq \frac{c}{R} \cdot \left( \frac{1}{P_{bb} - P_{ab}} - \frac{P_a}{P_{aa}P_{bb} - (P_{ab})^2} \right) \\ &\geq \frac{c}{R} \cdot \frac{P_b}{P_{aa}P_{bb} - (P_{ab})^2}, \end{aligned} \quad (16)$$

where the first and second inequalities are due to the facts that  $P_a \geq P_b$  and  $P_{aa} \geq P_{ab}$ .  $\square$

**COROLLARY 4.2.** *When  $P_a = P_b$  and  $P_{bb} - P_{ab} > \frac{c}{R}$ , the TA workload of the optimal DSIC RSS mechanism is equal to that of the optimal DSIC ROS mechanism.*

**PROOF.** The optimal DSIC ROS solution is  $\frac{c}{R} \cdot \frac{1}{P_{bb} - P_{ab}}$ . Also, when  $P_a = P_b$ ,  $P_a - P_{ab} = P_{aa} = P_b - P_{ab} = P_{bb}$ , and the spot-checking probabilities defined by (4) and (5) are equal. Thus, the TA workload that the RSS mechanism saves is:

$$\begin{aligned} &\frac{c}{R} \cdot \left( \frac{1}{P_{bb} - P_{ab}} - \frac{P_b}{P_{aa}P_{bb} - (P_{ab})^2} \right) \\ &= \frac{P_{aa}P_{bb} - (P_{ab})^2 - (P_{bb} + P_{ab})(P_{bb} - P_{ab})}{(P_{bb} - P_{ab})(P_{aa}P_{bb} - (P_{ab})^2)} = 0. \end{aligned}$$

$\square$

We now turn to comparing RSUS and RSS mechanisms. Because RSUS mechanisms are special cases of RSS mechanisms, it is obvious that the former perform weakly worse than the latter, all else being equal. Our second nontrivial main result shows something stronger: that the optimal RSS mechanism is strictly better than all RSUS mechanisms, even when the latter are strengthened by being required to satisfy only the ICCP solution concept. The proof is based on a linear program that lower bounds the workload of the optimal RSUS mechanism under the ICCP constraints. We show that this workload is still greater than that of the optimal RSS mechanism. Recall that RSUS mechanisms spot check all students whenever the TA is consulted. Our result thus implies that the TA workload can be decreased by choosing not to spot check certain students even when the TA spot checks other students.

**THEOREM 4.3.** *When  $P_a > P_b$  and  $P_{a|a} - P_a \geq \frac{c}{R}$ , the optimal RSS under the DSIC solution concept spot checks strictly less than the optimal RSUS mechanism under the ICCP solution concept.*

**Remark:** It is straightforward to see that in the case of  $P_a = P_b$ , the TA workload of the optimal ICCP RSUS mechanism is equal to the optimal DSIC RSS and ROS mechanisms.

**PROOF.** The RSUS mechanisms are special cases of the RSS mechanisms, where  $x_a(k) = x_b(k) = x(k)$  for  $k \in \{0, \dots, n\}$ . We prove this theorem by showing that the optimal solution of an adapted version of the RSS convex optimization problem for the RSUS mechanisms is strictly worse than the optimal DSIC RSS solution defined by (4) and (5). Once  $x_a(k) = x_b(k)$ , for any student  $i$ , we get that

$$\min_x \sum_{t \in Q} \Pr [q = t] \sum_{j=0}^n \binom{n}{j} (\Pr [s_i = a | q = t])^j \cdot (\Pr [s_i = b | q = t])^{n-j} x(k) \quad (17)$$

subject to:

$$-P_{ab}x(k+1) + P_{bb}x(k) \geq \frac{c}{R} \quad \forall k \in \{0, \dots, n-1\} \quad (18)$$

$$P_{aa}x(k+1) - P_{ab}x(k) \geq \frac{c}{R} \quad \forall k \in \{0, \dots, n-1\} \quad (19)$$

$$-P_{ab}x(k+1) + P_{bb}x(k) \geq 0 \quad \forall k \in \{0, \dots, n-1\} \quad (20)$$

$$P_{aa}x(k+1) - P_{ab}x(k) \geq 0 \quad \forall k \in \{0, \dots, n-1\} \quad (21)$$

$$0 \leq x(k) \leq 1 \quad \forall k \in \{0, \dots, n\} \quad (22)$$

The derivation of Constraints (18)–(21) is similar to that of Constraints(8)–(15) in the proof of Theorem 3.3. Observe that our optimization problem is now a linear program and that Constraints (20) and (21) are implied by Constraints (18) and (19); thus, they can be removed without changing the optimal solution.

We start from the optimal DSIC ROS solution  $x^*$  and decrease  $x(n)$  by  $\Delta x(n) = \frac{c}{R} \cdot \frac{P_a - P_b}{P_{bb} - P_{ab}} \cdot \frac{P_{bb}}{P_{aa}P_{bb} - (P_{ab})^2}$  and the rest of the spot checking probabilities by

$$\Delta x(k) = \left( \frac{P_{bb}}{P_{ab}} \right) \Delta x(k-1), \quad \forall k \in \{0, \dots, n\}. \quad (23)$$

At  $x^*$ , Constraint (18) becomes tight, following directly from our construction of the step size given by (23). Constraint (19) is tight when  $k = n - 1$ ; the rest of the constraints are easily satisfied. Hence, we only need to show that Constraint (19) will not be violated by the specified decrease in the decision variables. However, by (23) we get that for  $k \in \{0, \dots, n - 1\}$ , the ratios by which the gap between the left and right hand sides of Constraints (19) decreases is  $\left( \frac{P_{ab}}{P_{bb}} \right)^{(n-k)} \leq 1$  times the ratio by which the gap for the constraint corresponding to  $n = k - 1$  is decreasing. Therefore, the constraint corresponding to  $n = k - 1$  binds faster.

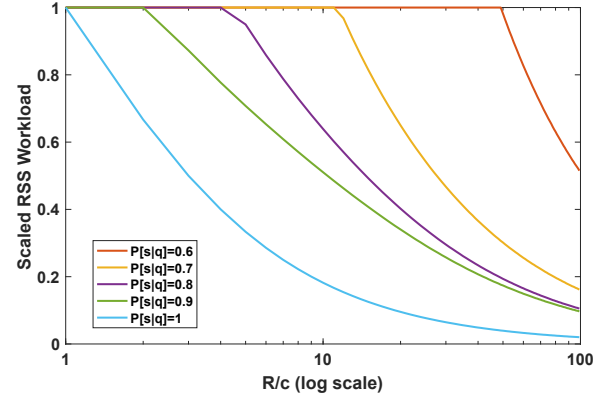
As a result, since Constraint (18) and (19) when  $k = n - 1$  are all binding, we can not decrease any of the decision variables anymore without increasing another. Since every objective coefficient is positive, increasing  $x(k)$  for any  $k \in \{0, \dots, n\}$  could be beneficial only if it resulted in decreasing the value of the rest of the decision variables, achieving an overall objective function improvement. However, if  $x(k + 1)$  is increased then  $x(k)$  needs to increase as well to preserve feasibility; if  $x(k)$  increases then  $x(k - 1)$  needs to increase as well; and so on. Thus, overall, there exist no local, objective-improving changes to the current values of  $x(k)$ , and so we have identified an optimal solution to our linear program. Finding the intersection point of the binding constraints shows that  $x(n) = x_a(k) \leq x(n - 1) = x_b(k) < x(n - 2), \dots, < x(0)$ , where  $x_a(k)$  and  $x_b(k)$  are defined by (4) and (5) for any  $k \in \{0, \dots, n\}$ . This leads to the result of the theorem statement.<sup>4</sup> □

## 4.2 Experimental Comparisons

In this section, we use numerical experiments to quantify performance differences between the optimal DSIC RSS mechanism and the optimal DSIC ROS mechanism. For simplicity we consider  $\Pr[s_i = a|q = a] = \Pr[s_i = b|q = b]$  (other settings yield qualitatively similar results but have more parameters over which to optimize).

First, we consider the case where each assignment is graded by three students, as occurs in various practical peer grading systems [2, 24]. Observe that TA workload can be computed given three quantities:  $\frac{R}{c}$ , the factor by which the reward exceeds the cost of effort;  $\Pr[q]$ , the prior over an assignment’s true quality; and  $\Pr[s|q]$ , the prior over the grader’s signal given a true quality. Given these quantities, let the *scaled RSS workload* be the per-assignment TA workload required by the optimal DSIC RSS mechanism, expressed as a fraction of the per-assignment TA workload required by the optimal DSIC ROS mechanism. (Note that a scaled RSS workload of 1 means that RSS offers no benefits over ROS, and scaled RSS workloads approaching 0 correspond to TA workloads under RSS that approach a 0% fraction of those under ROS.) To investigate the

<sup>4</sup>Note that we actually prove more than claimed in the theorem statement: our result holds not only for the ICCP solution concept, but also for all strategies in which no students invest grading effort.



**Figure 1: Scaled RSS workload as a function of  $R/c$  (on a log scale) and for different values of  $\Pr[s|q]$ , and for the value of  $\Pr[q]$  that minimizes scaled RSS workload.**

scaled RSS workload empirically, we varied  $\Pr[s|q]$  in 10% increments from 0.6 (the signal gives little information about the true grade) to 1 (the signal perfectly identifies the true grade) and then set  $\Pr[q]$  to minimize scaled RSS workload.<sup>5</sup> The result is shown in Figure 1. We make two key observations. First, scaled RSS workload starts out at 1 when rewards are similar to costs, but falls towards zero as  $R/c$  increases. Second, holding  $R/c$  constant, scaled RSS workload decreases as graders become more accurate (i.e., as  $\Pr[s|q]$  approaches 1).

Second, we investigate the relationship between the number of students grading each assignment and the scaled RSS workload. Observe that the optimal ROS mechanism gives rise to the same TA workload regardless of the number of students, while TA workload does change with the number of students under the optimal RSS mechanism. Figure 2 shows scaled RSS workload as a function of number of students for a setting with a moderately skewed prior distribution ( $\Pr[q = a] = 0.8$ ), relatively accurate graders ( $\Pr[s_i = a|q = a] = 0.9$ ), and a big gap between rewards and costs ( $\frac{R}{c} = 25$ ).<sup>6</sup> Overall, as the number of students grading each assignment increases, the scaled RSS workload increases (and hence the RSS mechanism offers less benefit over the ROS mechanism). To make this pattern more concrete, consider that when three students grade each assignment and follow their dominant strategies, the optimal ROS mechanism has a TA workload of 0.5 while the optimal RSS mechanism has a TA workload of 0.18 (a scaled RSS workload of 0.36); whereas, when 10 students grade each assignment, the ROS workload remains at 0.5 while the TA workload of the RSS mechanism increases to 0.23 (for a scaled RSS workload of 0.47). As

<sup>5</sup> We already know that from Corollary 4.2 that for every value of  $\Pr[s|q]$  there exists a value of  $\Pr[q]$  under which the scaled RSS workload is 1: the  $\Pr[q]$  for which  $\sum_q \Pr[s|q] \cdot \Pr[q]$  is uniform. Furthermore, note that for any fixed value of  $\Pr[q]$ , scaled RSS workload is unaffected by changes to  $R/c$  (consult Equations (3), (4), and (5)). However, Theorem 3.4 shows that as  $\frac{R}{c}$  increases, a wider range of  $\Pr[q]$  and  $\Pr[s|q]$  distributions become feasible, allowing for smaller scaled RSS workloads. We thus focus on characterizing the smallest RSS workload that can be achieved for given values of  $R/c$  and  $\Pr[s|q]$ .

<sup>6</sup>Other values of these parameters yield different quantitative results but the same qualitative trends.

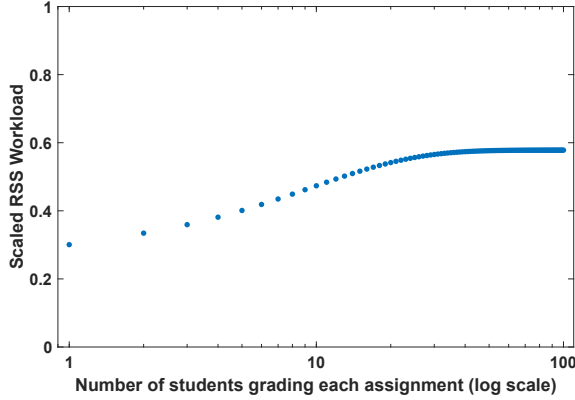


Figure 2: Scaled RSS Workload as a function of number of students.

the number of students grows, the scaled RSS workload approaches 0.58.

## 5 HETEROGENEOUS SIGNAL DISTRIBUTIONS

So far, our analysis has been limited to a model in which students' and TA's signals are drawn from the same signal distribution. In this section, we show that our main structural result, Theorem 3.3, in fact extends to a model in which students' and TA's signals are from different distributions.

More formally, suppose each student  $i$ , by exerting effort at cost  $c_i$ , can examine the submission and observe a signal  $s_i \in Q$ ; conditioning on the submission's quality  $q$ , each signal  $s_i$  is drawn independently from a commonly known distribution  $\Pr[s_i|q]$ . Also, the TA's signal  $s_{TA}$  is independently drawn from distribution  $\Pr[s_{TA}|q]$ . Note that the key difference from the setting in the first few sections is that  $\Pr[s_i|q]$  and  $\Pr[s_{TA}|q]$  may all be different from each other.

Let  $\vec{r}_{-i}$  be the vector of reports made by all students except  $i$ ;  $x_i^a(\vec{r}_{-i})$  denote the probability of spot checking student  $i$  when she reports  $a$ , while the other students report  $\vec{r}_{-i}$ ; similarly,  $x_i^b(\vec{r}_{-i})$  denotes the probability of spot checking student  $i$  when she reports  $b$ .

**THEOREM 5.1.** *If for any student  $i$  and signal  $s_i \in Q$ ,*

$$\min_{l \in Q} \{\Pr[s_i = l \mid s_{TA} = l] - \Pr[s_i = l]\} \geq \frac{c_i}{R},$$

*the PRSS mechanism with the following spot-checking probabilities is DSIC and minimizes the TA workload among all DSIC spot-checking mechanisms:*

$$x_i^a(\vec{r}_{-i}) = \frac{\frac{c_i}{R}}{\Pr[s_i = b \mid s_{TA} = b] - \Pr[s_i = b]} \quad \forall \vec{r}_{-i} \in Q^{n-1} \quad (24)$$

$$x_i^b(\vec{r}_{-i}) = \frac{\frac{c_i}{R}}{\Pr[s_i = a \mid s_{TA} = a] - \Pr[s_i = a]} \quad \forall \vec{r}_{-i} \in Q^{n-1}; \quad (25)$$

*if  $\min_{l \in Q} \{\Pr[s_i = l \mid s_{TA} = l] - \Pr[s_i = l]\} < \frac{c_i}{R}$  for any  $i$  and  $l$ , there exists no DSIC spot-checking mechanism.*

**PROOF SKETCH.** We first note that Eqs. (8) and (9) in the proof of Theorem 3.3 are defined for each student  $i$  and do not depend on the signal distribution of any student except student  $i$ . When the signal distribution for the TA and Students are different, the only change in (8) and (9) is that instead having  $n$  pair of such constraints and  $2n$  variables, there are  $2^{n-1}$  constraints and  $2^n$  variables, with  $x_i^a(\vec{r}_{-i})$  and  $x_i^b(\vec{r}_{-i})$  as a pair for each  $\vec{r}_{-i}$ . Let  $\vec{s}_{-i}$  be the signal observed by all students except student  $i$ . The objective function in minimization problem defined in the proof of Theorem 3.3 for minimizing the TA workload subject to (8) and (9) can be redefined as follows:

$$\min_{x_i^a, x_i^b} \sum_{\vec{s} \in Q^n} \sum_{t \in Q} \Pr[q = t] \cdot \prod_{j \in N} \Pr[s_j = l_j | q = t] \max_i \{x_i^{s_i}(\vec{s}_{-i})\}$$

with  $0 \leq x_i^a(\vec{r}_{-i}), x_i^b(\vec{r}_{-i}) \leq 1$  for all  $\vec{r}_{-i} \in Q^{n-1}$  and  $i$ , which is again a convex function due to the presence of the max functions. Since the constraints are again linear and with the same structure, the same argument about the optimality of the intersection of the constraints can be applied here. This means that for each  $\vec{r}_{-i} \in Q^{n-1}$ , the intersection of the corresponding pairs of constraints defined similar to (8) and (9), which are expressed by (24) and (25), give optimal solution to the above optimization problem. The arguments in the Step 2 of the proof of Theorem 3.3 are also directly applicable here. This completes the proof.  $\square$

## 6 CONCLUSION

We have investigated peer grading mechanisms that achieve dominant strategy incentive compatibility by using TAs to spot check students, and have focused on minimizing the required TA workload. We have explored mechanisms for *report-sensitive spot checking*: varying spot-checking probabilities based on the profile of grades that all students report for a given assignment. We proposed a simple optimal DSIC ‘‘PRSS’’ mechanism, and showed that it minimizes the required spot-checking budget (across both the ‘‘RSS’’ mechanisms and the more constrained ‘‘RSUS’’ mechanisms even under a weaker solution concept) and outperforms the (‘‘ROS’’) mechanisms that spot checks all students with a fixed, report-oblivious probability. We evaluated the performance of the optimal DSIC RSS mechanism both analytically and empirically. Finally, we extended our results to a setting which allowed each student and TA to observe the ground truth according to a different noise model.

We consider the most important direction for future work to be generalizing our results beyond two signals. We note that this would require a fundamentally different proof technique, as our convex programming formulation depends critically on the problem's two-signal structure. We expect that the multi-signal setting would also require other variations in the model. Notably, in such domains it becomes natural to impose an ordering over the signals and to reward agents according to the distance between their reports and that of the TA, rather than rewarding all ‘‘correct’’ reports equally. Another limitation of our work is the assumption that the prior distribution over signals is known to the mechanism designer. The derivation of prior-independent report-sensitive mechanisms is a second, worthwhile direction for future work. One possible strategy for building such mechanisms could be learning the prior in a repeated setting.



## REFERENCES

- [1] Anirban Dasgupta and Arpita Ghosh. 2013. Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd International Conference on World Wide Web*. ACM, 319–330.
- [2] Luca De Alfaro and Michael Shavlovsky. 2014. CrowdGrader: A tool for crowdsourcing the evaluation of homework assignments. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*. ACM, 415–420.
- [3] Luca De Alfaro, Michael Shavlovsky, and Vassilis Polychronopoulos. 2016. Incentives for truthful peer grading. *arXiv preprint arXiv:1604.03178* (2016).
- [4] Boi Faltings, Jason J Li, and Radu Jurca. 2012. Eliciting truthful measurements from a community of sensors. In *3rd IEEE International Conference on the Internet of Things*. IEEE, 47–54.
- [5] Alice Gao, Andrew Mao, Yiling Chen, and Ryan P. Adams. 2014. Trick or treat: putting peer prediction to the test. In *Proceedings of the 15th ACM Conference on Economics and Computation*. ACM, 507–524.
- [6] Alice Gao, James R Wright, and Kevin Leyton-Brown. 2016. Incentivizing evaluation via limited access to ground truth: Peer-prediction makes things worse. *arXiv preprint arXiv:1606.07042* (2016).
- [7] Naman Goel and Boi Faltings. 2019. Deep bayesian trust: A dominant and fair incentive mechanism for crowd. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 1996–2003.
- [8] Radu Jurca and Boi Faltings. 2005. Enforcing truthful strategies in incentive compatible reputation mechanisms. In *International Workshop on Internet and Network Economics*. Springer, 268–277.
- [9] Radu Jurca and Boi Faltings. 2009. Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research* 34 (2009), 209–253.
- [10] Vijay Kamble, David Marn, Nihar Shah, Abhay Parekh, and Kannan Ramachandran. 2015. Truth Serums for Massively Crowdsourced Evaluation Tasks. *arXiv preprint arXiv:1507.07045* (2015).
- [11] Yuqing Kong, Katrina Ligett, and Grant Schoenebeck. 2016. Putting Peer Prediction Under the Micro (economic) scope and Making Truth-telling Focal. In *International Conference on Web and Internet Economics*. Springer, 251–264.
- [12] Yang Liu and Yiling Chen. 2018. Surrogate Scoring Rules and a Dominant Truth Serum. *arXiv preprint arXiv:1802.09158* (2018).
- [13] Michael R Merrifield and Donald G Saari. 2009. Telescope time without tears: A distributed approach to peer review. *Astronomy & Geophysics* 50, 4 (2009), 4–16.
- [14] Nolan Miller, Paul Resnick, and Richard Zeckhauser. 2005. Eliciting informative feedback: The peer-prediction method. *Management Science* 51, 9 (2005), 1359–1373.
- [15] Dražen Prelec. 2004. A Bayesian truth serum for subjective data. *Science* 306, 5695 (2004), 462–466.
- [16] Goran Radanovic and Boi Faltings. 2013. A robust Bayesian truth serum for non-binary signals. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence*. 833–839.
- [17] Goran Radanovic and Boi Faltings. 2014. Incentives for truthful information elicitation of continuous signals. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence*. 770–776.
- [18] Blake Riley. 2014. Minimum truth serums with optional predictions. In *Proceedings of the 4th Workshop on Social Computing and User Generated Content*.
- [19] Victor Shnayder, Arpit Agarwal, Rafael Frongillo, and David C Parkes. 2016. Informed truthfulness in multi-task peer prediction. In *Proceedings of the 2016 ACM Conference on Economics and Computation*. ACM, 179–196.
- [20] Bo Waggoner and Yiling Chen. 2014. Output agreement mechanisms and common knowledge. In *2nd AAAI Conference on Human Computation and Crowdsourcing*.
- [21] Wanyuan Wang, Bo An, and Yichuan Jiang. 2018. Optimal Spot-Checking for Improving Evaluation Accuracy of Peer Grading Systems. In *32nd AAAI Conference on Artificial Intelligence*. 833–840.
- [22] Jens Witkowski, Yoram Bachrach, Peter Key, and David C Parkes. 2013. Dwelling on the negative: Incentivizing effort in peer prediction. In *1st AAAI Conference on Human Computation and Crowdsourcing*.
- [23] Jens Witkowski and David C Parkes. 2012. A robust Bayesian truth serum for small populations. In *26th AAAI Conference on Artificial Intelligence*. 1492–1498.
- [24] James R Wright, Chris Thornton, and Kevin Leyton-Brown. 2015. Mechanical TA: Partially automated high-stakes peer grading. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*. ACM, 96–101.
- [25] Yuanzhang Xiao, Florian Dörfler, and Mihaela Van Der Schaar. 2018. Incentive design in peer review: Rating and repeated endogenous matching. *IEEE Transactions on Network Science and Engineering* 6, 4 (2018), 898–908.