DECISION ANALYSIS
Vol. 9, No. 3, September 2012, pp. 253–273
ISSN 1545-8490 (print) | ISSN 1545-8504 (online)
http://dx.doi.org/10.1287/deca.1120.0247

Sharing Rewards Among Strangers Based on Peer Evaluations

Arthur Carvalho, Kate Larson
Cheriton School of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
{a3carval@cs.uwaterloo.ca, klarson@cs.uwaterloo.ca}

We study a problem where a new, unfamiliar group of agents has to decide how a joint reward should be shared among them. We focus on settings where the share that each agent receives depends on the evaluations of its peers concerning that agent's contribution to the group. We introduce a mechanism to elicit and aggregate evaluations as well as for determining agents' shares. The intuition behind the proposed mechanism is that each agent has its expected share maximized to the extent that it is well evaluated by its peers and that it is truthfully reporting its evaluations. For promoting truthfulness, the proposed mechanism uses a peer-prediction method built on strictly proper scoring rules. Under the assumption that agents are Bayesian decision makers, we show that our mechanism is incentive compatible and budget balanced. We also provide sufficient conditions under which the proposed mechanism is individually rational, resistant to some kinds of collusion, and fair.

Key words: fair division; peer-prediction methods; scoring rules
History: Received on February 13, 2012. Accepted on June 19, 2012, after 1 revision. Published online in *Articles in Advance* August 3, 2012.

1. Introduction

Understanding how agents can work together to achieve some common goal is a central research topic when dealing with multiagent systems (Shoham and Leyton-Brown 2009). Questions that are typically analyzed include how and which groups of agents should form (Rahwan et al. 2009); how agents should coordinate their actions once they have agreed to work together (Grosz and Kraus 1996); how to ensure that the group, once formed, does not disintegrate (Conitzer and Sandholm 2006); and how a joint reward should be divided among the group members (Moulin 2004). It is this last question that we address in this paper.

Commonly called *fair division*, the problem of dividing one or several goods among a set of agents, in a way that satisfies a suitable fairness criterion, has been studied in many literatures. In decision analysis, several fairness factors and practical issues are taken into account before constructing a decision analysis model for allocating goods (Keller et al. 2010), e.g., how uniform the allocation is, its impact on future generations, etc. In economics, the collective welfare approach is arguably the most influential application of the economic analysis to fair division. It uses the concepts of collective utility functions, in its cardinal interpretation, and social welfare orderings, in its ordinal interpretation, for deciding what makes a reasonable division (Moulin 2004). In computer science and, more specifically, artificial intelligence, the fair division problem is traditionally studied in settings where the underlying agents not only have preferences over alternative allocations of goods, but also actively participate in computing an allocation (Chevaleyre et al. 2006).

In this paper, we propose a game-theoretic model for sharing a joint, homogeneous reward based on the idea of *peer evaluations*. In detail, we consider scenarios where an unfamiliar group has been formed and has accomplished a task for which it is granted a reward, which in turn must be shared among the group members. In some sense, our work can be seen as a complement of the ongoing research on ad hoc teams (e.g., Stone et al. 2010). In such ad hoc team settings, an agent must be prepared to cooperate with previously unfamiliar teammates. Thus, team strategies cannot be developed a priori. Rather than focusing on the formation process of ad hoc teams, we propose a way of sharing rewards that might result from that collaboration.

After observing the individual contributions of the peers in accomplishing the joint task, each agent in the group is asked to *evaluate* the others. These evaluations are elicited and aggregated by a central, trusted entity called the *mechanism*, which is also responsible for sharing the joint reward. The share received by each agent from our proposed mechanism has two major components. The first one reflects the evaluations received by that agent. The second one is a truth-telling score, which is used to encourage agents to truthfully report their evaluations. For computing such scores, the mechanism uses a novel peer-prediction method built on strictly proper scoring rules. Under our method, truth-telling maximizes agents' expected scores when they do not have informative prior knowledge about the capabilities of their peers.

Thus, the intuition behind the proposed mechanism is that each agent has its expected share maximized to the extent that it is well evaluated and that it is telling the truth. Under the assumption that agents are Bayesian decision makers, we show that our mechanism is incentive compatible and budget balanced, and we present sufficient conditions under which it is individually rational, resistant to some kinds of collusion, and fair.

The rest of this paper is organized as follows. We review the literature related to our work in §2. In §3, we describe the basic model and concepts used throughout the paper. In §4, we introduce our mechanism and show that it satisfies important properties. In §5, we investigate the influence of the mechanism's parameter on agents' shares. In §6, we discuss some practical issues regarding our model and mechanism. We conclude in §7.

2. Related Work

Several resource allocation models have been proposed in the decision analysis literature, ranging from resource allocation in the development of military countermeasures (Golany et al. 2012) to resource allocation in healthcare (Griffin et al. 2008). A well-known resource allocation model is attributable to Pratt and Zeckhauser (1990). They described how to fairly divide a number of silver heirlooms among agents, where each agent's preferences for objects are represented by its utility function. Then, allocations to different agents are made as they would be in a market for probability shares in the objects, where each agent is assumed to have a fixed and equal budget for making purchases, and the objects' prices are the market-clearing equilibrium prices in a second-price auction. We observe that such an auction-based approach is unfeasible in our setting because the underlying object (reward) is homogeneous and divisible.

Several fairness factors and practical issues are also studied in decision analysis. Questions concerning whether or not it matters what different agents get, or how uniform the allocation is, are often taken into account before constructing decision analysis models and processes for allocating resources (Keller et al. 2010).

From a more theoretical perspective, the concept of fair division has long been studied in cooperative game theory. The Shapley (1953) value is a key concept in this field used to distribute a joint surplus (or cost) among a set of agents. Roughly speaking, the Shapley value assigns a share to each agent equal to that agent's marginal contribution to the group. We note that sharing schemes based on marginal contributions, like the Shapley value, are not appropriate in our setting. The idea of marginal contribution is not objectively defined in our model because individual contributions are subjective information.

Recent work in microeconomics has addressed the issue of sharing a value based on peer evaluations. Slightly different from our model, agents are asked to report the relative contributions of their peers in accomplishing a joint task; e.g., agent 1 might report that agent 3 deserves 10 times as much as agent 2. Three properties of mechanisms are very often studied: budget balance, impartiality, and consensuality. A mechanism is budget balanced if it neither takes a loss nor makes a profit. A mechanism is impartial if the amount paid to every agent is independent of that agent's report. A mechanism is consensual if the amount paid to agents respect the peer evaluations whenever they are consistent. With exactly three agents, de Clippel et al. (2008) show that there is a unique impartial and consensual mechanism, and that it is only budget balanced when agents' evaluations are consistent. Those authors also propose a family of mechanisms that satisfies the aforementioned properties whenever there are four or more agents. The follow-up paper by Tideman and Plassmann (2008) elaborates on the numerical and statistical properties of that family of mechanisms. Knoblauch (2009) determined the extent to which an impartial, budget balanced mechanism must deviate from consensuality in the three-agent case.

A crucial assumption in the aforecited works is that the underlying agents are not strategic. The rationale behind such an assumption is that because the evaluations reported by an agent do not affect its share of the joint reward (impartiality), then that agent should have no incentive to lie. First, we note that this assumption makes the resulting mechanisms extremely susceptible to collusive behavior. Furthermore, it does not help to prevent situations where agents are not willing to exert the required effort to evaluate their peers, or when agents' evaluations are biased.

This last point is illustrated in the work done by Kaufman et al. (2000). These authors use a peer-rating system to account for individual performance in teams of students. In detail, each team member initially receives a common grade as the result of a joint academic work. Later, each grade is weighted by a weighting factor, which is equal to the average evaluation received by the underlying student divided by the team average. This peer-rating system was used in assignments of two sophomore-level courses. The results strongly suggest that both gender and racial bias might influence the reported ratings. Furthermore, several cases of collusion between students are reported, i.e., team members agreeing to give each other identical ratings.

We believe that the above-mentioned problems might be even worse in our scenario because the joint reward must be shared among the agents. Hence, the truth-telling score is an essential component of our mechanism because it encourages rational agents to truthfully reveal their evaluations. To the best of our knowledge, the first mechanism that explicitly deals with strategic agents when sharing rewards based on peer evaluations is attributable to Carvalho and Larson (2011). For promoting truthfulness, they use the "Bayesian truth serum method" proposed by Prelec (2004). This method brings two major drawbacks. First, it requires the population of agents to be large. Second, besides reporting evaluations, agents must also make predictions about how their peers are evaluated.

In this work, we eliminate both problems by using a novel peer-prediction method built on strictly proper scoring rules. The first of such methods is attributable to Miller et al. (2005). In their work, a number of agents experience a product and rate its quality. A mechanism then collects the ratings and makes the payments based on them. This method makes use of the stochastic correlation between the signals observed by the agents from the product to achieve a Bayes–Nash equilibrium where every agent tells the truth. Jurca and Faltings (2009) extended the original method by also considering several kinds of collusion.

The major problem with the original peer-prediction method when applied to our setting is that it depends on previous, historical data. In detail, after agent i reports its evaluation of agent j, say s_i^j, the mechanism estimates agent i's prediction of the evaluation reported by another agent k, $P(s_k^j \mid s_i^j)$, which is then evaluated and rewarded using a scoring rule and agent k's actual reported evaluation. The mechanism needs to have a history of previous evaluations for computing $P(s_k^j \mid s_i^j)$. In our scenario, this is unreasonable to assume whenever an agent is being evaluated for the first time.

Witkowski and Parkes (2011) proposed a way to circumvent this problem by removing some crucial common knowledge assumptions. In their model, agents possess private beliefs regarding both the quality of the product and the likelihood of a positive experience given a particular quality. Agents are required to report their beliefs before and after experiencing the product. A mechanism estimates the perceived quality from the direction of the belief change. Each estimation is used as the observed outcome ("ground truth") when rewarding another agent's report. This approach might create many practical problems in our setting. First, it implies that agents have to evaluate their peers before working together, which is a troublesome situation whenever they are working together for the first time. Furthermore, they have to report their entire beliefs (probability distributions) twice. Finally, the estimated perceived quality is binary, which implies that an agent could only do a good job or a poor one.

We balance the budget of our peer-prediction method by using an adaptation of the technique proposed by Goel et al. (2009), which in turn is based on competitive scoring rules (Kilgour and Gerchak 2004). Goel et al. (2009) also proposed a peer-prediction method called the "collective revelation mechanism," which not only admits a truth-telling Bayes–Nash equilibrium, but also weights agents' ratings by their relative information content. The major drawback of that mechanism is that agents must provide an extra piece of information for the mechanism, namely, their willingness to update their believes in light of hypothetical new evidence.

Finally, we note that our setting bears a tenuous relation to bargaining games (Nash 1950), cost-sharing schemes from the mechanism design literature (Moulin 1999, Moulin and Shenker 2001), and the cake-cutting problem (Brams and Taylor 1996, Robertson and Webb 1998), where different fairness criteria are studied (e.g., proportionality, equitability, and envy freeness).

3. Model and Background

A set of unknown agents[1] $N = \{1, \ldots, n\}$, for $n \geq 4$, has accomplished a task for which it is granted a *reward* $V \in \Re^+$. Every agent is assumed to want more of the reward. Therefore, we can identify an agent's share with its welfare. We are interested in settings where the share of V that an agent receives depends on the *evaluations* of its peers concerning that agent's contribution to the group.

We model the private information of each agent as $n - 1$ private signals that that agent observes from its peers. These signals are direct assessments of the peers' performance in accomplishing the joint task. We call them *truthful evaluations*. To avoid a biased self-judgment, agents do not do self-evaluations. Formally, given a positive integer parameter $M \geq 1$ and for each agent $j \in N$, let $\boldsymbol{\Omega}_j$ be a multinomial distribution[2] with *unknown* parameter $\boldsymbol{\omega}_j \in \Delta^M$ (a unit simplex in $\Re^M$) that represents the truthful evaluations for agent j. The signals observed by agent i are represented by the vector $\mathbf{t}_i = (t_i^1, \ldots, t_i^{i-1}, t_i^{i+1}, \ldots, t_i^n)$, where $t_i^j \in \{1, \ldots, M\}$ represents the signal observed by agent i coming from agent j, i.e., $t_i^j \sim \boldsymbol{\Omega}_j$. Therefore, $\mathbf{t}_i$ is the vector with the truthful evaluations made by agent i regarding the contributions of its peers in accomplishing the joint task.

The parameter M represents the top possible evaluation that an agent can give or receive, and we assume that its value is common knowledge. Intuitively, as M increases, the evaluations might be more fine grained in that small differences between agents can be recognized by their peers. However, this increased expressivity might be burdensome for some agents because they will have more freedom to evaluate their peers, which can make the evaluation process more challenging. We argue that the underlying application might help to determine appropriate settings for M. By assuming that the lowest possible evaluation that an agent can receive is equal to 1, we implicitly assume that every agent contributed to the task. Thus, agents who did not contribute at all must not participate in the sharing process. We make the following additional assumptions in our model:

1. *Self-interest*. Agents act to maximize their expected shares.
2. *Common prior*. For all $j \in N$, there exists a common prior distribution over $\boldsymbol{\omega}_j$, $P(\boldsymbol{\omega}_j)$. We assume that this distribution is a *noninformative Dirichlet distribution*.
3. *Rationality*. Every agent $i \in N$, with truthful evaluation t_i^j, forms a posterior by applying Bayes' rule to the common prior $P(\boldsymbol{\omega}_j)$, i.e., $P(\boldsymbol{\omega}_j \mid t_i^j)$.
4. *Independent signals*. The signals observed by an agent are independent of each other. Formally, given $i, j, k \in N$, and $x, y \in \{1, \ldots, M\}$, $P(t_i^j = x \mid t_i^k = y) = P(t_i^j = x)$.

The first assumption means that agents are *risk neutral* (Mas-Colell et al. 1995). The second assumption implies that agents have common prior distributions over distributions of truthful evaluations. We discuss the meaning and validity of the assumption of noninformative Dirichlet priors in §3.5. The third

[1] Because agents are not necessarily human beings (e.g., computational agents), we refer to a single one of them as *it*.

[2] We use the term *multinomial distribution* (also known as a categorical distribution) for the generalization of the Bernoulli distribution for discrete random variables with any constant number of values. The parameter of that distribution is a probability vector that specifies the probability of each possible outcome.

assumption means that the posterior distributions are consistent with Bayesian updating, i.e.,

$$P(\boldsymbol{\omega}_j \mid t_i^j) = KP(t_i^j \mid \boldsymbol{\omega}_j)P(\boldsymbol{\omega}_j), \tag{1}$$

where K is a normalizing constant that ensures that the posterior adds up to 1. The first three assumptions are traditional in both game theory and multiagent systems literature (Osborne and Rubinstein 1994, Shoham and Leyton-Brown 2009). Together, they imply that agents are *Bayesian decision makers*. Finally, the last assumption implies that the truthful evaluation of an agent for a peer does not influence that agent's truthful evaluation of other peers.

A consequence of self-interest is that agents may deliberately lie when reporting their evaluations. For example, an agent might intentionally give all other agents a low evaluation so that, in comparison, it looks good and receives a greater share of V. Therefore, we distinguish between the truthful evaluations made by each agent $i \in N$, $\mathbf{t}_i$, and the evaluations that agent i reports, $\mathbf{s}_i(\mathbf{t}_i) = (s_i^1(t_i^1), \ldots, s_i^{i-1}(t_i^{i-1}), s_i^{i+1}(t_i^{i+1}), \ldots, s_i^n(t_i^n))$. We call $\mathbf{s}_i(\mathbf{t}_i)$ the strategy of agent i. For clarity's sake, we use $\mathbf{s}_i$ to represent agent i's strategy when $\mathbf{t}_i$ is irrelevant or clear from the context. S_i is the set of strategies available to agent i, and $S = S_1 \times \cdots \times S_n$. Each vector $\mathbf{s} = (\mathbf{s}_1(\mathbf{t}_1), \ldots, \mathbf{s}_n(\mathbf{t}_n)) \in S$ is a *strategy profile*.

As customary, let the subscript "$-i$" denote a vector without agent i's component, e.g., $\mathbf{s}_{-i}(\mathbf{t}_{-i}) = (\mathbf{s}_1(\mathbf{t}_1), \ldots, \mathbf{s}_{i-1}(\mathbf{t}_{i-1}), \mathbf{s}_{i+1}(\mathbf{t}_{i+1}), \ldots, \mathbf{s}_n(\mathbf{t}_n))$. If the evaluations reported by agent i are equal to its truthful evaluations, i.e., $\mathbf{s}_i(\mathbf{t}_i) = \mathbf{t}_i$, then we say that agent i's strategy is *truthful*. If for all $i \in N, \mathbf{s}_i(\mathbf{t}_i) = \mathbf{t}_i$, then we say that $\mathbf{s}$ is the *collective truth-telling strategy profile*. We represent it by $\overset{*}{\mathbf{s}}$. Evaluations are elicited and aggregated by a central, trusted entity called the *mechanism*, which is also responsible for sharing the reward among the agents. Formally:

Definition 1 (Mechanism). A mechanism is a sharing function, $\Gamma\colon S \to \Re^n$, which maps each strategy profile to a vector of shares.

Given the strategy profile $\mathbf{s}$, we denote the share of V given to agent i by $\Gamma_i(\mathbf{s})$. We use Γ_i when $\mathbf{s}$ is either irrelevant or clear from the context. Throughout this paper, we use the solution concept called *Bayes–Nash equilibrium*.

Definition 2 (Bayes–Nash Equilibrium). A strategy profile $\mathbf{s}$ is a Bayes–Nash equilibrium if for each agent i and strategy $\mathbf{s}'_i \neq \mathbf{s}_i \in S_i$, $E[\Gamma_i(\mathbf{s}_i(\mathbf{t}_i), \mathbf{s}_{-i}(\mathbf{t}_{-i})) \mid \mathbf{t}_i] \geq E[\Gamma_i(\mathbf{s}'_i(\mathbf{t}_i), \mathbf{s}_{-i}(\mathbf{t}_{-i})) \mid \mathbf{t}_i]$.

In words, for each agent $i \in N$, $\mathbf{s}_i(\mathbf{t}_i)$ is the best response, in an expected sense, that agent i has to $\mathbf{s}_{-i}(\mathbf{t}_{-i})$ given its truthful evaluations $\mathbf{t}_i$. The expectation is taken with respect to the posterior predictive distributions, discussed in §3.5. When the inequality in Definition 2 holds strictly (with "$>$" instead of "$\geq$"), then we say that the strategy profile $\mathbf{s}$ is a *strict Bayes–Nash equilibrium*.

3.1. Numerical Example

In this subsection, we illustrate some of the concepts defined so far. The same example will be extended in subsequent sections to illustrate new concepts and results. Consider four agents indexed by the letters A, B, C, D, a joint reward $V = 100$, and the parameter $M = 10$. Furthermore, suppose that the truthful evaluations are the ones shown in Table 1, where each numeric cell can be interpreted as the signal observed by the agent in the row coming from the agent in the column, e.g., the emphasized number 2 represents t_A^B, i.e., the signal observed by agent A coming from agent B. Now, suppose that agent A wants to downgrade others' contributions. Hence, it could report the vector $\mathbf{s}_A = (1, 1, 1)$ instead of telling the truth and reporting $\mathbf{s}_A = \mathbf{t}_A = (2, 9, 6)$.

3.2. Properties

There are several key properties we wish mechanisms to have. We introduce them in this section.

Definition 3 (Fairness). Consider any strategy profile $\mathbf{s} \in S$, where $s_k^i > s_k^j$ for every agent $k \neq i, j$, and $s_j^i > s_i^j$. Then, we say that a mechanism is fair if $\Gamma_i(\mathbf{s}) > \Gamma_j(\mathbf{s})$.

Table 1 **Truthful Evaluations**

	Reported evaluations			
Agent	*A*	*B*	*C*	*D*
A	—	*2*	9	6
B	8	—	10	9
C	7	5	—	6
D	7	1	10	—

In words, if an agent unanimously receives better evaluations than a peer, then that agent should also receive a greater share of the joint reward than its peer.

Definition 4 (Budget Balance). If $\forall \mathbf{s} \in S$, $\sum_{i=1}^{n} \Gamma_i(\mathbf{s}) = V$, then the mechanism is budget-balanced.

In words, a budget-balanced mechanism allocates the entire reward V back to the agents. As stated, this is a strong definition because we do not put constraints on $\mathbf{s}$, e.g., we do not require $\mathbf{s}$ to be an equilibrium strategy profile.

Definition 5 (Individual Rationality). A mechanism is individually rational if $\forall i \in N$, $\forall \mathbf{s} \in S$, $\Gamma_i(\mathbf{s}) \geq 0$.

This condition requires the share received by each agent to be greater than or equal to zero. In other words, all agents are weakly better off participating in the sharing process than not participating at all.

Definition 6 (Incentive Compatibility). A mechanism is incentive compatible if collective truth-telling is an equilibrium strategy profile.

Because we are working with Bayes–Nash equilibria, an incentive-compatible mechanism implies that it is best, in an expected sense, for each agent to tell the truth provided that the others are also doing so.

Definition 7 (Collusion Resistance). A mechanism is collusion resistant if agents have no incentive to enter into a priori agreements to undermine the mechanism.

In the following subsection, we extend our discussion on collusions. By no means do we argue that the properties defined in this section are exhaustive. However, we believe that they are among the most desirable ones in practical applications.

3.3. Collusion

We consider that collusion occurs when a group of agents agree to deliberately lie about their evaluations of each other. Formally:

Definition 8 (Collusion). Consider a group of agents G so that $|G| = x$, for $2 \leq x \leq n$. We say that an (x, y)-collusion occurs when $\forall i, j \in G, \forall k, l \notin G$, $s_i^j = y$, for $y \in \{1, \ldots, M\}$, $s_l^j = t_l^j$, $s_i^k = t_i^k$, and $s_l^k = t_l^k$.

In words, an (x, y)-collusion means that x agents deviate from the collective truth-telling strategy profile by giving the evaluation y to each other. For the discussion in this section, let $\tilde{\mathbf{s}}$ be the strategy profile after a collusion among a group of agents $G \subseteq N$. Because of the self-interest assumption, an agent will only collude if its expected share increases by doing so. Consequently, a collusion will only happen if $E[\sum_{i \in G} \Gamma_i(\tilde{\mathbf{s}})] > E[\sum_{i \in G} \Gamma_i(\mathring{\mathbf{s}})]$. Now, we are ready to formally define the *collusion-resistance* property.

Definition 9 (Collusion Resistance). Let $\mathring{\mathbf{s}}$ be the collective truth-telling strategy profile, and let $\tilde{\mathbf{s}}$ be any strategy profile resulting from an (x, y)-collusion among a group of agents G, for $x = |G|$ and $y \in \{1, \ldots, M\}$. We say that a mechanism is (x)-collusion resistant if $E[\sum_{i \in G} \Gamma_i(\tilde{\mathbf{s}})] \leq E[\sum_{i \in G} \Gamma_i(\mathring{\mathbf{s}})]$.

We can strengthen the above definition by saying that a (x, y)-collusion, for any $y \in \{1, \ldots, M\}$, will never be profitable, i.e., $\sum_{i \in G} \Gamma_i(\tilde{\mathbf{s}}) \leq \sum_{i \in G} \Gamma_i(\mathring{\mathbf{s}})$. If a mechanism satisfies this inequality, then we say that it is ex post (x)-collusion resistant. There is an interesting relationship between budget balance and ex post collusion resistance, as shown in the following proposition.

Proposition 1. *If a mechanism is budget balanced, then it is ex post* (n)*-collusion-resistant.*

Proof. If a mechanism is budget balanced, then $\forall \mathbf{s} \in S, \sum_{i=1}^{n} \Gamma_i(\mathbf{s}) = V$. Consequently, a budget-balanced mechanism is trivially ex post (n)-collusion resistant because no strategy profile different from the collective truth-telling strategy profile can make some agent better off without making some other agent worse off. □

Incentive compatibility and collusion resistance are complementary to each other: whereas the former deals with single agents deviating from the collective truth-telling strategy profile, the latter deals with groups of agents. When collective truth telling is a Bayes–Nash equilibrium, our definition of collusion resistance is similar to the concept of resilient equilibrium, e.g., we say that an equilibrium strategy profile is k-resilient if it tolerates deviations by collusions of up to k agents (Halpern 2008). In this paper, we focus on collusions of size 2 because they are more tractable (from a mathematical point of view) and, arguably, more likely to occur than bigger ones.

An interesting point to discuss is the relationship between the collusion-resistance property defined in this paper and the group strategy-proof concept from the mechanism design literature (e.g., Moulin and Shenker 2001). Roughly speaking, a group strategy-proof mechanism in our setting implies that there is no collusion G, of any size, such that every agent in G is at least as well off as when it truthfully reports its evaluations, and at least one agent in G is strictly better off. Although both concepts deal with a group of colluders deviating from the collective truth-telling strategy profile, group strategy-proofness is stronger in a sense that a group strategy-proof mechanism prevents any coalition of agents to gain by lying.

3.4. Scoring Rules

A crucial component of the proposed mechanism is a scoring method for promoting truthfulness, which is based on *strictly proper scoring rules*. Consider an uncertain quantity with possible outcomes $o_1, \ldots, o_M$ and a probability vector $\mathbf{q} = (q_1, \ldots, q_M)$. A *scoring rule* $R(\mathbf{q}, e)$ is a function that provides a score for the assessment $\mathbf{q}$ upon observing the event o_e. A scoring rule is called *strictly proper* when an agent receives its maximal expected score if and only if its stated assessment $\mathbf{q}$ corresponds to its true assessment $\mathbf{z} = (z_1, \ldots, z_M)$ (Savage 1971). The *expected score* of $\mathbf{q}$ at $\mathbf{z}$ for a real value scoring rule $R(\mathbf{q}, e)$ is

$$C(\mathbf{q} \mid \mathbf{z}) = \sum_{e=1}^{M} z_e R(\mathbf{q}, e), \tag{2}$$

and the *expected score loss* is defined by the equation

$$L(\mathbf{q} \mid \mathbf{z}) = C(\mathbf{z} \mid \mathbf{z}) - C(\mathbf{q} \mid \mathbf{z}). \tag{3}$$

Arguably, the best-known strictly proper scoring rules, together with their scoring ranges, are

$$\begin{aligned} \text{logarithmic:} \quad & R(\mathbf{q}, i) = \log q_i \quad (-\infty, 0]; \\ \text{quadratic:} \quad & R(\mathbf{q}, i) = 2q_i - \sum_{e=1}^{M} q_e^2 \quad [-1, 1]; \\ \text{spherical:} \quad & R(\mathbf{q}, i) = \frac{q_i}{\sqrt{\sum_{e=1}^{M} q_e^2}} \quad [0, 1]. \end{aligned} \tag{4}$$

A well-known property of strictly proper scoring rules is that they are still strictly proper under positive affine transformations (for example, see Gneiting and Raftery 2007).

Lemma 1. *If $R(\mathbf{q}, e)$ is a strictly proper scoring rule, then a positive affine transformation of R, i.e., $\gamma R(\mathbf{q}, e) + \phi$, for $\gamma > 0$ and $\phi \in \Re$, is also strictly proper.*

3.5. Dirichlet Distributions

An important assumption in our model is that agents have *noninformative Dirichlet priors* over distributions of truthful evaluations. A *Dirichlet distribution* can be seen as a continuous distribution over parameter vectors of a multinomial distribution. Because $\boldsymbol{\omega}_j$, $\forall j \in N$, is the unknown parameter of the multinomial distribution that describes the truthful evaluations for agent j, then it is natural to consider a Dirichlet distribution as a prior for $\boldsymbol{\omega}_j$. Let $\boldsymbol{\omega}_j = (\omega_{j,1}, \ldots, \omega_{j,M})$. Given a vector of positive reals, $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_M)$, that determines the shape of the Dirichlet distribution, the probability density function of the Dirichlet distribution over $\boldsymbol{\omega}_j$ is

$$P(\boldsymbol{\omega}_j \mid \boldsymbol{\sigma}) = \frac{1}{Z(\boldsymbol{\sigma})} \prod_{k=1}^{M} \omega_{j,k}^{\sigma_k - 1}, \tag{5}$$

where

$$Z(\boldsymbol{\sigma}) = \frac{\prod_{k=1}^{M} (\sigma_k - 1)!}{(\sum_{k=1}^{M} \sigma_k - 1)!}.$$

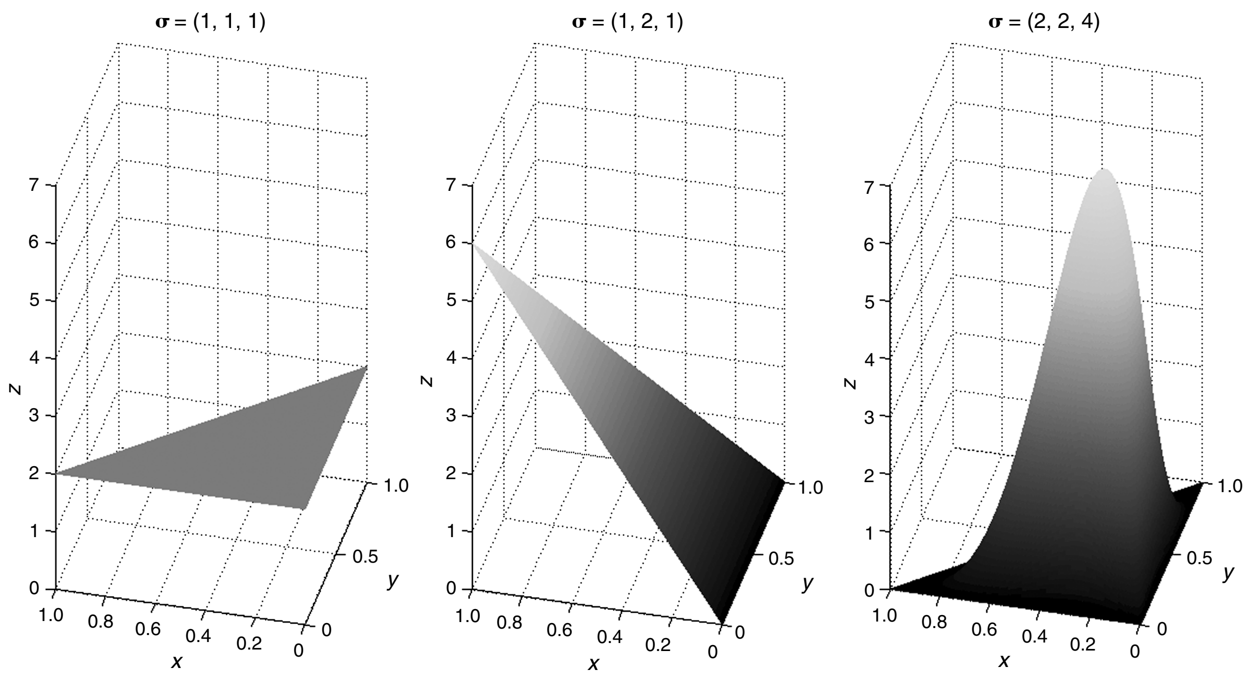
Figure 1 shows the above probability density when $M = 3$ for various parameter vectors $\boldsymbol{\sigma}$. For the Dirichlet distribution in Equation (5), the expected value of, say, $\omega_{j,i}$ is $E[\omega_{j,i} \mid \boldsymbol{\sigma}] = \sigma_i / \sum_{k=1}^{M} \sigma_k$. The probability vector $(E[\omega_{j,1} \mid \boldsymbol{\sigma}], \ldots, E[\omega_{j,M} \mid \boldsymbol{\sigma}])$ is called the *expected distribution* regarding $\boldsymbol{\omega}_j$.

An interesting property of the Dirichlet distribution is that it is the *conjugate prior* of the multinomial distribution (Bernardo and Smith 1994), i.e., the posterior distribution $P(\boldsymbol{\omega}_j \mid \boldsymbol{\sigma}, t_i^j)$, for $i \neq j$, is itself a Dirichlet distribution. This relationship is often used in Bayesian statistics to estimate the hidden parameters of a multinomial distribution. To illustrate this point, suppose that agent i observes the signal $t_i^j = 1$ from agent j. After applying Bayes' rule to its prior (Equation (1)), agent i's posterior distribution is $P(\boldsymbol{\omega}_j \mid \boldsymbol{\sigma}, t_i^j = 1) = P(\boldsymbol{\omega}_j \mid (\sigma_1 + 1, \sigma_2, \ldots, \sigma_M))$. Consequently, the new expected distribution is

$$\boldsymbol{\Theta}_i^j = \left(\frac{\sigma_1 + 1}{1 + \sum_{k=1}^{M} \sigma_k}, \frac{\sigma_2}{1 + \sum_{k=1}^{M} \sigma_k}, \ldots, \frac{\sigma_M}{1 + \sum_{k=1}^{M} \sigma_k} \right).$$

We call the above probability vector agent i's *posterior predictive distribution* regarding $\boldsymbol{\omega}_j$ because it

Figure 1 **Probability Densities of Dirichlet Distributions When $M = 3$ for Different Parameter Vectors**



provides the distribution of future outcomes given the observed data t_i^j. In this way, we can regard the parameters $\sigma_1, \ldots, \sigma_M$ as "pseudocounts" from "pseudodata," where each σ_k can be interpreted as the number of times that the $\omega_{j,k}$-probability event has been observed before. We say that the Dirichlet distribution is *noninformative* (or uniform) when all of the elements making up the vector $\boldsymbol{\sigma}$ have the same value (for simplicity's sake, we assume that this value is equal to 1). Noninformative priors are used when there is no prior knowledge favoring one probability event over another.

In our scenario, a noninformative prior means that before observing, say, agent j's performance in accomplishing the joint task, the initial belief of each agent i regarding the evaluation deserved by agent j is unbiased. Mathematically, this means that the expected distribution, $E[\boldsymbol{\omega}_j]$, is uniform over the set $\{1, \ldots, M\}$. Consequently, the posterior distribution of agent i, $P(\boldsymbol{\omega}_j \mid \boldsymbol{\sigma}, t_i^j)$, indicates that t_i^j is the evaluation most likely to be deserved by agent j because

$$E[\omega_{j,k} \mid \boldsymbol{\sigma}, t_i^j] = \begin{cases} 2/(M+1) & \text{if } t_i^j = k, \\ 1/(M+1) & \text{otherwise.} \end{cases}$$

In short, the assumption of noninformative Dirichlet priors means that each agent's relevant information consists exclusively of its assessments of the peers' performance in accomplishing the joint task. Intuitively, this assumption makes sense because in our model agents are strangers to each other before working together.

We note that other priors could have been used. However, the inference process would not necessarily be analytically tractable.[3] By using noninformative Dirichlet priors, belief updating can be expressed as an updating of the parameters of the prior distribution. Furthermore, we can estimate agents' posterior distributions based solely on their reported evaluations, a point that is explored by the proposed mechanism when computing truth-telling scores.

[3] In general, tractability can be obtained by using conjugate distributions. Hence, another modeling choice is to consider that truthful evaluations follow normal distributions with unknown parameters. Assuming exchangeability, we can then use either the normal-gamma distribution or the normal-scaled inverse gamma distribution as conjugate priors (Bernardo and Smith 1994). The major drawback with this approach is that continuous evaluations might bring extra complexity to the evaluation process.

4. The Mechanism

In this section, we propose a mechanism for sharing rewards based on peer evaluations. For each vector with evaluations, $\mathbf{s}_i$, the mechanism creates a second vector, $\boldsymbol{\zeta}_i = (\zeta_i^1, \ldots, \zeta_i^{i-1}, \zeta_i^{i+1}, \ldots, \zeta_i^n)$, by scaling the elements of $\mathbf{s}_i$ so that they sum up to V. Mathematically,

$$\boldsymbol{\zeta}_i = \left(s_i^1 \frac{V}{\sum_{q \neq i} s_i^q}, \ldots, s_i^{i-1} \frac{V}{\sum_{q \neq i} s_i^q}, s_i^{i+1} \frac{V}{\sum_{q \neq i} s_i^q}, \ldots, s_i^n \frac{V}{\sum_{q \neq i} s_i^q} \right). \quad (6)$$

This simple preprocessing step ensures that the sum of the resulting shares is not orders of magnitude less (or greater) than the reward V. Furthermore, it helps to avoid some deceptive strategies, e.g., when an agent, aiming to look better than its peers and thus to receive a greater share of the joint reward, always gives the lowest possible evaluation to the others. Because the reported evaluations are always scaled to sum up to V, such a deceptive strategy might be ineffective.

The share received by each agent i from the mechanism has two major components. The first one, $\bar{\zeta}^i$, reflects agent i's received evaluations. It is calculated by summing the scaled evaluations received by agent i and dividing the sum by the total number of agents n:

$$\bar{\zeta}^i = \frac{\sum_{j \neq i} \zeta_j^i}{n}. \quad (7)$$

The second component of agent i's share is a truth-telling score τ_i. The intuition behind such scores is that agents maximize their expected values by telling the truth. We provide details about how these scores are calculated in §4.1. Finally, the share of agent i is a linear combination of $\bar{\zeta}^i$ and τ_i:

$$\Gamma_i = \bar{\zeta}^i + \alpha \tau_i, \quad (8)$$

where the constant α, for $\alpha > 0$, fine-tunes the weight given to the truth-telling score τ_i. Thus, the intuition behind the proposed mechanism is that agents have their expected shares maximized to the extent that they are well evaluated and that they truthfully report their evaluations.

4.1. Truth-Telling Scores

Our scoring method is built on scoring rules. As shown in §3.4, scoring rules require an outcome, or a "reality," to score an assessment. If the mechanism knew a priori agents' truthful evaluations, it could compare them to the reported ones and reward agreement. However, because of the subjective nature of the evaluations, we are facing a situation where this objective truth is practically unknowable. Our solution to this issue is to compare reported evaluations and reward agreement. This type of solution is commonly called the *peer-prediction* method (Miller et al. 2005). In the following subsection, we restrict our discussion on truth-telling scores to how to compute a single *peer-prediction score*, μ_i^j, for agent i given its reported evaluation s_i^j. The truth-telling score τ_i is just the arithmetic mean of peer-prediction scores, i.e.,

$$\tau_i = \frac{\sum_{j \neq i} \mu_i^j}{n-1}.$$

4.2. Peer-Prediction Scores

To compute peer-prediction scores based on evaluations for agent j, consider a random ordering of all agents, except agent j. This ordering is known only by the mechanism. For ease of exposition, suppose that agent i is in position i, and that agents i_{+1} and i_{-1} are, respectively, right after and right before agent i in the ordering. Furthermore, consider that agents wrap around for $i+1 > n-1$ and $i-1 < 1$. Intuitively, we use the evaluation given by agent i_{+1} to agent j as a *reference* when computing the peer-prediction score μ_i^j. Let $\boldsymbol{\Phi}_i^j = (\phi_1, \ldots, \phi_M)$ be the *estimated posterior predictive distribution* of agent i regarding $\boldsymbol{\omega}_j$. It is calculated by the mechanism based on agent i's reported evaluation s_i^j, i.e.,

$$\phi_k = \begin{cases} 2/(M+1) & \text{if } s_i^j = k, \\ 1/(M+1) & \text{otherwise.} \end{cases}$$

In short, our peer-prediction method uses the evaluation given by agent i_{+1} to agent j as the realized outcome of an uncertain event (the evaluation deserved by agent j) and rewards agent i's estimated posterior predictive distribution by using a strictly proper scoring rule, i.e.,

$$R(\boldsymbol{\Phi}_i^j, s_{i_{+1}}^j). \quad (9)$$

Technically, R can be any bounded strictly proper scoring rule. Because it has impact on some properties of the proposed mechanism, from now on we assume that R is the quadratic scoring rule[4] shown in Equation (4). The final step toward computing μ_i^j is to normalize the score $R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j)$ by using an adaptation of the technique proposed by Goel et al. (2009) for balancing budgets:

$$\mu_i^j = R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j) - \frac{1}{n-3} \sum_{y \neq i, j, i_{-1}} R(\mathbf{\Phi}_y^j, s_{y_{+1}}^j); \qquad (10)$$

that is, we normalize agent i's score by using the original scores of other agents that cannot be affected by agent i's report (that is why we require $n \geq 4$). In some sense, we are taking into account the group's performance when rewarding agent i. The main reason for doing this is that we can "balance the budget" of our peer-prediction method, as proved in the following proposition.

Proposition 2. *For each agent* $j \in N$, $\sum_{i \neq j} \mu_i^j = 0$.

Proof. The sum of the peer-prediction scores based on the evaluations given to agent j is

$$\begin{aligned}\sum_{i \neq j} \mu_i^j &= \sum_{i \neq j} \left(R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j) - \frac{1}{n-3} \sum_{y \neq i, j, i_{-1}} R(\mathbf{\Phi}_y^j, s_{y_{+1}}^j) \right) \\ &= \sum_{i \neq j} R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j) - \frac{1}{n-3} \sum_{i \neq j} \sum_{y \neq i, j, i_{-1}} R(\mathbf{\Phi}_y^j, s_{y_{+1}}^j) \\ &= 0. \quad \square\end{aligned}$$

An important property of our scoring method is that under the assumptions of *self-interest*, *common prior*, and *rationality*, agents maximize their expected peer-prediction scores by telling the truth, as proved in the following lemma.

Lemma 2. *Under the assumptions of self-interest, a common prior, and rationality, each agent* $i \neq j$ *strictly maximizes its expected peer-prediction score,* $E[\mu_i^j]$, *when* $s_i^j = t_i^j$.

Proof. Because the term

$$\frac{1}{n-3} \sum_{y \neq i, j, i_{-1}} R(\mathbf{\Phi}_y^j, s_{y_{+1}}^j)$$

cannot be affected by the evaluation s_i^j, Equation (10) can be seen as a positive affine transformation of the strictly proper scoring rule $R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j)$. If we take the expectation with respect to $\mathbf{\Theta}_i^j$ (agent i's posterior predictive distribution), then, according to Lemma 1, agent i strictly maximizes its expected score when $\mathbf{\Phi}_i^j = \mathbf{\Theta}_i^j$, which happens when $s_i^j = t_i^j$. $\square$

Another way to interpret the above result is to imagine that agent i is betting on the evaluation deserved by agent j. Because the only relevant information available to agent i is the signal t_i^j, its best strategy (in an expected sense) is to bet on that signal, i.e., to bet on its truthful evaluation. In what follows, we provide bounds for the scores returned by the proposed peer-prediction method. We note that this result does not depend on the assumptions made in our model.

Proposition 3. *For each agent* $i \neq j$, $\mu_i^j \in [-2/(M+1), 2/(M+1)]$.

Proof. See the appendix.

4.3. Numerical Example

Proceeding with the previous example, assume that $\alpha = 50$. Furthermore, assume that all agents are telling the truth, i.e., the reported evaluations are the ones shown in Table 1. The first step taken by the mechanism for sharing $V = 100$ is to scale the reported evaluations of each agent so that they sum up to V (see Equation (6)). For example, the scaled evaluation $\zeta_A^B = s_A^B \times V/(s_A^B + s_A^C + s_A^D) = 2 \times (100/17)$ approximately equals 11.76. The resulting scaled evaluations are shown in Table 2.

After that, the mechanism computes the original, nonnormalized peer-prediction scores (Equation (9)). For illustration's sake, consider the original score received by agent C, $R(\mathbf{\Phi}_C^D, s_{C_{+1}}^D)$, calculated based

Table 2 Scaled Evaluations

	Scaled evaluations			
Agent	*A*	*B*	*C*	*D*
A	—	11.76	52.94	35.30
B	29.63	—	37.04	33.33
C	38.89	27.78	—	33.33
D	38.89	5.55	55.56	—

[4] The proof that the quadratic scoring rule is indeed strictly proper as well as some of its interesting properties can be seen in the paper by Selten (1998).

Table 3 Random Orderings

Agent	Random ordering
A	*CBD*
B	*CDA*
C	*ABD*
D	*BCA*

on its evaluation of agent D, s_C^D. For computing such score, the mechanism starts by estimating agent C's posterior predictive distribution regarding $\boldsymbol{\omega}_D$ as follows:

$$\phi_k = \begin{cases} 2/(10+1) & \text{if } s_C^D = k, \\ 1/(10+1) & \text{otherwise}, \end{cases}$$

which results in the probability vector $\boldsymbol{\Phi}_C^D = (1/11, 1/11, 1/11, 1/11, 1/11, 2/11, 1/11, 1/11, 1/11, 1/11)$ because $s_C^D = 6$. Now, consider the random orderings in Table 3. Each one is indexed by the agent not included in the ordering. According to Table 3, the agent right after agent C in the random ordering indexed by D is agent A. Thus, the mechanism scores the estimated posterior predictive distribution $\boldsymbol{\Phi}_C^D$ using the evaluation given by agent A to agent D as a reference, i.e.,

$$R(\boldsymbol{\Phi}_C^D, s_A^D) \approx 0.256,$$

where R is the quadratic scoring rule shown in Equation (4). The resulting original, nonnormalized peer-prediction scores can be seen in Table 4, where each numeric cell can be interpreted as the score received by the agent in the row based on its evaluation of the agent in the column.

In the following step, the mechanism computes the peer-prediction scores by normalizing the original scores (Equation (10)). The peer-prediction score received by agent C, μ_C^D, based on its evaluation of agent D, s_C^D, is equal to its original score minus the mean of the original scores that are based on evaluations given to agent D and whose values cannot be affected by s_C^D. According to Table 4, the agent right before agent C in the random ordering indexed by D is agent B. Consequently, s_C^D can affect agent B's original peer-prediction score because it is used as a reference when computing that score. In this way, agent C's normalized score is

$$\mu_C^D = R(\boldsymbol{\Phi}_C^D, s_A^D) - \frac{1}{n-3} \sum_{y \neq C, D, B} R(\boldsymbol{\Phi}_y^D, s_{y_{+1}}^D).$$
$$\approx 0.182$$

The resulting normalized peer-prediction scores are also shown in Table 4. Finally, the mechanism computes agents' shares. For doing this, it aggregates the scaled evaluations received by each agent (Equation (7)), e.g., $\bar{\zeta}^A = (29.63 + 38.89 + 38.89)/4 \approx 26.86$. Thereafter, the mechanism computes agents' truth-telling scores as the average of their peer-prediction scores, e.g., $\tau_A = (0 - 0.182 + 0)/3 \approx -0.061$. Finally, the share of each agent is a linear combination of the aggregation of the scaled evaluations received by that agent and its truth-telling score, e.g., $\Gamma_A = \bar{\zeta}^A + \alpha\tau_A \approx 23.81$. The resulting shares and their major components can be seen in Table 5.

Table 5 Resulting Shares

Agent	$\bar{\zeta}^i$	τ_i	Γ_i
A	26.86	−0.061	23.81
B	11.27	−0.061	8.22
C	36.38	0.061	39.43
D	25.49	0.061	28.54

4.4. Properties

The proposed mechanism satisfies important properties. We start by stating our main result.

Theorem 1. *Under the assumptions of self-interest, a common prior, rationality, and independent signals, truth telling strictly maximizes agents' expected shares.*

Table 4 Original and Normalized Peer-Prediction Scores

	Original peer-prediction scores				Normalized peer-prediction scores			
Agent	*A*	*B*	*C*	*D*	*A*	*B*	*C*	*D*
A	—	0.074	0.074	0.074	—	0.000	−0.182	0.000
B	0.074	—	0.256	0.074	−0.182	—	0.182	−0.182
C	0.074	0.074	—	0.256	0.000	0.000	—	0.182
D	0.256	0.074	0.074	—	0.182	0.000	0.000	—

Proof. Consider the share received by agent i:

$$\Gamma_i = \bar{\zeta}^i + \alpha \frac{\sum_{j\neq i} \mu_i^j}{n-1} = \bar{\zeta}^i + \alpha' \sum_{j\neq i} \mu_i^j,$$

for $\alpha' > 0$. Because of the assumption of independent signals, we can restrict ourselves to analyze a single evaluation reported by agent i, say, its evaluation of agent j. Consequently, we can restrict our analysis to the expression $\bar{\zeta}^i + \alpha' \mu_i^j$. Now, we note that μ_i^j can be seen as a positive affine transformation of a strictly proper scoring rule (see the proof of Lemma 2). Furthermore, we observe that agent i's reported evaluations cannot affect $\bar{\zeta}^i$ (see Equation (7)). Consequently, $\bar{\zeta}^i + \alpha' \mu_i^j$ can be seen as another positive affine transformation of a strictly proper scoring rule. We show in Lemma 2 that if we take the expectation with respect to $\mathbf{\Theta}_i^j$ (agent i's posterior predictive distribution), then agent i strictly maximizes $E[\mu_i^j]$ when $s_i^j = t_i^j$, i.e., when it is telling the truth. Because $\bar{\zeta}^i + \alpha' \mu_i^j$ is a positive affine transformation of a strictly proper scoring rule, Lemma 1 implies that its expected value is also strictly maximized when $s_i^j = t_i^j$. □

A straightforward corollary of Theorem 1 is that *the collective truth-telling strategy profile is a strict Bayes–Nash equilibrium*. In other words, the proposed mechanism is incentive compatible. The theorem itself is stronger because, under the assumptions of the proposed model, it says that each agent strictly maximizes its expected share by following the truth-telling strategy, and the evaluations reported by others do not matter. However, it is not as strong as ex post equilibrium, which means that no agent would ever want to deviate from the collective truth-telling strategy profile even if it knew the truthful evaluations of its peers.

It is also interesting to note that by setting $\alpha = 0$, the collective truth-telling strategy profile becomes a weak dominant-strategy equilibrium, i.e., each agent weakly maximizes its share by following the truth-telling strategy. In this case, given a fixed vector $\mathbf{s}_{-i}$, agent i always receives the same share, no matter what it reports. In other words, although agents do not have direct incentives for misreporting evaluations, they also do not have incentives for telling the truth. Consequently, several problems may arise, e.g., collusions, biased evaluations, etc. That is why we require α to be strictly greater than zero. By doing this, rational agents are encouraged to disclose their truthful evaluations. Interestingly, Theorem 1 holds for any $\alpha > 0$, i.e., the weight of the truth-telling scores on agents' shares does not compromise the incentive-compatibility property of the mechanism.

Proposition 4. *The proposed mechanism is budget balanced.*

Proof. See the appendix.

Because agents' truth-telling scores can be negative (see Proposition 3), their shares can also be negative. In the following proposition, we present a scheme for avoiding negative shares. Intuitively, it means that the resulting shares can always be positive, regardless of the reported evaluations, if we reduce the influence of the truth-telling scores on agents' shares.

Proposition 5. *If* $\alpha \leq V/(2n)$, *then the proposed mechanism is individually rational.*

Proof. See the appendix.

Related to the fairness criterion (Definition 3), we note that there are two major issues with our mechanism. First, by using truth-telling scores, we employ part of the joint reward to promote truthfulness. Consequently, an agent that unanimously receives better evaluations than a peer will not necessarily receive a greater share of the reward than its peer because that agent's truth-telling score can be lower. Intuitively, we need to reduce the weight of the truth-telling scores on agents' shares for making the proposed mechanism fair, so that those shares will depend almost entirely on the reported evaluations.

The second issue is about scaling evaluations. Because we take into account the evaluations given by two agents to each other in our fairness criterion, then an agent that universally receives better evaluations than a peer will not necessarily have *all* scaled evaluations greater than its peer's scaled evaluations. In other words, given two agents i and j, where $s_z^i > s_z^j$, for all $z \neq i, j$, we can see from Equation (6) that the scaled evaluation ζ_z^i is greater than ζ_z^j. However, $s_j^i > s_i^j$ does not imply that $\zeta_j^i > \zeta_i^j$. Thus, an agent that unanimously receives better evaluations than a peer will not necessarily receive a greater share

Table 6 A Scenario Where the Proposed Mechanism Fails to Be Fair

	Reported evaluations				Scaled evaluations						
Agent	*A*	*B*	*C*	*D*	*A*	*B*	*C*	*D*	$\bar{\zeta}^i$	τ_i	Γ_i
A	—	3	1	1	—	60.00	20.00	20.00	12.24	−0.06	9.24
B	4	—	9	9	18.18	—	40.91	40.91	18.85	−0.06	15.84
C	2	1	—	10	15.38	7.69	—	76.92	34.46	0.06	37.46
D	2	1	10	—	15.38	7.69	76.92	—	34.46	0.06	37.46

Notes. Agent *B* receives a greater share of the joint reward than agent *A*, even though agent *A* unanimously receives better evaluations than agent *B*. The parameters $V = 100$, $M = 10$, and $\alpha = 50$ are used as well as the random permutations in Table 3.

of the joint reward than its peer because the aggregation of the scaled evaluations received by that agent can be lower.

To illustrate this last point, consider the evaluations shown in Table 6 and the parameters $V = 100$, $M = 10$, and $\alpha = 50$. As can be seen, $\forall z \neq A, B$, $s_z^A > s_z^B$. Furthermore, $s_B^A > s_A^B$. Thus, agent *A* unanimously receives better evaluations than agent *B*. Consequently, for the mechanism to be fair, the share received by agent *A* must be greater than the share received by agent *B*, i.e., $\Gamma_A > \Gamma_B$. However, if we use the random orderings in Table 3, Γ_B turns out to be greater than Γ_A. Because agent *A* and *B*'s truth-telling scores are the same ($\tau_A = \tau_B = -0.06$), the major reason for the difference in their shares is that the scaled evaluation $\zeta_A^B = 60.00$ is greater than the scaled evaluation $\zeta_B^A = 18.18$, even though the reported evaluation $s_A^B = 3$ is less than $s_B^A = 4$. In what follows, we present a sufficient condition under which the proposed mechanism is fair.

Proposition 6. *If* $M^2 + 2 \leq n \leq \sqrt{V/(4\alpha)}$, *then the mechanism is fair.*

Proof. See the appendix.

Next, we propose a scheme to avoid collusions of size 2. Intuitively, it means that if truth-telling scores have a high weight on agents' shares, then those agents will not have strong incentives to collude.

Proposition 7. *If* $\alpha \geq V(M+1)^3/(2n)$, *then the proposed mechanism is* (2)*-collusion-resistant*.

Proof. See the appendix.

Because the proposed mechanism is budget balanced, it is also ex-post (n)-collusion resistant (Proposition 1). From Propositions 5–7, we can see that there is a major trade-off between individual rationality/fairness and collusion resistance. Actually, it is impossible to set α in such a way that the proposed mechanism always satisfies all properties (we skip the proof because it can be easily deduced from the proofs of the aforementioned propositions). In other words, for any value of α, we can always come up with a scenario (reported evaluations, reward, etc.) in which the proposed mechanism does not simultaneously satisfy all properties.

Finally, we note that the proofs of the aforementioned propositions are based on worst-case scenarios, which we believe are fairly unlikely to happen in practice. In other words, they provide sufficient, but not necessary, conditions. We argue that it is possible to use higher (respectively, lower) values for α than the ones suggested in Propositions 5 and 6 (respectively, Proposition 7) and still be able to obtain individual rationality and fairness (respectively, collusion resistance) in practice, as we empirically show in the following section.

5. Simulation Results

The parameter α of the proposed mechanism fine-tunes the weight given to the truth-telling scores. To better understand its influence on agents' shares, we performed the following experiment. We shared the reward $V = 100$ among 100 agents using the proposed mechanism and $M = 10$ as the top possible evaluation that an agent can give or receive. A large population was intentionally used to decrease the average share. We chose values for α from the set {1, 10, 25, 50, 75, 100, 250, 500}.

We ran this experiment 100 independent times. At each simulation step, we randomly generated permutations of agents and truthful evaluations. In detail, for each agent $i \in N$, we drew $n-1$ samples from a normal distribution, and rounded each of them to the nearest integer. These rounded samples were the signals (truthful evaluations) observed by agent i's peers regarding its performance in accomplishing the joint task. The mean of each normal distribution was uniformly selected from the set $\{1, \ldots, M\}$. The variance of each of them was set to 1. Each rounded sample less than 1 or greater than M was discarded and a new one was drawn.

After generating evaluations, we computed the resulting shares when all agents were telling the truth and when there was a $(2, M)$-collusion between two fixed agents, i.e., when they always gave the top possible evaluation to each other. Thereafter, we calculated the *loss by lying*, i.e., the difference between the sum of the shares received by the fixed agents when they were telling the truth and when they were colluding. To show the statistical significance of the obtained results, we performed the directional t-test. Our null hypothesis was that the average joint share when the fixed agents were telling the truth was equal to the average joint share when they were colluding. Our alternative hypothesis was that the former was greater than the latter, i.e., the average loss by lying was greater than zero. Because each noncolluder was always telling the truth, then the joint share when the fixed agents were telling the truth and the joint share when they were colluding were correlated. Because of that, we used the t-test for correlated samples (also known as paired t-test). This test allowed us to remove irrelevant, extraneous information from the analyzed shares.

Table 7 presents the results of the experiment. We start by noting that if $\alpha \geq 666$, then we can guarantee that the proposed mechanism is (2)-collusion resistant (Proposition 7). The experimental results show that even using much lower values for α (e.g., $\alpha = 25$), the average joint share can be greater when the fixed agents are telling the truth than when they are colluding. When the value of α increases, the mechanism becomes stronger against $(2, M)$-collusions, i.e., the average loss by lying increases. The standard deviation also increases because the range of the shares becomes larger.

In our experiments, we also calculated the average number of negative and unfair shares returned by the mechanism when all agents were telling the truth. An agent's share is considered unfair if that agent unanimously receives better evaluations than a peer, but its share is smaller than the peer's share. Thus, a mechanism is fair if it does not return any unfair shares (see Definition 3). To compute the number of unfair shares, we made pairwise comparisons at each simulation step where each returned share was compared to the others for determining whether or not it was unfair.

Table 7 **Loss by Lying**

α	Avg.	Std. dev.	p-value
1	−0.0161	0.0041	1
10	−0.0100	0.0110	1
25	0.0002	0.0316	0.4748
50	0.0172	0.0525	0.0039
75	0.0412	0.0952	0.0011
100	0.0512	0.1050	0.0001
250	0.1531	0.2626	≈0
500	0.3230	0.5254	≈0

Table 8 presents the experimental results related to fairness and individual rationality. We first note that we need to set $\alpha \leq 0.0025$ to theoretically ensure that the proposed mechanism is always fair (Proposition 6). The experimental results show that even with much higher values for α (e.g., $\alpha = 10$), the mechanism might still be fair. Similar results occur with the individual-rationality property. Proposition 5 says that we must set $\alpha \leq 0.5$ to guarantee that the shares returned by the mechanism are always greater than or equal to zero. From Table 8, we can see that even with much higher values for α (e.g., $\alpha = 10$), the shares returned by the mechanism might be nonnegative. As expected, the average number of unfair and negative shares increase as the value of α increases.

Summarizing, this experiment helps to illustrate two important points. First, it shows the trade-off between fairness/individual rationality and collusion resistance. We show that when α increases, the mechanism becomes stronger against $(2, M)$-collusions, but at the expense of increasing the number of unfair and negative shares. Intuitively, this happens because the mechanism is putting more weight on the truth-telling scores. Second, because Propositions 5–7 are based on worst-case scenarios, this experiment highlights that we can set values for α in a different way

Table 8 Average Number of Unfair and Negative Shares

α	Unfair shares	Negative shares
1	0	0
10	0	0
25	2.12	1.02
50	45.90	5.96
75	117.82	10.83
100	168.79	15.56
250	311.35	32.12
500	366.42	40.74

than suggested by those propositions and still be able to obtain the underlying properties. These points are rather general. They also appear in different scenarios (i.e., different values for V, n, and M) as well as when using some other distributions of truthful evaluations (e.g., U-shaped distributions).

6. Discussion

We discuss in this section some practical considerations regarding our mechanism and model as well as an extension of them to scenarios where agents observe their peers' performance several times before evaluating them.

6.1. Deployment

The intuition behind the proposed mechanism is pretty clear: Each agent has its expected share maximized to the extent that it is well evaluated by its peers and that it is truthfully reporting its evaluations. When it comes to the mathematics behind the proposed mechanism, the first component of agents' shares (Equation (7)) is also very intuitive: It is a normalized value proportional to the received evaluations. Therefore, agents can be informed that the higher their received evaluations are, the greater this first component will be.

However, it might be difficult to explain to the agents how the peer-prediction scores are calculated. But if we look closely, Equation (9) works by making pairwise comparisons between two reported evaluations. Therefore, instead of teaching agents the complex mathematics behind the peer-prediction scores, one can simply show a flowchart that summarizes how the method works, as in Figure 2, and honestly suggest that truthful reporting is the best strategy for three reasons: (1) the reported evaluations will be rewarded based on whether or not they are in agreement with other reported evaluations; (2) each agent does not know against whom its reported evaluations will be compared to; and (3) even if it knew the identity of the other agent, the former would not necessarily know about the latter's reported evaluations.

6.2. Multiple Signals

We consider an extension of our model where agents may observe their peers' performance several times before evaluating them. Let $\rho_i^j \in \mathbb{Z}^+$ be the number of signals observed by agent i from agent j. Instead of a single number, t_i^j is now a vector, i.e., $\mathbf{t}_i^j = (t_i^j[1], \ldots, t_i^j[\rho_i^j])$, where $t_i^j[k] \in \{1, \ldots, M\}$, for $1 \leq k \leq \rho_i^j$. The basic assumptions (self-interest, common-prior, rationality, and independent signals) are still the same. As a way to abstract agents' private information, the mechanism only elicits the most common signals observed by each agent. In detail, let $H(x, k)$ be an indicator function, i.e.,

$$H(x, k) = \begin{cases} 1 & \text{if } x = k, \\ 0 & \text{otherwise.} \end{cases}$$

We say that agent i is truthfully reporting its evaluation of agent j when

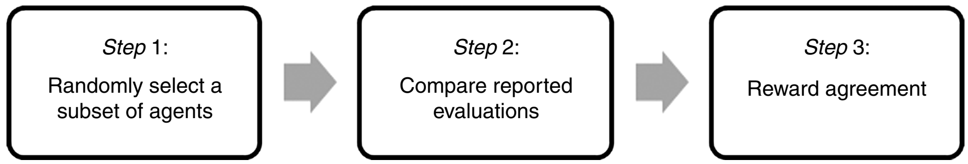
$$s_i^j = \underset{x \in \{1, \ldots, M\}}{\arg\max} \sum_{k=1}^{\rho_i^j} H(x, t_i^j[k]).$$

In words, agent i reports the most common signal coming from agent j. Ties are broken randomly. In this new model, agent i's posterior predictive distribution regarding $\boldsymbol{\omega}_j$, $\boldsymbol{\Theta}_i^j$, is defined as

$$\left(\frac{\sigma_1 + \sum_{k=1}^{\rho_i^j} H(1, t_i^j[k])}{\sum_{k=1}^{M} \sigma_k + \rho_i^j}, \frac{\sigma_2 + \sum_{k=1}^{\rho_i^j} H(2, t_i^j[k])}{\sum_{k=1}^{M} \sigma_k + \rho_i^j}, \ldots, \frac{\sigma_M + \sum_{k=1}^{\rho_i^j} H(M, t_i^j[k])}{\sum_{k=1}^{M} \sigma_k + \rho_i^j} \right). \quad (11)$$

Once again, we assume that $\sigma_k = 1$, for $k \in \{1, \ldots, M\}$ (noninformative priors). For illustration's sake, consider that agent i observes five signals from agent j ($\rho_i^j = 5$) and that $\mathbf{t}_i^j = (1, 5, 4, 5, 5)$ for $M = 5$. Consequently, agent i's posterior predictive distribution regarding $\boldsymbol{\omega}_j$ is $\boldsymbol{\Theta}_i^j = (2/10\ 1/10,\ 1/10,\ 2/10,\ 4/10)$, and agent i's truthful evaluation of

Figure 2 **Flowchart of How the Original Truth-Telling Scoring Method Works**



agent j is 5 because this is the most common signal from agent j.

Interestingly, the proposed mechanism still satisfies important properties under this new model, even without knowing a priori the number of observed signals. First, because Propositions 4–6 do not depend on agents' private information, the proposed mechanism is still budget balanced, and the sufficient conditions under which it is fair and individually rational are still the same. To show that each agent strictly maximizes its expected share by telling the truth, we can restrict ourselves to show that truth-telling strictly maximizes the expected peer-prediction scores. By showing this, the proof of Theorem 1 can be used to show that expected shares are also strictly maximized.

Proposition 8. *When observing multiple signals and under the assumptions of self-interest, a common prior, and rationality, each agent $i \neq j$ strictly maximizes its expected peer-prediction score, $E[\mu_i^j]$, when*

$$s_i^j = \underset{x \in \{1, \dots, M\}}{\arg\max} \sum_{k=1}^{\rho_i^j} H(x, t_i^j[k]).$$

Proof. See the appendix.

Finally, we note that Proposition 7 does not hold anymore. Collusion resistance depends on the expected scoring loss when an agent is lying instead of telling the truth (see Equation (16) in the proof of Proposition 7). From Proposition 8, we can deduce that the expected scoring loss is always nonnegative. However, it is easy to show that the loss can be arbitrarily small as the number of observed signals increases. Therefore, an arbitrarily large part of the joint reward must be used as truth-telling scores to avoid collusions of size 2.

6.3. Random Orderings

The computation of the peer-prediction scores is highly dependent on the underlying ordering of the agents. For example, even though an agent's reported evaluation is in agreement with the majority of other reported evaluations, that agent can receive a low peer-prediction score because of a single "unlucky" pairwise comparison. This fact can generate a sense of unfairness because agents can lose a decent payoff due to an unfavorable ordering. A simple way to circumvent the above problem is by using more than one agent as reference when scoring a reported evaluation. For example, the peer-prediction method can independently use the evaluations given by agents i_{+1} and i_{+2} to agent j as the outcomes of an uncertain event (the evaluation deserved by agent j) and reward agent i's estimated posterior predictive distribution as follows:

$$\hat{\mu}_i^j = R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j) + R(\mathbf{\Phi}_i^j, s_{i_{+2}}^j).$$

Then, the normalized peer-prediction score (Equation (10)) becomes

$$\mu_i^j = \hat{\mu}_i^j - \frac{1}{n-4} \sum_{y \neq i,\, j,\, i_{-1},\, i_{-2}} \hat{\mu}_y^j$$

for $n \geq 5$. The above change has no impact on Theorem 1 and Proposition 4, i.e., the proposed mechanism is still incentive compatible and budget balanced. The mechanism can use the reports from up to $k < n/2 + 1$ agents as reference and those properties are still valid. However, this approach has impact on the range of the peer-prediction scores (Proposition 3). In the above example, the new range is $[-(4/(M+1)), 4/(M+1)]$, i.e., two times the old range. Because the scoring range has changed, the sufficient conditions under which the proposed mechanism is individually rational, fair, and resistant to collusions of size 2 also change. The new conditions can be easily obtained by updating both the lower bound and the upper bound of the peer-prediction scores in the proofs of Propositions 5–7.

6.4. Scoring Rules Sensitive to Distance

Scoring rules, as defined in §3.4, are used to both provide an ex ante incentive for truthful reporting and to measure how informative the probabilities look ex post, i.e., after the true outcome is observed. However, the canonical definition of scoring rules does not take into account any ordering of the underlying outcomes. In our setting, this implies that any two different reported evaluations will receive the same score from Equation (9) whenever they are not similar to the evaluation used as reference.

That happens because, as noted before, the proposed peer-prediction method essentially works by making pairwise comparisons, i.e., an agent's score is high when its reported evaluations are equal to the evaluations used as reference. This approach can be too restrictive and, to some degree, unfair when the top possible evaluation is high. For example, when $M = 10$ and the evaluation used as reference is also equal to 10, a reported evaluation equal to 9 seems to be more accurate than a reported evaluation equal to 1. One effective way to deal with this issue is by using a strictly proper scoring rule in Equation (9) that is *sensitive to distance*.

Using the notation from §3.4, recall that given M possible outcomes, $\mathbf{q} = (q_1, \ldots, q_M)$ is some reported probability distribution. Given that the outcomes are ordered, we denote the cumulative probabilities by a capital letter: $\mathbf{Q}_i = \sum_{j \le i} q_j$. We first define the notion of distance between two probability vectors. We say that a probability vector $\mathbf{q}'$ is more distant from the jth outcome than a probability vector $\mathbf{q}$, for $\mathbf{q} \neq \mathbf{q}'$, if

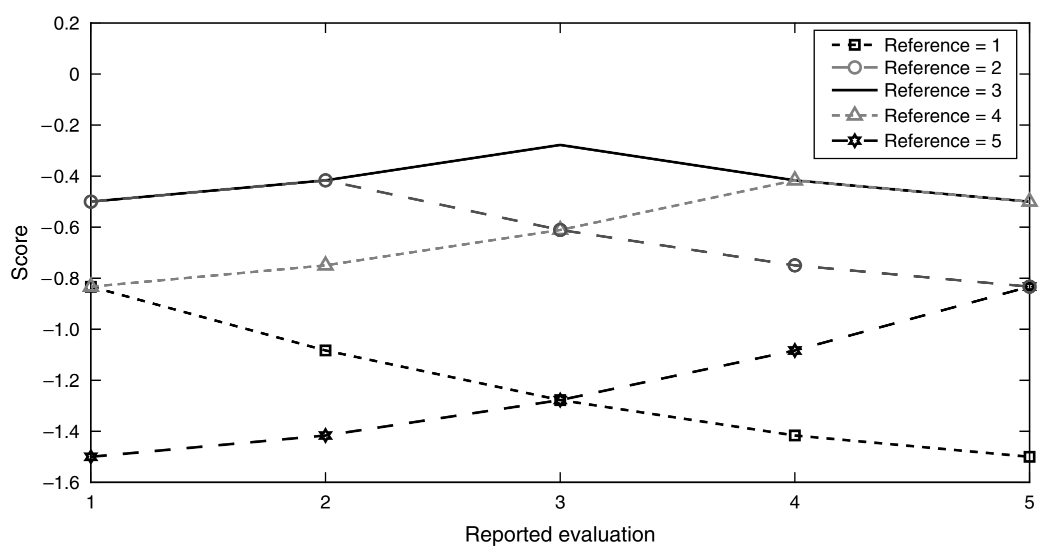
$$\mathbf{Q}'_i \ge \mathbf{Q}_i, \quad \text{for } i = 1, \ldots, j-1;$$
$$\mathbf{Q}'_i \le \mathbf{Q}_i, \quad \text{for } i = j, \ldots, M-1$$

(Staël von Holstein 1970). Intuitively, the above definition means that $\mathbf{q}$ can be obtained from $\mathbf{q}'$ by successively moving probability mass towards the jth outcome from other outcomes more distant from the jth outcome. A scoring rule R is said to be *sensitive to distance* if $R(\mathbf{q}, j) > R(\mathbf{q}', j)$ whenever $\mathbf{q}'$ is more distance from $\mathbf{q}$ for all j. Epstein (1969) introduces the ranked probability score (RPS), a strictly proper scoring rule that is sensitive to distance. Using Murphy's (1970) formulation, we have

$$\text{RPS}(\mathbf{q}, j) = -\sum_{i=1}^{j-1} \mathbf{Q}_i^2 - \sum_{i=j}^{M-1} (1 - \mathbf{Q}_i)^2.$$

Figure 3 illustrates the scores returned by Equation (9) for different reported evaluations and reference evaluations (outcomes) when $M = 5$ and R is equal to the above equation. When using the RPS as the scoring rule, agents are rewarded based on how close their reported evaluations are to the evaluations used as reference: The closer one reported evaluation is, the higher its score will be. For example, when the

Figure 3 **Scores Returned by Equation** (9) **When** $M = 5$ **and the RPS Is the Strictly Proper Scoring Rule**



Note. Each line represents a different evaluation used as reference (outcome).

evaluation used as reference is equal to 1 (see the dotted line with squares in Figure 3), the returned score monotonically decreases as the reported evaluation increases. Because RPS is strictly proper, Theorem 1 is still valid, i.e., the proposed mechanism is still incentive compatible. It can be easily seen that the RPS scoring range is $[-M+1, 0]$. Because it is bounded, the proposed mechanism is still budged balanced (Proposition 4) when using RPS. Furthermore, the new sufficient conditions under which the proposed mechanism is individually rational, fair, and resistant to collusions of size 2 can be easily obtained by updating the lower bound and the upper bound of the peer-prediction scores in the proofs of Propositions 5–7.

7. Conclusion

We proposed a game-theoretic model for sharing a joint, homogeneous reward among an unfamiliar group of agents based on peer evaluations. We introduced a mechanism to elicit and aggregate evaluations as well as for determining agents' shares. The intuition behind the proposed mechanism is that each agent has its expected share maximized to the extent that it is well evaluated by its peers and that it is truthfully reporting its evaluations. For promoting truthfulness, we proposed a peer-prediction method built on strictly proper scoring rules. Under the main assumptions that agents are Bayesian decision makers and that they do not have informative prior knowledge about the competence of their teammates, we showed that our mechanism is incentive compatible and budget balanced. We also provided sufficient conditions under which the mechanism is individually rational, resistant to collusions of size 2, and fair.

An interesting point regarding our mechanism is the trade-off between fairness/individual rationality and collusion resistance; that is, if we overly promote truthfulness, then the mechanism becomes stronger against collusions, but at the expense of increasing the number of unfair and negative shares. A major reason for this fact is that avoiding collusions is expensive, i.e., a large part of the joint reward must be used to promote truthfulness. Thus, an exciting direction for extending this work is to study new schemes that guarantee collusion resistance, but using less of the joint reward. Another interesting line of research is to investigate other kinds of collusive behavior that might arise and how to prevent them.

Acknowledgments

The authors are grateful to Pascal Poupart for useful discussions about Dirichlet distributions. They are also grateful to the editor-in-chief, associate editor, and the referee for their many constructive comments.

Appendix. Proofs

Proof of Proposition 3

By construction, the estimated posterior predictive distribution $\mathbf{\Phi}_i^j$ contains one element equal to $2/(M+1)$ and $M-1$ elements equal to $1/(M+1)$. Because the original, nonnormalized score (Equation (9)) is the result of the quadratic scoring rule (Equation (4)) applied to $\mathbf{\Phi}_i^j$, we have that

$$\frac{2}{M+1} - \frac{M+3}{(M+1)^2} \leq R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j) \leq \frac{4}{M+1} - \frac{M+3}{(M+1)^2}$$
$$\Leftrightarrow \frac{M-1}{(M+1)^2} \leq R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j) \leq \frac{3M+1}{(M+1)^2}.$$

Recall that

$$\mu_i^j = R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j) - \frac{1}{n-3} \sum_{y \neq i, j, i_{-1}} R(\mathbf{\Phi}_y^j, s_{y_{+1}}^j).$$

Consequently,

$$\frac{M-1}{(M+1)^2} - \frac{3M+1}{(M+1)^2} \leq \mu_i^j \leq \frac{3M+1}{(M+1)^2} - \frac{M-1}{(M+1)^2}$$
$$\Leftrightarrow -\frac{2}{M+1} \leq \mu_i^j \leq \frac{2}{M+1}. \quad \square$$

Proof of Proposition 4

The sum of the shares received by the agents is equal to

$$\begin{aligned}\sum_{i=1}^{n}(\bar{\zeta}^i + \alpha\tau_i) &= \sum_{i=1}^{n}\bar{\zeta}^i + \alpha\sum_{i=1}^{n}\tau_i \\ &= \sum_{i=1}^{n}\frac{\sum_{j\neq i}\zeta_j^i}{n} + \alpha\sum_{i=1}^{n}\frac{\sum_{j\neq i}\mu_i^j}{n-1} \\ &= \sum_{j=1}^{n}\frac{\sum_{i\neq j}\zeta_j^i}{n} + \alpha\sum_{j=1}^{n}\frac{\sum_{i\neq j}\mu_i^j}{n-1} \\ &= n\frac{V}{n} + \frac{\alpha}{n-1}\left(\sum_{j=1}^{n}\sum_{i\neq j}\mu_i^j\right).\end{aligned}$$

The last equality follows from the fact that the scaled evaluations sum up to V (Equation (6)). From Proposition 2, we know that $\sum_{i\neq j}\mu_i^j = 0$, thus completing the proof. $\square$

Proof of Proposition 5

We start the proof by observing that $\forall\, i \in N,\ \bar{\zeta}^i > 0$ (Equation (7)). Consequently, if agents' truth-telling scores are positive, then their shares will also be positive (recall that $\Gamma_i = \bar{\zeta}^i + \alpha\tau_i$). In this way, we restrict ourselves to the scenario where the truth-telling scores are negative. In detail, for the proposed mechanism to be individually rational, the following inequality must be true for every agent $i \in N$ and $\tau_i < 0$:

$$\bar{\zeta}^i + \alpha\tau_i \geq 0 \equiv \frac{\bar{\zeta}^i}{-\tau_i} \geq \alpha. \tag{12}$$

In what follows, we compute a lower-bound for the fraction in (12). Starting with the numerator, we have

$$\begin{aligned}\bar{\zeta}^i &= \frac{\sum_{j\neq i} s_j^i (V/(\sum_{q\neq j} s_j^q))}{n} \\ &\geq \frac{\sum_{j\neq i} s_j^i (V/M(n-1))}{n} \\ &\geq \frac{V(n-1)}{Mn(n-1)}.\end{aligned}$$

The inequalities follow from the fact that $\forall\ i, j,\ s_i^j \in \{1, \ldots, M\}$. Focusing on the denominator of the fraction in (12), because τ_i is the average of $n-1$ peer-prediction scores, we can restrict ourselves to find the lowest negative score that can be returned by our peer-prediction method. From Proposition 3, we know that this value is equal to $-2/(M+1)$, which is greater than $-2/M$. Thus, we conclude that if

$$\alpha \leq \frac{V/(Mn)}{-(-2/M)} = \frac{V}{2n},$$

then the proposed mechanism is individually rational. □

Proof of Proposition 6

Consider a pair of agents $i, j \in N$, and any strategy profile $\mathbf{s} \in S$ where $s_j^i > s_i^j$ and, for every other agent $z \neq i, j,\ s_z^i > s_z^j$. For the mechanism to be considered fair (Definition 3), the shares $\Gamma_i(\mathbf{s})$ and $\Gamma_j(\mathbf{s})$ must satisfy the following inequality:

$$\Gamma_i(\mathbf{s}) > \Gamma_j(\mathbf{s}) \Leftrightarrow \bar{\zeta}^i + \alpha\tau_i > \bar{\zeta}^j + \alpha\tau_j \Leftrightarrow \alpha < \frac{\bar{\zeta}^i - \bar{\zeta}^j}{\tau_j - \tau_i}. \tag{13}$$

In what follows, we compute a lower bound for the above fraction. Starting with the numerator, we have

$$\begin{aligned}&\bar{\zeta}^i - \bar{\zeta}^j \\ &\quad= \frac{\sum_{z\neq i,j}(s_z^i - s_z^j)(V/\sum_{q\neq z} s_z^q) + s_j^i(V/\sum_{q\neq j} s_j^q) - s_i^j(V/\sum_{q\neq i} s_i^q)}{n} \\ &\quad\geq \frac{V}{n}\left(\frac{n-2}{(n-1)M} + \frac{1}{(n-1)M} - \frac{M}{(n-1)}\right) \\ &\quad= \frac{V}{n}\left(\frac{n-2+1-M^2}{(n-1)M}\right) \\ &\quad\geq \frac{V}{n(n-1)M} \\ &\quad\geq \frac{V}{n^2 M}.\end{aligned}$$

The first inequality follows from the facts that for every agent $z \neq i, j,\ s_z^i > s_z^j$ and that $\forall\, i, j,\ s_i^j \in \{1, \ldots, M\}$. The second inequality follows from the assumption that $M^2 + 2 \leq n \equiv M \leq \sqrt{n-2}$. Focusing on the denominator of the fraction in (13), because $\forall\, q \in N,\ \tau_q$ is the average of $n-1$ results from our peer-prediction method, the difference between τ_j and τ_i is always less than or equal to the difference between the highest and the lowest peer-prediction score. According to Proposition 3, this difference is equal to $4/(M+1)$, which is less than $4/M$. Thus, we conclude that if

$$\alpha \leq \frac{V/(Mn^2)}{4/M} = \frac{V}{4n^2}$$

and $M \leq \sqrt{n-2}$, which is equivalent to say that $M^2 + 2 \leq n \leq \sqrt{V/(4\alpha)}$, then the proposed mechanism is fair. □

Proof of Proposition 7

Let $\mathring{\mathbf{s}}$ be the collective truth-telling strategy profile, where $\Gamma_i(\mathring{\mathbf{s}}) = \sum_{z\neq i} \mathring{\zeta}_z^i/n + \alpha(\sum_{z\neq i} \mathring{\mu}_i^z/(n-1))$ is agent i's share when it is truthfully reporting its evaluations. Furthermore, let $\tilde{\mathbf{s}}$ be the resulting strategy profile after a $(2, M)$-collusion between agent i and agent j, so that $\Gamma_i(\tilde{\mathbf{s}}) = \sum_{z\neq i} \tilde{\zeta}_z^i/n + \alpha(\sum_{z\neq i} \tilde{\mu}_i^z/(n-1))$. We show that for certain values of α, $E[\Gamma_i(\tilde{\mathbf{s}})] \leq E[\Gamma_i(\mathring{\mathbf{s}})]$, i.e., agent i has no incentive to collude. We start by noting that $\forall\, z \neq i, j,\ \mathring{\zeta}_z^i = \tilde{\zeta}_z^i$ and $\mathring{\mu}_i^z = \tilde{\mu}_i^z$ because only agents i and j are colluding, i.e., both the evaluations given by noncolluders to agent i and agent i's peer-prediction scores resulting from its evaluations of noncolluders do not change because of agent i's collusive behavior. Consequently, we can restrict ourselves to the inequality

$$E\left[\frac{\tilde{\zeta}_j^i}{n} + \alpha\frac{\tilde{\mu}_i^j}{n-1}\right] \leq E\left[\frac{\mathring{\zeta}_j^i}{n} + \alpha\frac{\mathring{\mu}_i^j}{n-1}\right],$$

which, after some algebraic manipulations, leads to

$$\alpha \geq \frac{E[\tilde{\zeta}_j^i - \mathring{\zeta}_j^i]}{E[\mathring{\mu}_i^j - \tilde{\mu}_i^j]}\ \frac{n-1}{n}. \tag{14}$$

In what follows, we compute an upper bound for the numerator of (14). From Equation (6) we have that

$$\begin{aligned}E[\tilde{\zeta}_j^i - \mathring{\zeta}_j^i] &= \tilde{s}_j^i\left(\frac{V}{\sum_{q\neq j} \tilde{s}_j^q}\right) - \mathring{s}_j^i\left(\frac{V}{\sum_{q\neq j} \mathring{s}_j^q}\right) \\ &\leq \frac{MV}{n-1} - \frac{V}{(n-1)M} \\ &= \frac{V(M^2-1)}{(n-1)M} \\ &= \frac{V(M-1)(M+1)}{(n-1)M} \\ &\leq \frac{V(M+1)}{(n-1)}.\end{aligned} \tag{15}$$

The first inequality follows from the fact that $\forall\, i, j,\ s_i^j \in \{1, \ldots, M\}$. Focusing on $E[\mathring{\mu}_i^j - \tilde{\mu}_i^j]$ in (14), from Equation (10) we can deduce that

$$E[\mathring{\mu}_i^j - \tilde{\mu}_i^j] = E[R(\mathbf{\Theta}_i^j, s_{i_{+1}}^j) - R(\mathbf{\Phi}_i^j, s_{i_{+1}}^j)],$$

where $\mathbf{\Theta}_i^j = (\theta_1, \ldots, \theta_M)$ is agent i's actual posterior predictive distribution, and $\mathbf{\Phi}_i^j = (\phi_1, \ldots, \phi_M)$ is agent i's estimated posterior predictive distribution when it is lying, i.e.,

$$\theta_k = \begin{cases} 2/(M+1) & \text{if } \mathring{s}_i^j = k, \\ 1/(M+1) & \text{otherwise,} \end{cases} \quad \text{and}$$

$$\phi_k = \begin{cases} 2/(M+1) & \text{if } \tilde{s}_i^j = k, \\ 1/(M+1) & \text{otherwise.} \end{cases}$$

In words, $E[\mathring{\mu}_i^j - \tilde{\mu}_i^j]$ is equal to the expected scoring loss (Equation (3)) when agent i enters into a collusion with agent j, instead of following the collective truth-telling strategy profile. By setting R equal to the quadratic scoring rule (Equation (4)) and taking the expectation with respect to agent i's actual posterior predictive distribution, we obtain the following equation after simple algebraic manipulations:

$$\begin{aligned} E[\mathring{\mu}_i^j - \tilde{\mu}_i^j] &= \sum_{k=1}^{M} \theta_k \left(2\theta_k - \sum_{x=1}^{M} \theta_x^2 - 2\phi_k + \sum_{x=1}^{M} \phi_x^2\right) \\ &= \sum_{k=1}^{M} \theta_k (2\theta_k - 2\phi_k) \\ &= \frac{1}{M+1}\left(\frac{2}{M+1} - \frac{4}{M+1}\right) \\ &\quad + \frac{2}{M+1}\left(\frac{4}{M+1} - \frac{2}{M+1}\right) \\ &= \frac{2}{(M+1)^2}. \end{aligned} \tag{16}$$

The second and third equalities follow from the fact that there are only two elements in $\mathbf{\Theta}_i^j$ and in $\mathbf{\Phi}_i^j$ that have different values (by construction of the mechanism). Combining (14), (15), and (16), we have that if

$$\alpha \geq \frac{V(M+1)}{n-1}\ \frac{(M+1)^2}{2}\ \frac{n-1}{n} = \frac{V(M+1)^3}{2n},$$

then the proposed mechanism is resistant to $(2, M)$-collusions. Although we focused on $(2, M)$-collusions in this proof, it is easy to see that the same result applies to $(2, y)$-collusions, for $y \in \{1, \ldots, M\}$, because (15) is an upper bound on the difference $E[\tilde{\zeta}_j^i - \mathring{\zeta}_j^i]$, and $E[\mathring{\mu}_i^j - \tilde{\mu}_i^j]$ is not affected by y. $\square$

Proof of Proposition 8

We show that, when observing multiple signals, agent i strictly maximizes $E[\mu_i^j]$ by telling the truth, i.e., when

$$s_i^j = \underset{x \in \{1, \ldots, M\}}{\arg\max} \sum_{k=1}^{\rho_i^j} H(x, t_i^j[k]).$$

Without loss of generality, assume that

$$\underset{x \in \{1, \ldots, M\}}{\arg\max} \sum_{k=1}^{\rho_i^j} H(x, t_i^j[k]) = 1,$$

i.e., "1" is the most common signal observed by agent i coming from agent j. Recall that $\mathbf{\Theta}_i^j = (\theta_1, \ldots, \theta_M)$ is agent i's actual posterior predictive distribution regarding $\boldsymbol{\omega}_j$ (Equation (11)). Consequently, $\theta_1 \geq \theta_k$, for $2 \leq k \leq M$. In this proof, let $\mathring{\mathbf{\Phi}}_i^j = (\mathring{\phi}_1, \ldots, \mathring{\phi}_M)$ be agent i's estimated posterior predictive distribution when it is telling the truth, i.e.,

$$\mathring{\phi}_k = \begin{cases} 2/(M+1) & \text{if } k = 1, \\ 1/(M+1) & \text{otherwise.} \end{cases}$$

For contradiction's sake, suppose that agent i maximizes $E[\mu_i^j]$ by misreporting its evaluation. Without loss of generality, assume that agent i reports $s_i^j = 2$. Because "2" is not the most common signal observed by agent i from agent j, then $\theta_1 > \theta_2$. Let $\tilde{\mathbf{\Phi}}_i^j = (\tilde{\phi}_1, \ldots, \tilde{\phi}_M)$ be agent i's estimated posterior predictive distribution when it is misreporting its evaluation, i.e.,

$$\tilde{\phi}_k = \begin{cases} 2/(M+1) & \text{if } k = 2, \\ 1/(M+1) & \text{otherwise.} \end{cases}$$

It is important to note that $\mathring{\phi}_k = \tilde{\phi}_k$ for $3 \leq k \leq M$. A consequence of our assumption that agent i maximizes $E[\mu_i^j]$ by misreporting its evaluation is that

$$\begin{aligned} &E\left[R(\tilde{\mathbf{\Phi}}_i^j, s_{i_{+1}}^j) - \frac{1}{n-3} \sum_{y \neq i, j, i_{-1}} R(\mathbf{\Phi}_y^j, s_{y_{+1}}^j)\right] \\ &\quad \geq E\left[R(\mathring{\mathbf{\Phi}}_i^j, s_{i_{+1}}^j) - \frac{1}{n-3} \sum_{y \neq i, j, i_{-1}} R(\mathbf{\Phi}_y^j, s_{y_{+1}}^j)\right] \\ &\quad \Leftrightarrow E[R(\tilde{\mathbf{\Phi}}_i^j, s_{i_{+1}}^j)] \geq E[R(\mathring{\mathbf{\Phi}}_i^j, s_{i_{+1}}^j)]. \end{aligned}$$

By setting R equal to the quadratic scoring rule (Equation (4)) and taking the expectation with respect to agent i's actual posterior predictive distribution, the above inequality becomes equivalent to

$$\begin{aligned} &\sum_{k=1}^{M} \theta_k \left(2\tilde{\phi}_k - \sum_{x=1}^{M} (\tilde{\phi}_x)^2\right) \geq \sum_{k=1}^{M} \theta_k \left(2\mathring{\phi}_k - \sum_{x=1}^{M} (\mathring{\phi}_x)^2\right) \\ &\quad \Leftrightarrow \sum_{k=1}^{M} 2\theta_k \tilde{\phi}_k \geq \sum_{k=1}^{M} 2\theta_k \mathring{\phi}_k \\ &\quad \Leftrightarrow \theta_1 \tilde{\phi}_1 + \theta_2 \tilde{\phi}_2 \geq \theta_1 \mathring{\phi}_1 + \theta_2 \mathring{\phi}_2 \\ &\quad \Leftrightarrow \theta_2 \geq \theta_1\ \frac{\mathring{\phi}_1 - \tilde{\phi}_1}{\tilde{\phi}_2 - \mathring{\phi}_2}. \end{aligned}$$

The second and third lines follow, respectively, from the facts that $\sum_{x=1}^{M}(\tilde{\phi}_x)^2 = \sum_{x=1}^{M}(\overset{*}{\phi}_x)^2$ and that $\tilde{\phi}_k = \overset{*}{\phi}_k$ for $3 \leq k \leq M$. Regarding the last line, we have by construction that $\overset{*}{\phi}_1 = 2/(M+1)$, $\tilde{\phi}_1 = 1/(M+1)$, $\overset{*}{\phi}_2 = 1/(M+1)$, $\tilde{\phi}_2 = 2/(M+1)$. Consequently, we obtain that $\theta_2 \geq \theta_1$. As noted before, because "2" is not the most common signal observed by agent i from j, then $\theta_1 > \theta_2$. Hence, we have a contradiction. So, $E[R(\tilde{\mathbf{\Phi}}_i^j, s_{i_{+1}}^j)] < E[R(\overset{*}{\mathbf{\Phi}}_i^j, s_{i_{+1}}^j)]$, i.e., agent i strictly maximizes $E[\mu_i^j]$ by telling the truth. □

References

Bernardo JM, Smith AFM (1994) *Bayesian Theory* (John Wiley & Sons, West Sussex, UK).

Brams SJ, Taylor AD (1996) *Fair Division: From Cake-Cutting to Dispute Resolution* (Cambridge University Press, Cambridge, UK).

Carvalho A, Larson K (2011) A truth serum for sharing rewards. *Proc. 10th Internat. Conf. Autonomous Agents and Multiagent Systems* (International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC), 635–642.

Chevaleyre Y, Dunne PE, Endriss U, Lang J, Lemaître M, Maudet N, Padget J, Phelps S, Rodríguez-Aguilar JA, Sousa P (2006) Issues in multiagent resource allocation. *Informatica* 30:3–31.

Conitzer V, Sandholm T (2006) Complexity of constructing solutions in the core based on synergies among coalitions. *Artificial Intelligence* 170(6):607–619.

de Clippel G, Moulin H, Tideman N (2008) Impartial division of a dollar. *J. Econom. Theory* 139(1):176–191.

Epstein ES (1969) A scoring system for probability forecasts of ranked categories. *J. Appl. Meteorology* 8(6):985–987.

Gneiting T, Raftery AE (2007) Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* 102(477):359–378.

Goel S, Reeves DM, Pennock DM (2009) Collective revelation: A mechanism for self-verified, weighted, and truthful predictions. *Proc. 10th ACM Conf. Electronic Commerce* (Association for Computing Machinery, New York), 265–274.

Golany B, Kress M, Penn M, Rothblum UG (2012) Network optimization models for resource allocation in developing military countermeasures. *Oper. Res.* 60(1):48–63.

Griffin S, Claxton K, Sculpher M (2008) Decision analysis for resource allocation in health care. *J. Health Services Res. Policy* 13(3):23–30.

Grosz BJ, Kraus S (1996) Collaborative plans for complex group action. *Artificial Intelligence* 86(2):269–357.

Halpern JY (2008) Beyond Nash equilibrium: Solution concepts for the 21st century. *Proc. Twenty-Seventh ACM Sympos. Principles of Distributed Comput.* (Association for Computing Machinery, New York), 1–10.

Jurca R, Faltings B (2009) Mechanisms for making crowds truthful. *J. Artificial Intelligence Res.* 34:209–253.

Kaufman DB, Felder RM, Fuller H (2000) Accounting for individual effort in cooperative learning teams. *J. Engrg. Ed.* 89(2):133–140.

Keller LR, Wang Y, Feng T (2010) Fairness and equity in societal decision analysis. Cochran JJ, Cox LA, Keskinocak P, Kharoufeh JP, Smith JC, eds. *Wiley Encyclopedia of Operations Research and Management Science* (John Wiley & Sons, New York).

Kilgour DM, Gerchak Y (2004) Elicitation of probabilities using competitive scoring rules. *Decision Anal.* 1(2):108–113.

Knoblauch V (2009) Three-agent peer evaluation. *Econom. Lett.* 105(3):312–314.

Mas-Colell A, Whinston MD, Green JR (1995) *Microeconomic Theory* (Oxford University Press, New York).

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

Moulin H (1999) Incremental cost sharing: Characterization by coalition strategy-proofness. *Soc. Choice Welfare* 16(2):279–320.

Moulin H (2004) *Fair Division and Collective Welfare* (MIT Press, Cambridge, MA).

Moulin H, Shenker S (2001) Strategyproof sharing of submodular costs: Budget balance versus efficiency. *Econom. Theory* 18(3):511–533.

Murphy AH (1970) A note on the ranked probability score. *J. Appl. Meteorology* 10(1):155–156.

Nash JF (1950) The bargaining problem. *Econometrica* 18(2):155–162.

Osborne MJ, Rubinstein A (1994) *A Course in Game Theory* (MIT Press, Cambridge, MA).

Pratt JW, Zeckhauser RJ (1990) The fair and efficient division of the Winsor family silver. *Management Sci.* 36(11):1293–1301.

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

Rahwan T, Ramchurn SD, Jennings NR, Giovannucci A (2009) An anytime algorithm for optimal coalition structure generation. *J. Artificial Intelligence Res.* 34(1):521–567.

Robertson J, Webb W (1998) *Cake-Cutting Algorithms: Be Fair If You Can* (A.K. Peters, Natick, MA).

Savage LJ (1971) Elicitation of personal probabilities and expectations. *J. Amer. Statist. Assoc.* 66(336):783–801.

Selten R (1998) Axiomatic characterization of the quadratic scoring rule. *Experiment. Econom.* 1(1):43–61.

Shapley LS (1953) A value for *N*-person games. Kuhn HW, Tucker AW, eds. *Contributions to the Theory of Games*, Vol. 2 (Princeton University Press, Princeton, NJ), 307–317.

Shoham Y, Leyton-Brown K (2009) *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations* (Cambridge University Press, Cambridge, UK).

Staël von Holstein C-AS (1970) A family of strictly proper scoring rules which are sensitive to distance. *J. Appl. Meteorology* 9(3):360–364.

Stone P, Kaminka G, Kraus S, Rosenschein JS (2010) Ad hoc autonomous agent teams: Collaboration without pre-coordination. *Proc. Twenty-Fourth Conf. Artificial Intelligence* (Association for the Advancement of Artificial Intelligence, Palo Alto, CA), 1504–1509.

Tideman TN, Plassmann F (2008) Paying the partners. *Public Choice* 136(1):19–37.

Witkowski J, Parkes DC (2011) Peer prediction with private beliefs. *Proc. Workshop Soc. Comput. User Generated Content, San Jose, CA*.