Simple bets to elicit private signals*

Aurelien Baillon[a] and Yan Xu[a]

[a]Erasmus School of Economics, Erasmus University Rotterdam

December 2019

Abstract

This paper introduces two simple betting mechanisms, Top-Flop and Threshold betting, to elicit unverifiable information from crowds. Both mechanisms offer agents bets on the scores of two items, one about which they have received a private signal and the other one about which they have not. For instance, in Top-Flop betting, agents bet on or against a movie they just watched having a higher score than another, random movie. Alternatively, in Threshold betting, agents bet which movie will exceed a threshold score. We characterize conditions for the chosen bet to reveal the agents' private signal (e.g. their truthful assessment of the movie). We further establish micro-economic foundations of the scores in a game setting, in which the scores underlying the bets are endogenously determined by the actions of other agents. In the game setting, we relax standard assumptions of the literature such as a common prior and homogeneous and risk neutral agents, and we still obtain that bet choices reveal private signals.

Keywords: Bets; Private signals; Bayesian game; Elicitation.

*This research was made possible by European Research Council Starting Grant 638408 BayesianMarkets. Email: baillon@ese.eur.nl, xu@ese.eur.nl.

1 Introduction

Suppose the manager of a customer-care call center wants to assess her employees through some customer satisfaction measures. At the end of each call, she invites customers to take a one-question survey about whether or not they are satisfied with the services. She can reward participation with a small prize (voucher or fidelity points) but this is not enough. She would also like to have the customers think carefully about the question and provide truthful answers. If she were able to verify the answer, incentivizing truth-telling would be easy. However, only the customers themselves know whether they are actually satisfied or not, making it difficult to align rewards with truth-telling. We propose the following solution. The manager can reformulate the survey question and ask customers to bet whether the employee they talked to has a higher or lower satisfaction rate than another, randomly selected, employee from the call center. Customers who win the bet receive the prize.

We call the aforementioned method Top-Flop betting and show that it provides incentives for agents to truthfully reveal private information. We consider two cases. In the first case, the bets are defined on a pre-existing satisfaction rating, which may be biased as long as it is informative enough (as specified later). In the second case, the rating is a function of the bets chosen by other customers. Another method introduced in this paper and that we call Threshold betting, induces truth-telling by making customers bet on which employee (the one they talked to or a random one) is more likely to get a satisfaction rate exceeding a given threshold.

It is easy to implement Top-Flop and Threshold betting in many settings in which people receive private binary signals, in the form of tastes or experiences. An application, which we will use as a leading example, is to elicit whether people liked or disliked a movie after previewing it. Previewers are offered bets on some future performance measures of the movie, like the Rotten Tomatoes score or the number of tickets sold, versus those of another movie of the same type. To put it simply, our mechanisms ask people to bet on the *relative* performance of the previewed movie. Doing so alleviates the concern of Keynesian beauty-contest type of herding, when agents act upon what they think others will think, rather than upon their own signal. With a betting mechanism on absolute performance, as in a prediction market, agents' decisions are jointly determined by their private signals and their prior expectations about movie performance. Betting on relative performance, as in our mechanisms, disentangles the private signal from prior expectations, as we will show.

This paper introduces simple betting mechanisms (Top-Flop betting and Threshold betting) and determines sufficient conditions for the chosen bets to reveal private signals. The first part of the paper considers a setting where a single agent receives a signal about one item and bets on its score relative to that of another item belonging to a collection of similar items. In this setting, we assume that the scores are exogenous random variables. There are two key conditions for the agent to reveal his signal through his betting behavior. First, the score of an item must be more informative about the signals related to that item than the scores of other items are. For instance, learning that the previewed movie grossed more than $500M on its first weekend is more informative about the probability to like that specific movie than learning that another movie exceeded the same milestone is. Second, the agent has the same prior for all items of the collection. That is, the agent has no reason to prefer one movie over the other ex ante. Our results do not require the agent to be risk neutral (or even a risk-averse expected-utility maximizer), but simply to choose the bet giving the highest chance to win. Hence, our results are valid for any decision model satisfying first-order stochastic dominance.

In the second part of the paper, we consider a game setting with at least four agents and provide a theoretical foundation for the score. For a given agent, the score for an item in the collection is determined by betting choices of other agents. Similarly to the single-agent case, each agent in a betting game receives a signal about one item in the collection. We again establish sufficient conditions for agents to reveal their signals. Specifically, we do not require that they fully agree on how signals are generated and how signals of any two agents are related. Agents may think they all have a different prior probability to like a given movie. They may even disagree about what these probabilities are. They do agree that the signals of two agents are more positively correlated when the signals are for the same item than for different items. However, they may disagree on the exact degree of correlation.

Several methods have been proposed to reveal unverifiable signals in survey settings (Prelec, 2004; Witkowski and Parkes, 2012b; Radanovic and Faltings, 2013; Baillon, 2017; Cvitanić et al., 2019). They provide truth-telling incentives by asking each agent two questions regarding a single item. One of the questions is directly about the signal and the other one is about predicting other agents' answers. These methods are based on a common-prior assumption, requiring that agents only differ in the signal they received. With these methods, truthful signal reporting is a Bayesian Nash equilibrium when agents are risk neutral. By using more than one item, we can relax the common prior assumption and replace it by an assumption about how the items are related. In other words, in our model, priors may differ across agents but have to agree across items.

Witkowski and Parkes (2012a) also introduced a method that relaxed the common prior assumption but it required to elicit priors before agents receive their signals. We do not require such additional elicitation. Our approach also allows us to use a payment rule that is simpler than the aforementioned mechanisms and is robust to risk aversion, certainty effects, and other behavioral phenomena. Finally, the game-theoretic version of our mechanisms is based on assumptions that are close to those of Dasgupta and Ghosh (2013) and Shnayder et al. (2016). These authors also used cross-item correlations to incentivize truthful signal reporting, also for non-binary signals (Shnayder et al., 2016) and $n \leq 2$ agents, but they needed that all agents get signals for at least two items. The literature is further discussed in Section 4.

We conclude our paper with examples of practical implementations and potential applications of our methods. We show how Threshold betting can be implemented as a financial derivative (an option) of prediction markets. We also explain how our simple bets can be used to assess whether people are willing to pay a given amount for product features that are yet to be developed.

2 Betting on exogenous scores

2.1 Signals, scores, and beliefs

We first consider a setting of a single agent ("he"). There is a *collection of items* $\mathcal{K} \equiv \{1, \ldots, K\}$ with $K \geq 2$. For one fixed $l \in \mathcal{K}$, the agent receives a private signal, modeled as a realization $t \in \mathcal{T} = \{0, 1\}$ of a random variable T. A *center* ("she") wishes to elicit t. For instance, $\mathcal{K}$ is a collection of movies, the agent watches movie l, and the center wants to know whether he liked it ($t = 1$) or not ($t = 0$). Each item $k \in \mathcal{K}$ has a score, reflecting its quality and taking values from $\mathcal{S}$, a countable subset of the reals. The scores are unknown to the agent and to the center when the agent receives t. Furthermore, neither the agent nor the center can influence the scores. Hence, scores are modeled as bounded[1] random variables Y_k with generic realization $y_k \in \mathcal{S}$.

We assume that all the random variables (scores and signals) are defined on the same probability space $(\Omega, \mathcal{F}, P)$. By Kolomogoroff (1933), this can always been assumed. For simplicity, we avoid measure-theoretic complications and assume that Ω is countable, that $\mathcal{F}$ is the sigma-algebra of all subsets of Ω (called *events*), and that P is countably additive.[2] The random variables (and P) need not describe some objective processes but rather the agent's beliefs. His prior probability to get signal 1 is $P\left(t=1\right)$ and H_k denotes the distribution function of his prior about the score.

Assumption 1 (Unique prior)**.** *For any $k \in \mathcal{K} \setminus \{l\}$, Y_k and Y_l are identically distributed, with $H_k = H_l$.*

Let H ($\equiv H_l$) be the unique prior as defined in Assumption 1. Assumption 1 means that the agent has the same expectations about the items in the collection before he receives a signal about item l. In practice, it requires that items are similar. In the movie example, if the score is a performance measure such as reviews or gross revenue, the collection should not mix blockbusters with independent movies because the agent may have very different expectations of the score for the two categories. Dasgupta and Ghosh (2013) and Shnayder et al. (2016) argued for the unique prior assumption when the agent is ignorant about the collection and items are randomly assigned. They typically considered agents completing multiple tasks that are crowd-sourced, such as image labeling, peer-assessment in online courses, or reporting features of hotels and restaurants.

A subset of the score space, useful for what follows, is $\mathcal{S}' = \{y \in \mathcal{S} : 0 <$

[1] A real-valued random variable $Y_k = Y_k(\omega)$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is bounded if there exists a constant M such that $\mid Y_k(\omega) \mid \leq M$ for all $\omega \in \Omega$.

[2] For instance, Ω may be the Cartesian product of the score space and the signal space, $\Omega = (\Pi_{k \in \mathcal{K}} \mathcal{S}) \times \mathcal{T}$.

$H(y) < 1\}$. It excludes all scores that are so low or so high that the agent believes they will never occur. It also excludes the maximum score the agent believes may occur (the smallest y such that $H(y) = 1$). We consider the non-trivial case where the agent believes that more than one score may occur, i.e. $\mathcal{S}'$ not empty.

Assumption 2 (Comparative informativeness)**.** *For all* $k \in \mathcal{K}\backslash\{l\}$ *and* $y \in \mathcal{S}$, $P\,(t = 1 \mid Y_l > y) > P\,(t = 1 \mid Y_k > y)$.

In the mechanism design literature, private signals are linked to states of nature by a signal technology. Here, the possible scores play the role of the states of nature. The signal technology is (believed by the agent to be) such that that the score of item l is more positively associated with receiving a signal 1 about l than the score of item k is.[3] Let the collection of items be, for instance, all movies of a franchise. If the agent learns that movie $l = 4$ scores at least 70% on Rotten Tomatoes, he may update his probability of liking that movie upwards. If instead, he learns that another movie, e.g. $k = 3$, scores at least 70%, he may also update his probability to like movie 4 upwards but less so. He may even decrease his probability to like movie 4 if he thinks that a great movie 3 means a less good movie 4. Our assumption allows for biases or distrust of the underlying scores. For instance, the agent may think that the score is biased by the presence of trolls as long as the biases neither eliminate nor reverse the stronger relation between a high score of l and a signal 1 than between a high score of k and a signal 1.

Once the agent learns his signal t, he updates his beliefs about the scores, which yields the posterior distribution function $F_k^t(y) = P(Y_k \leq y \mid T = t)$. Assumptions 1 and 2 guarantees that the signal influences his expectations

[3]Assumption 2 also implies $P\,(t = 1) \in (0, 1)$ because a degenerate prior would give the same posterior no matter what Y_l and Y_k would be.

about Y_l in a very specific way relatively to any other Y_k. For any two cumulative distribution functions F and G with domain $\mathcal{S}$, we write $F \succeq_{SD} G$ ($F \succ_{SD} G$) and say that F *(strictly) first-order stochastically dominates* G when $F(y) \leq G(y)$ for all $y \in \mathcal{S}$ (with $F(y) < G(y)$ for some y).

Lemma 1. *Assumptions 1 and 2 imply* $F_l^1(y) \succ_{SD} F_k^1(y)$ *and* $F_k^0(y) \succ_{SD} F_l^0(y)$ *for all* $k \neq l$.

The proof of Lemma 1, as all other proofs, is in Appendix. Intuitively, a signal $t = 1$ is more associated with high scores of item l than with high scores of item k and therefore shifts more posterior F_l^1 to the right than posterior F_k^1. Note that we could have immediately assumed the implications of Lemma 1, which would be more general than Assumptions 1 and 2. The advantage of providing sufficient conditions is to clarify what types of items and scores can be used. If the agent believes the score of l is more positively correlated with the signal than the score of k is and views all items of the collection as equivalent, ex ante, in terms of scores, then his beliefs about the scores of l and of any $k \neq l$ once he has received his signal will satisfy the stochastic dominance properties spelled out in Lemma 1. These properties guarantee that signals can be identified from beliefs. Before we design bets based on this identification strategy, we introduce an additional assumption that we will use in some of our results, in which we need the random variables Y_k and Y_l to be not only identically distributed but also independent.

Assumption 3 (Independence)**.** *For any* $k \in \mathcal{K}$ *with* $k \neq l$, Y_k *and* Y_l *are independent, and conditionally independent given* T.

We could also replace conditional independence in Assumption 3, using the fact that Y_k and Y_l are independent, by:

$$\frac{P(t=1 \mid Y_l, Y_k)}{P(t=1 \mid Y_l)} = \frac{P(t=1 \mid Y_k)}{P(t=1)}. \tag{1}$$

In other words, how information about Y_k changes the probability of a positive signal is invariant to information about Y_l.

2.2 The bets

Let π be a *prize* (money, a gift, or... an actual pie) that the agent likes. The absence of prize is denoted by 0. Let $\mathcal{E}$ be an event, an element of $\mathcal{F}$. A *bet on* $\mathcal{E}$ assigns π to $\mathcal{E}$ and 0 to the complement of $\mathcal{E}$. The agent has preferences over bets. If we do not explicitly mention that preferences are strict, we mean weak preferences.

Assumption 4 (Probabilistic sophistication)**.** *For any three events $\mathcal{E}$, $\mathcal{E}'$, and $\mathcal{G} \in \mathcal{F}$, the agent prefers a bet on $\mathcal{E}$ to a bet on $\mathcal{E}'$ when he knows that $\mathcal{G}$ occurred if and only if $P(\mathcal{E} \mid \mathcal{G}) \geq P(\mathcal{E}' \mid \mathcal{G})$.*

Assumption 4 says that the agent is probabilitically sophisticated in the sense of Machina and Schmeidler (1992), and furthermore, that preferences are consistent with P, the (subjective) probability measure that underlies the random variables. He may be risk neutral, or be a risk-averse expected utility maximizer, or even transform his probabilities as long as the transformation is strictly increasing in P so as to satisfy stochastic dominance (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Assumption 4 implies that the agent strictly prefers π (a bet on Ω) to nothing (a bet on $\varnothing$).

Definition 1. *A* Top bet *is a bet on* $\{\omega \in \Omega : Y_l(\omega) > Y_k(\omega)\}$ *and a* Flop bet *is a bet on* $\{\omega \in \Omega : Y_l(\omega) < Y_k(\omega)\}$.

The center proposes a Top bet and a Flop bet to the agent, who may choose one of them (or reject both).

Lemma 2. *Under Assumptions 1 to 4, the agent, before learning t, is indifferent between the Top and the Flop bet but strictly prefers any of them to nothing.*

Ex ante, the agent has the same belief H about the distribution of Y_k and Y_l (Assumption 1), which are also independent (Assumption 3), and there is no reason to prefer betting on one score being higher rather than the other (Assumption 4). Furthermore, the agent does not expect the scores to be equal with certainty, and therefore expects that both bets have a nonnull chance to yield the prize. The agent wants to participate in the betting. When he learns his signal, he has a clear preference for one of the bets, as established by the next Theorem.

Theorem 1. *Under Assumptions 1 to 4, for any $k \in \mathcal{K} \setminus \{l\}$, the agent strictly prefers the Top bet if $t = 1$ and the Flop bet if $t = 0$.*

The following corollary makes explicit that the agent does not need to know k, which can be selected from the collection of items with a random device. We assume, here and whenever we will refer to such exogenous random devices, that they are independent of all the random variables described so far and also conditionally independent given T, and that all elements of the collection have a positive probability to be drawn.

Corollary 1. *Theorem 1 remains valid if k is unknown to the agent and will be randomly drawn from $\mathcal{K} \setminus \{l\}$.*

Our results for the Top and Flop bets rely on (conditional) independence of the scores. The center can also propose another type of simple bets to the agents, which still reveal signals but without relying on independence, only on the stochastic dominance conditions established in Lemma 1. For instance, the agent could be asked to bet on whether the score of item l or the score of item k will exceed some threshold. We call this approach *Threshold betting.*

Definition 2. *A* Threshold-y bet *on* k *is a bet on* $\{\omega \in \Omega : Y_k(\omega) > y\}$.

If the scores are taken from Rotten Tomatoes, a Threshold-60 bet would yield the prize only if the score of the movie exceeds 60%. Ex ante, the agent is indifferent between the items on which the Threshold-y bets are based.

Lemma 3. *Under Assumptions 1 and 4, for any* $y \in \mathcal{S}'$ *and* $k \in \mathcal{K} \setminus \{l\}$*, the agent, before learning* t*, is indifferent between a Threshold-*y *bet on* k *and a Threshold-*y *bet on* l*, but strictly prefers any of them to nothing.*

Lemmas 2 and 3 provide a way to test our assumptions, including the unique prior assumption. Before previewing a movie, the agent should be indifferent between the bets.

Theorem 2. *Under Assumptions 1, 2, and 4, for any* $y \in \mathcal{S}$ *and* $k \in \mathcal{K} \setminus \{l\}$*, the agent strictly prefers a Threshold-*y *bet on* l *to a Threshold-*y *bet on* k *if* $t = 1$ *and a Threshold-*y *bet on* k *to a a Threshold-*y *bet on* l *if* $t = 0$.

Corollary 2. *Theorem 2 remains valid if* k *is unknown to the agent and will be randomly drawn from* $\mathcal{K} \setminus \{l\}$ *and/or if* y *is unknown to the agent and will be randomly drawn from* $\mathcal{S}$.

A challenge of Theorem 2 is to find a value from the support to use as threshold, because the support, unlike the domain, is subjective. The center can mitigate the problem by avoiding extreme values. Corollary 2 solves the challenge by proposing to randomly draw a value from $\mathcal{S}$ after the agent chooses a bet.

Before receiving a signal, the agent is indifferent between Top and Flop bets (Lemma 2) and also between Threshold-y bets on l and Threshold-y bets on k (Lemma 3). No matter which signal he receives, his winning probability always increases if he chooses optimally. With Threshold-y bets, the winning probability with optimal choices is $P(t = 1)P(Y_l > y \mid t = 1) + P(t = 0)P(Y_k > y \mid t = 0)$, which strictly exceeds the no-signal chance of winning $P(Y_l > y)$ $(= P(Y_k > y))$.[4] The difference between the two gives us the ex ante value of the signal (in terms of winning chances). The same reasoning applies to Top-Flop betting.

Now imagine that the agent has to pay a cost (or provide an effort) to acquire the signal. He will compare this cost to the benefit–the increase in the probability of getting π.

Remark 1. *The ex ante value of the signal is positive. Hence, under common regularity assumptions (continuity in utility), there exists a non-degenerate range of costs that the agent is willing to pay to acquire the signal.*

How much (effort) the agent is willing to spend on the signal will depend on his whole utility function. Calculating it would require specifying further assumptions about the decision model of the agent (beyond Assumption 4). Obviously, we can expect that increasing the value of the prize will increase the maximum cost the agent is willing to pay. What we claim is that our simple bets can stimulate signal acquisition. In practice, they can be used to motivate people to look for a piece of information, preview a movie, or carefully evaluate a product.

[4] Proof: $P(Y_l > y) = P(t = 1)P(Y_l > y \mid t = 1) + P(t = 0)P(Y_l > y \mid t = 0)$ by definition. Replacing the $P(Y_l > y \mid t = 0)$ by the strictly larger $P(Y_k > y \mid t = 0)$ (according to Theorem 2) establishes the result.

3 Betting on endogenous scores

3.1 Agents, their signals, and their beliefs

We now consider multiple agents $i \in \mathcal{I} = \{1, \cdots, Kn\}$, i.e., $n \geq 2$ agents per item. In the simplest case, with two items, we need a minimum of 4 agents. In this section, most variables and objects from the previous section become agent-specific, which will be indicated by subscript i. Each agent i gets a signal $T_i \in \mathcal{T} = \{0, 1\}$, about item $l_i \in \mathcal{K}$. The set of agents with a signal about k is $\mathcal{I}_k \equiv \{j \in \mathcal{I} : l_j = k\}$ and it has cardinality n. The state space is $\Omega = \mathcal{T}^{Kn}$, where a state ω is the vector of signals received by the Kn agents. (We need not specify scores here, as will become apparent later.)

Agent i will be offered to bet on scores based on the others' actions in the games to be defined in the next subsection. For item $k = l_i$, "the others" mean $\mathcal{I}_{i,k} \equiv \mathcal{I}_k \setminus \{i\}$. In what follows, it will be desirable to consider sets of agents with the same cardinality as this set of others. We, therefore, define for items $k \neq l_i$, $\mathcal{I}_{i,k} \equiv \mathcal{I}_k \setminus \{j\}$ with $j = \max \mathcal{I}_k$ (any other j could have been chosen as well). We can now define the analog of the random variables Y_k of the preceding section. For all i and k,

$$Y_{i,k} = \sum_{j \in \mathcal{I}_{i,k}} T_j. \tag{2}$$

The random variable $Y_{i,k}$ is, for agent i, the number of other agents who received signal 1 for item k. As in the previous section, agent i's belief P_i, defined over Ω, generates distribution priors $H_{i,k}$ about $Y_{i,k}$. The domain of $H_{i,k}$ is $\mathcal{S}_i = \mathcal{S} = \{0, \ldots, n-1\}$ because $Y_{i,k}$ can take values between 0 and $n - 1$. The sets $\mathcal{S}'_i$ is defined similarly as $\mathcal{S}'$ in the preceding section.

Example 1. *The simplest case of our setting is $n = K = 2$, involving four* *agents. State ω is a quadruplet of signals (t_1, t_2, t_3, t_4). With $l_1 = l_2 = 1$, $l_3 = l_4 = 2$, $\mathcal{I}_{1,2} = \{3\}$, and $\omega = (t_1, t_2, t_3, t_4)$, we have $Y_{1,1}(\omega) = t_2$ and $Y_{1,2}(\omega) = t_3$.*

Assumption 5 (Common knowledge)**.** *It is common knowledge that Assumption 4 holds for all agents $i \in \mathcal{I}$, with all P_is themselves common knowledge.*

Assumption 5 means that agents may all have different P_i but they know that everyone satisfies first order stochastic dominance with respect to their own beliefs. Furthermore, if Assumptions 1, 2, and 3 hold for all P_i, then this fact is automatically common knowledge because the beliefs P_is are themselves common knowledge. Assumptions 1, 2, and 5 do not require that all agents in $\mathcal{I}_k$ have the same probability to get a signal 1. Agent i can think everyone is different, and even that some people dislike everything (trolls). What we need is that each agent i perceives T_i and $Y_{i,k}$ more associated when $k = l_i$ than when $k \neq l_i$. Independence (Assumption3) can now be justified if, for instance, signals of any two agents i and j are independent when $l_i \neq l_j$.

3.2 The games

We first define a generic game with the same *action set* $\mathcal{A} = \{0, 1\}$ for all agents, with a_i the action of agent i. The *payoff function* of the game for agent i is $\Pi_i : \mathcal{A}^{Kn} \longrightarrow \{0, \pi\}$. Each agent chooses a *strategy*, which is a pair of actions $(a_i^0, a_i^1) \in \mathcal{A}^2$, where a^0 will be implemented in state ω if $T_i(\omega) = 0$ and a^1 will be implemented if $T_i(\omega) = 1$. A *strategy profile*, i.e. the strategy of all agents, is denoted by $(a^0, a^1) \in (\mathcal{A}^2)^{Kn}$. The *implemented action* for agent i in state ω is $a_i^{T_i(\omega)}$, which we write a_i^ω for short. We similarly denote $a^\omega \in \mathcal{A}^{Kn}$ the *profile of implemented actions*.

Example 1 (continued)**.** *A strategy profile is of the form* $((a_1^0, a_1^1), (a_2^0, a_2^1), (a_3^0, a_3^1), (a_4^0, a_4^1))$. *If the realized state is* $\omega = (0, 1, 1, 0)$*, then the profile of implemented actions is* $a^\omega = (a_1^0, a_2^1, a_3^1, a_4^0)$*. The payoff function* Π_i *of agent* i *assigns either 0 or* π *to any such quadruplet.*

The agents have preferences over strategy profiles, conditional on their signal and denoted by $\succsim_{i|T_i}$. Assumption 5, which includes Assumption 4, implies that it is common knowledge that $(a^0, a^1) \succsim_{i|T_i} (b^0, b^1)$ if and only if

$$P_i\left(\{\omega \in \Omega : \Pi_i\left(a^\omega\right) = \pi\} \mid T_i\right) \geq P_i\left(\{\omega \in \Omega : \Pi_i\left(b^\omega\right) = \pi\} \mid T_i\right). \qquad (3)$$

In Equation 3, the agent first determines which are the states ω yielding π if the strategy profile is (a^0, a^1), and if the strategy profile is (b^0, b^1). The agent then compares the probability (given his signal) of the states yielding π when the strategy profile is (a^0, a^1) to the probability obtained if the strategy profile is (b^0, b^1). Agent i finally chooses the strategy profile that gives the higher chance to get π.

With $\mathcal{I}$, Ω, $\mathcal{A}$, $\mathcal{T}$, T_i, P_i, and $\succsim_{i|T_i}$, we have defined a Bayesian game (Osborne and Rubinstein, 1994, Definition 25.1).[5] Let $(b_i^0, b_i^1; a^0, a^1)$ be the strategy profile, which replaces a_i^0 and a_i^1 by b_i^0 and b_i^1 in (a^0, a^1). Following (Osborne and Rubinstein, 1994, Definition 26.1), a strategy profile (a^0, a^1) is a *Nash equilibrium* of the Bayesian game if for all $i \in \mathcal{I}$, $(a^0, a^1) \succsim_{i|T_i} (b_i^0, b_i^1; a^0, a^1)$ for all $(b_i^0, b_i^1) \in \mathcal{A}^2$. We say that the Nash equilibrium is *strict* if, in addition and for all i, $(a^0, a^1) \succ_{i|T_i=0} (b_i^0, a_i^1; a^0, a^1)$ for all $b_i^0 \in \mathcal{A} \setminus \{a^0\}$ and $(a^0, a^1) \succ_{i|T_i=1} (a_i^0, b_i^1; a^0, a^1)$ for all $b_i^1 \in \mathcal{A} \setminus \{a^1\}$. Strict means that the implemented action is strictly preferred (even though the not-implemented action is only weakly preferred).

[5]We assume common knowledge of Ω, $\mathcal{I}$, $\mathcal{T}$, $\mathcal{A}$, and the Π_is.

We can now define Top-Flop and Threshold-y games. Each agent i will be offered bets on (individualized) scores $\widehat{Y}_{i,k}$ defined as a function of an action profile $a \in \mathcal{A}^{Kn}$ by:

$$\widehat{Y}_{i,k} = \sum_{j \in \mathcal{I}_{i,k}} a_j. \tag{4}$$

In Section 2, the scores were exogenous and agents had beliefs about them. In the present section, we provide a game-theoretic foundation for the scores, which are endogenously defined by the actions of others. Agents now have beliefs about signals, which translate into beliefs about scores $\widehat{Y}_{i,k}$ for a given strategy profile. The payoff function of the game is defined on the $\widehat{Y}_{i,k}$s. We first assign h_i to each agent i, given by $h_i = l_i + 1$ if $l_i < K$ and $h_K = 1$.

Definition 3. *In a* Top-Flop game, Π_i *assigns* π *to* $\left\{a \in \mathcal{A}^{Kn} : a_i = 1 \ \& \ \left(\widehat{Y}_{i,l_i} > \widehat{Y}_{i,h_i}\right)\right\}$ *(Top case) and to* $\left\{a \in \mathcal{A}^{Kn} : a_i = 0 \ \& \ \left(\widehat{Y}_{i,l_i} < \widehat{Y}_{i,h_i}\right)\right\}$ *(Flop case). It assigns* 0 *to all other elements of* $\mathcal{A}^{Kn}$*.*

The payoff function is defined such that choosing action 1 is equivalent to choosing a Top bet; it pays π if $\widehat{Y}_{i,l_i} > \widehat{Y}_{i,h_i}$. Similarly, choosing action 0 is equivalent to choosing a Flop bet, which pays off if $\widehat{Y}_{i,l_i} < \widehat{Y}_{i,h_i}$.

Example 1 (continued)**.** *With* $l_1 = l_2 = 1$, $l_3 = l_4 = 2$, *agents 1 and 2 get a signal about item 1, and agents 3 and 4 get a signal about item 2. Furthermore,* $\widehat{Y}_{1,1} = a_2$ *and* $\widehat{Y}_{1,2} = a_3$, *which means agent 1 bets on the actions of agents 2 and 3. The following table describes* Π_1*.*

$\widehat{Y}_{1,1}$	$\widehat{Y}_{1,2}$	$a_1 = 0$	$a_1 = 1$
$a_2 = 0$	$a_3 = 0$	*0*	*0*
$a_2 = 0$	$a_3 = 1$	π	*0*
$a_2 = 1$	$a_3 = 0$	*0*	π
$a_2 = 1$	$a_3 = 1$	*0*	*0*

First note that for agent 1, the action of agent 4 does not affect his payment. Second, he wins π in two cases: (i) if he and agent 2 report 0 while agent 3 reports 1 and (ii) if he and agent 2 report 1 while agent 3 reports 0. Case (i) is a Flop bet, where item 2 gets a higher score ($\widehat{Y}_{1,2} = 1$) than item 1 ($\widehat{Y}_{1,1} = 0$). Symmetrically, case (ii) is a Top bet.

Theorem 3. *If P_i satisfies Assumptions 2 to 3 for all $i \in \mathcal{I}$ and if Assumption 5 holds, then (a^0, a^1) with $a_i^0 = 0$ and $a_i^1 = 1$ for all $i \in \mathcal{I}$ is a strict Nash equilibrium of a Top-Flop game.*

In the proof (Appendix B), we first establish that if every $j \neq i$ plays $(0, 1)$, then $\widehat{Y}_{i,k} = Y_{i,k}$ for all k. By Theorem 1, the best response of agent i is then to choose a Flop bet if $T_i = 0$ and a Top bet if $T_i = 1$, hence picking strategy profile $(0, 1)$. All this is common knowledge, so the agents' beliefs are consistent with the Nash equilibrium.

Corollary 3. *Under the assumptions of Theorem 3, all agents strictly prefer the equilibrium of a Top-Flop game in which all agents play* $(0, 1)$ *to all agents playing* $(0, 0)$ *or all agents playing* $(1, 1)$.

By construction, degenerate strategy profiles where everyone plays $(0, 0)$ or everyone plays $(1, 1)$ yields payoff 0. Hence, the equilibrium $(0, 1)$ is preferred because it gives a chance to get π. We now turn to Threshold-y betting that we similarly transform into a game.

Definition 4. *In a* Threshold-y game, *for* $y \in \{0, \ldots, n-2\}$, Π_i *assigns* π *to* $\left\{a \in \mathcal{A}^{Kn} : a_i = 1 \ \& \ \left(\widehat{Y}_{i,l_i} > y\right)\right\}$ *and to* $\left\{a \in \mathcal{A}^{Kn} : a_i = 0 \ \& \ \left(\widehat{Y}_{i,h_i} > y\right)\right\}$. *It assigns* 0 *to all other elements of* $\mathcal{A}^{Kn}$.

With the payoff functions of a Threshold-y game, agent i gets π when playing 1 if item l_i exceeds threshold y and when playing 0 if item h_i exceeds threshold y. The threshold can be any value up to $n-1$ because $\widehat{Y}_{i,k}$ can never exceed n.

Example 2 (continued). *With four agents, only a Threshold-0 game is possible.*[6] *Agent 1 still bets on the actions of agents 2 and 3 but* Π_1 *is now:*

$\widehat{Y}_{1,1}$	$\widehat{Y}_{1,2}$	$a_1 = 0$	$a_1 = 1$
$a_2 = 0$	$a_3 = 0$	*0*	*0*
$a_2 = 0$	$a_3 = 1$	π	*0*
$a_2 = 1$	$a_3 = 0$	*0*	π
$a_2 = 1$	$a_3 = 1$	π	π

Agent 1 wins π *in two cases: (i) if he and agent 2 play 1* ($a_1 = a_2 = 1$) *and (ii) if he plays 0 while agent 3 plays 1* ($a_1 = 0$ *and* $a_3 = 1$). *Case (i) is a bet on the score of item 1 (= the action of agent 2) exceeding 0 and case (ii) a bet on the score of item 2 (= the action of agent 3) exceeding 0. The last row of the table differs from the Top-Flop game.*

Theorem 4. *If* P_i *satisfies Assumptions 1 and 2 for all* $i \in \mathcal{I}$ *and if Assumption 5 holds, then* (a^0, a^1) *with* $a_i^0 = 0$ *and* $a_i^1 = 1$ *for all* i *is a strict Nash equilibrium of a Threshold-*y *game when* $y \in \mathcal{S}_i'$ *for all* i.

Corollary 4. *Under the assumptions of Theorem 4,* (a^0, a^1) *with* $a_i^0 = 0$ *and* $a_i^1 = 1$ *for all* i *is a strict Nash equilibrium of a Threshold-*y *game when* y *is randomly drawn from* $\mathcal{S}$.

Theorem 4 has two main limitations. First, all agents must think the threshold is not trivial, neither too high nor too low. A solution, given by Corollary 4 is to randomly draw the threshold ex post. Second, unlike in the Top-Flop game, there exists an equilibrium that would be preferred by all

[6] $\widehat{Y}_{i,k}$ can only be 0 or 1, and therefore can only strictly exceed 0.

agents to playing $(1, 0)$. If they all play $(1, 1)$, they can all win with certainty. This equilibrium can be excluded by altering Π_i such that it is 0 if $\widehat{Y}_{i,l_i} = \widehat{Y}_{i,h_i} = n - 1$ (the maximum score). This modification of the payoff function is not anodyne though, and requires to bring back Assumption 3.[7]

4 Discussion

4.1 Limitations and related literature

In the exogenous-score setting, it is important that the agent does not expect the center to have control over Y_k. A suspicious agent would then enter a game with the center. Suspicion can be avoided or at least mitigated by using scores controlled by an independent third party or involving a multitude of people. For instance, the score can be the price established on a large prediction market at a given time. This would make clear that influencing the score would cost more to the center than paying π to the agent.

Our exogenous-score setting relates to the literature on canonical contract design for adverse selection problems as in Mirrlees (1971), Maskin and Riley (1984) and Baron and Myerson (1982). For instance, in the classical monopoly setting, the principal (the center in our setting) does not know the agent's private information, but she can screen different types of agents by offering them an incentive compatible menu of contracts, under which the agent will pick the one revealing his true type. Since the screening is achieved by leveraging the structure of agents' preferences, the principal is required to know the preference for each type and its distribution. Our methods do not require that because our screening techniques are mainly based on the complementarity

[7] The probability of getting π does not depend anymore on either $\widehat{Y}_{i,l_i}$ if $a_i = 1$ or $\widehat{Y}_{i,h_i}$ if $a_i = 0$, but on both $\widehat{Y}_{i,l_i}$ and $\widehat{Y}_{i,h_i}$ for all a_i.

between the score and the private signal for each agent. This is possible because, in our setting, agents have no other incentives (to either reveal or hide the signals) than trying to win the prize.

Our Bayesian game setting relates to a strand of literature in mechanism design, for instance, as in Cremer and McLean (1988) and Maskin (1999). Both mechanisms construct truth-telling equilibrium by exploiting the correlation of private information across agents. More recently, Bayesian methods to elicit private signals in surveys or on crowd-sourcing platforms have been proposed by Prelec (2004), Miller et al. (2005), Witkowski and Parkes (2012b), Radanovic and Faltings (2013), Baillon (2017), and Cvitanić et al. (2019). All these papers rely on common prior assumptions, sometimes weakly relaxing them. Our common knowledge assumption is much weaker, allowing all agents to disagree on the probability to observe some signals. Note that for the Nash equilibrium to be credible, the key point is not so much that agents know the priors of all other agents but rather that they know that these priors are well-behaved as described by Assumptions 2 to 3. Our beliefs assumptions are very close to those of Dasgupta and Ghosh (2013) and Radanovic et al. (2016). These papers consider a signal correlation matrix and assume that it describes the beliefs of all agents. However, Radanovic et al. (2016) do point out that only the structure of the correlations matters and therefore heterogeneity in beliefs would be possible (their footnote 7 and subsection 5.4). Unlike the present paper, Dasgupta and Ghosh (2013) and Radanovic et al. (2016) only consider game settings and require that each agent receives signals about two items (in their setting, performs two tasks) whereas our agents receive a signal about only one item.

A major limitation of our paper, which is shared by Dasgupta and Ghosh (2013) but not by Radanovic et al. (2016), is that we can only handle binary signals. Extending our results to non-binary signals is not trivial and would require much heavier assumptions about beliefs, especially correlations between signals and scores. With binary signals, signal 1 being associated with high scores means that signal 0 is associated with low scores. With non-binary signals, such implications do not hold anymore. Imagine that signals are satisfaction levels $\{1, 2, 3\}$ and that we have, for each item k, three scores Y_k^1, Y_k^2, and Y_k^3 (for instance, the number of other agents reporting signals 1, 2, and 3 respectively). An agent with satisfaction level 3 can reasonably increase the probability that Y_k^3 is at least y but also the probability that Y_k^2 is at least y. A possible approach is to split the agent sample between three groups. Some agents get the possibility to bet on Y_k^3 vs Y_l^3, which can reveal whether their signal was 3 or not 3. Other agents get the possibility to bet on Y_k^2 vs Y_l^2 and the last ones on Y_k^1 vs Y_l^1.

Throughout the paper, we implicitly assumed that the center, offering the bets or organizing the games, is willing to pay up to π for each signal. Often, participation in surveys or experiments is rewarded. What we propose here is to use this reward as prize π, to make agents reveal their signal instead of only rewarding them for providing *any* answer. Our results from the game setting assume that agents cannot communicate. If they could, a full coalition can make sure they get π with probability 1 if K is even and all agents with even items play 1 and all agents with odd items play 0. A way to deter such coalitions is to make the game zero-sum.

4.2 Practical implementation and examples

Organizing Top-Flop or Threshold betting on exogenous scores is easier to implement in practice than the respective game versions. Threshold betting can, for instance, be combined with prediction markets. When people predict the score of a movie or the results of a song contest, they do not report their own taste but their beliefs about others. Threshold betting, where the score is defined as the price in the prediction markets for items l and k at a given time, reveals people's own taste (under the assumptions and setting of Section 2). A threshold-y bet on prediction market k is a digital option that pays π if the price reaches y. In other words, Top-Flop and Threshold bets can be implemented as derivatives of existing markets.

Let us conclude with two other examples. The director of a company hesitates where to invest in research and development. There is a set $\mathcal{K}$ of possible product features that could be developed. The director would like to know for which feature the consumers would be willing to pay $100 more. These features do not exist yet and therefore cannot be sold to consumers. Hence, eliciting the willingness-to-pay cannot be incentivized, for instance with the Becker-deGroot-Marschak mechanism (Becker et al., 1964), because it would require actually selling the features. Instead, the director could implement a Top-Flop game among a panel of consumers, organized in K subgroups. Each subgroup of panelists are informed about a feature, and have to bet Top or Flop, not knowing what the other possible innovative features are. A final example of possible application concerns environmental research. It is not always possible to incentivize the elicitation of the willingness-to-pay to save (or the willingness-to-accept for not saving) endangered species. Our simple bets can help there as well by providing subgroups of respondents with information about one (rare) species and ask them whether more people would pay a given amount to save the species they were informed about rather than another random species.

5 Conclusion

This paper introduced two methods, Top-Flop and Threshold betting, to elicit private signals. The first part of the paper showed how to transform pre-existing scores, which may be biased or only partially-informative, into a mechanism incentivizing truthful revelation of signals. An agent betting on the scores need not fully trust them, but only believe that they are somewhat associated with the signals. In the second part of the paper, the scores naturally arise from the other agents' betting decisions. In retrospect, our bets, and therefore our mechanisms, look quite simple but they have been overlooked so far, in favor of more complex approaches. The payment rules of Top-Flop and Threshold bets are transparent, with a unique, fixed prize assigned to a well-defined event. We established conditions ensuring that Top-Flop and Threshold betting properly reveal signals. These conditions are to be milder in terms of individual preferences than typically assumed in the literature, and therefore more likely to be satisfied in practical applications.

Appendix A Proofs for the single-agent setting

A.1 Proof of Lemma 1

Proof. The posterior cumulative distribution for item l is $F_l^1(y) = 1 - P(Y_l > y \mid t = 1)$. By Bayes rule, we have

$$P(Y_l > y \mid t = 1) = \frac{P(t = 1 \mid Y_l > y)}{P(t = 1)} \times P(Y_l > y). \tag{5}$$

By definition, $P(Y_l > y) = 1 - H_l(y)$, and by by Assumption 1, $1 - H_l(y) = 1 - H_k(y) = P(Y_k > y)$. Furthermore, Assumption 2 states that $P(t = 1 \mid Y_l > y) > P(t = 1 \mid Y_k > y)$ if $y \in \mathcal{S}'$. Hence, we have

$$P(Y_l > y \mid t = 1) > \frac{P(t = 1 \mid Y_k > y)}{P(t = 1)} \times P(Y_k > y) = P(Y_k > y \mid t = 1) \tag{6}$$

if $y \in \mathcal{S}'$ and

$$P(Y_l > y \mid t = 1) = P(Y_k > y \mid t = 1) = P(Y_k > y) \tag{7}$$

otherwise. As a conclusion, $F_l^1(y) \succ_{SD} F_k^1(y)$ for all $y \in \mathcal{S}$.

We now consider $t = 0$. By definition,

$$P(Y_l > y \mid t = 0) = \frac{P(Y_l > y) - P(Y_l > y \mid t = 1) P(t = 1)}{P(t = 0)} \tag{8}$$

and

$$P(Y_k > y \mid t = 0) = \frac{P(Y_k > y) - P(Y_k > y \mid t = 1) P(t = 1)}{P(t = 0)}. \tag{9}$$

By Assumption 1, $P(Y_l > y) = P(Y_k > y)$ and By Eqs. 6 and 7, $F_k^1(y) \succ_{SD} F_l^1(y)$ for all $y \in \mathcal{S}$. □

A.2 Proof of Lemma 2

Proof.

$$\begin{aligned} P(\{\omega \in \Omega : Y_l(\omega) < Y_k(\omega)\}) &= P\left(\bigcup_{s \in \mathcal{S}} \{\omega \in \Omega : Y_l(\omega) = s\} \cap \{\omega \in \Omega : Y_k(\omega) > s\}\right) \\ &= \sum_{s \in \mathcal{S}} P(\{\omega \in \Omega : Y_l(\omega) = s\} \cap \{\omega \in \Omega : Y_k(\omega) > s\}) \\ &= \sum_{s \in \mathcal{S}} P(\{\omega \in \Omega : Y_l(\omega) = s\}) \times P(\{\omega \in \Omega : Y_k(\omega) > s\}) \\ &= \sum_{s \in \mathcal{S}} P(Y_l = s) \times (1 - H_k(s)) . \end{aligned} \tag{10}$$

The second equality comes from events $\{\omega \in \Omega : Y_l(\omega) = s\}$ for any two s being disjoint. Independence (Assumption 3) implies the third equality. Because Y_l and Y_k are identically distributed, $P(Y_l = s) = P(Y_k = s)$ and $H_k(s) = H_l(s)$ for all s and therefore, $P(\{\omega \in \Omega : Y_l(\omega) < Y_k(\omega)\}) = P(\{\omega \in \Omega : Y_l(\omega) > Y_k(\omega)\})$. By Assumption 4, the agent is indifferent between the Top and the Flop bet.

By Assumption 4, the agent would prefer a bet on $\varnothing$ to the Top bet or to the Flop bet if and only if $P(\{\omega \in \Omega : Y_l(\omega) > Y_k(\omega)\}) = 0$ or $P(\{\omega \in \Omega : Y_l(\omega) < Y_k(\omega)\}) = 0$. We have just shown that $P(\{\omega \in \Omega : Y_l(\omega) > Y_k(\omega)\}) = P(\{\omega \in \Omega : Y_l(\omega) < Y_k(\omega)\})$. Hence, the agent would prefer a bet on $\varnothing$ if and only if $P(\{\omega \in \Omega : Y_l(\omega) = Y_k(\omega)\}) = 1$. This implies $P(\{\omega \in \Omega : Y_l(\omega) = Y_k(\omega)\} \mid t = 1) = 1$ and therefore, $F_l^1(y) = F_k^1(y)$. The latter contradicts $F_l^1(y) \succ_{SD} F_k^1(y)$ and according to Lemma 1, is therefore incompatible with Assumptions 1 and 2. As a consequence, under Assumptions 1 to 4, the agent must strictly prefer any of the bets he is offered to nothing. □

A.3 Proof of Theorem 1

Proof. Assume $t = 1$.

$$\begin{aligned}
&P\left(\{\omega \in \Omega : Y_l(\omega) < Y_k(\omega)\} \mid t=1\right) \\
=&P\left(\bigcup_{s\in\mathcal{S}} \{\omega \in \Omega : Y_l(\omega) = s\} \cap \{\omega \in \Omega : Y_k(\omega) > s\} \mid t=1\right) \\
=&\sum_{s\in\mathcal{S}} P\left(\{\omega \in \Omega : Y_l(\omega) = s\} \cap \{\omega \in \Omega : Y_k(\omega) > s\} \mid t=1\right) \\
=&\sum_{s\in\mathcal{S}} P\left(\{\omega \in \Omega : Y_l(\omega) = s\} \mid t=1\right) \times P\left(\{\omega \in \Omega : Y_k(\omega) > s\} \mid t=1\right) \\
=&\sum_{s\in\mathcal{S}} P\left(Y_l = s \mid t=1\right) \times \left(1 - F_k^1(s)\right).
\end{aligned} \tag{11}$$

The second equality comes from events $\{\omega \in \Omega : Y_l(\omega) = s\}$ being disjoint for any two s. Conditional independence (Assumption 3) implies the third equality.

$$\begin{aligned}
&P\left(\{\omega \in \Omega : Y_l(\omega) > Y_k(\omega)\} \mid t=1\right) \\
=&\sum_{s\in\mathcal{S}} P\left(Y_k = s \mid t=1\right) \times \left(1 - F_l^1(s)\right) \\
>&\sum_{s\in\mathcal{S}} P\left(Y_k = s \mid t=1\right) \times \left(1 - F_k^1(s)\right) \\
\geq&\sum_{s\in\mathcal{S}} P\left(Y_l = s \mid t=1\right) \times \left(1 - F_k^1(s)\right) \\
=&P\left(\{\omega \in \Omega : Y_l(\omega) < Y_k(\omega)\} \mid t=1\right)
\end{aligned} \tag{12}$$

The first equality comes from Eq. 11 (reversing l and k) and the following inequality from Lemma 1 because $F_l^1(s) \succ_{SD} F_k^1(s)$ means that $F_l^1(s) \leq F_k^1(s)$ with a strict inequality for some s. Notice that stochastic dominance also implies that Y_l can be obtained from Y_k by moving probability mass from low values of $\mathcal{S}$ to high values of $\mathcal{S}$. The weights $(1 - F_k^1(s))$ are lower for high values of $\mathcal{S}$ than for low values and therefore, replacing Y_k by Y_l decreases the whole sum, which justifies the fourth line of the equation. The final line is obtained from Eq. 11.

Together with Assumption 4, Eq. 12 implies that the agent prefers the Top bet when his signal is $t = 1$. The proof from $t = 0$ is symmetric. □

A.4 Proof of Corollary 1

Proof. If k is randomly chosen in $\mathcal{K} \setminus \{l\}$, with the random device being independent of all random variables and conditionally independent given T, then the winning probability of the Top and Flop bets does not change and the preferences given in Theorem 1 still hold. □

A.5 Proof of Lemma 3

Proof. Under Assumption 1, $H_k(y) = H_l(y) > 0$ for all $y \in \mathcal{S}'$. This, together with Assumption 4, gives the result. □

A.6 Proof of Theorem 2

Proof. From Lemma 1, we know $F_l^1(y) \succ_{SD} F_k^1(y)$ and $F_k^0(y) \succ_{SD} F_l^0(y)$ for all $k \neq l$. More precisely, the proof showed $F_l^1(y) < F_k^1(y)$ for all $y \in \mathcal{S}'$, and by symmetry, $F_l^0(y) > F_k^0(y)$. We obtain, for all $y \in \mathcal{S}'$, $P(Y_l > y \mid t = 1) > P(Y_k > y \mid t = 1)$ and $P(Y_l > y \mid t = 0) < P(Y_k > y \mid t = 0)$. Assumption 4 then implies the preferences described in the theorem. □

A.7 Proof of Corollary 2

Proof. If k is randomly chosen in $\mathcal{K} \setminus \{l\}$, with the random device being independent of all random variables and conditionally independent given T, then the winning probability of bets do not change and the preferences given in Theorem 1 remain.

If y is drawn from $\mathcal{S}$, either $y \in \mathcal{S}'$ and the strict preferences mentioned in Theorem 2 hold, or the events are equally likely and the agent would be indifferent. Hence, before knowing y, the strict preferences mentioned in Theorem 2 hold. $\square$

Appendix B Proofs for the game setting

B.1 Proof of Theorem 3

Proof. Consider $(b_i^0, b_i^1; a^0, a^1)$ with $a_j^0 = 0$ and $a_j^1 = 1$ for all $j \neq i$ and $(b_i^0, b_i^1) \in \mathcal{A}^2$. Hence, in state ω, $\widehat{Y}_{i,k} = \sum_{j \in \mathcal{I}_{i,k}} a_j^{T_j(\omega)} = \sum_{j \in \mathcal{I}_{i,k}} T_j(\omega)$, which implies $\widehat{Y}_{i,k} = Y_{i,k}(\omega)$ for all k, and noticeably for l_i and h_i. Under Assumption 5, it is common knowledge that Assumptions 1 to 4 are satisfied. Therefore, applying Theorem 1, it is also common knowledge that agent i strictly prefers $a_i^1 = 1$ to $b_i^1 = 0$ (when b_i^0 is fixed) if $T_i = 1$ and strictly prefers $a_i^0 = 0$ to $b_i^0 = 1$ (when b_i^1 is fixed) if $T_i = 0$. $P_i(T_i = 0 \mid T_i = 1) = P_i(T_i = 1 \mid T_i = 0) = 0$ implies that the agent is indifferent between $a_i^0 = 1$ and $b_i^0 = 0$ (when b_i^1 is fixed) if $T_i = 1$ and between $a_i^1 = 0$ and $b_i^1 = 1$ (when b_i^0 is fixed) if $T_i = 0$. Hence, it is common knowledge that a best response of i to $a_j^0 = 0$ and $a_j^1 = 1$ for all $j \neq i$ is $a_i^0 = 0$ and $a_i^1 = 1$ and therefore, (a^0, a^1) is a Nash equilibrium. It is a strict Nash equilibrium because we showed $(0, 1)$ is strictly preferred to $(1, 1)$ given $T_i = 0$ and $(0, 1)$ is strictly preferred to $(1, 0)$ given $T_i = 1$. $\square$

B.2 Proof of Corollary 3

Proof. Note that the strategy profiles with $b_i^0 = b_i^1 = 0$ for all i gives payment 0 to everyone. The same is true for $b_i^0 = b_i^1 = 1$. By contrast, the equilibrium in Theorem 3 is strict, which would not be possible if the payment was 0. □

B.3 Proof of Theorem 4

Proof. The proof is similar to that of Theorem 3, simply using Theorem 2 instead of Theorem 1. □

B.4 Proof of Corollary 4

Proof. The proof is similar to that of Corollary 2. □

References

Baillon, A. (2017). Bayesian markets to elicit private information. *Proceedings of the National Academy of Sciences*, 114(30):7958–7962.

Baron, D. P. and Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica: Journal of the Econometric Society*, pages 911–930.

Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, 9(3):226–232.

Cremer, J. and McLean, R. P. (1988). Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica: Journal of the Econometric Society*, pages 1247–1257.

Cvitanić, J., Prelec, D., Riley, B., Tereick, B., et al. (2019). Honesty via choice-matching. *American Economic Review: Insights*, 1(2):179–92.

Dasgupta, A. and Ghosh, A. (2013). Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13, pages 319–330, New York, NY, USA. ACM.

Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2):263.

Kolomogoroff, A. N. (1933). *Grundbegriffe der wahrscheinlichkeitsrechnung*. Springer, Berlin. Translated into English by Nathan Morrison (1950) *Foundations of the Theory of Probability*, Chelsea, New York.

Machina, M. J. and Schmeidler, D. (1992). A more robust definition of subjective probability. *Econometrica: Journal of the Econometric Society*, pages 745–780.

Maskin, E. (1999). Nash equilibrium and welfare optimality. *The Review of Economic Studies*, 66(1):23–38.

Maskin, E. and Riley, J. (1984). Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196.

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

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208.

Osborne, M. J. and Rubinstein, A. (1994). *A course in game theory.* MIT press.

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

Radanovic, G. and Faltings, B. (2013). A robust Bayesian truth serum for non-binary signals. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, AAAI'13, pages 833–839. AAAI Press.

Radanovic, G., Faltings, B., and Jurca, R. (2016). Incentives for effort in crowdsourcing using the peer truth serum. *ACM Trans. Intell. Syst. Technol.*, 7(4):48:1–48:28.

Shnayder, V., Agarwal, A., Frongillo, R., and Parkes, D. C. (2016). Informed truthfulness in multi-task peer prediction. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, EC '16, pages 179–196, New York, NY, USA. ACM.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

Witkowski, J. and Parkes, D. C. (2012a). Peer prediction without a common prior. In *Proceedings of the 2012 ACM Conference on Electronic Commerce*, EC '12, pages 964–981, New York, NY, USA. ACM.

Witkowski, J. and Parkes, D. C. (2012b). A robust Bayesian truth serum for small populations. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*.