

Crowdsourced Outcome Determination in Prediction Markets

Rupert Freeman
Duke University
rupert@cs.duke.edu

Sébastien Lahaie
Microsoft Research
slahaie@microsoft.com

David M. Pennock
Microsoft Research
dpennock@microsoft.com

Abstract

A prediction market is a useful means of aggregating information about a future event. To function, the market needs a trusted entity who will verify the true outcome in the end. Motivated by the recent introduction of decentralized prediction markets, we introduce a mechanism that allows for the outcome to be determined by the votes of a group of arbiters who may themselves hold stakes in the market. Despite the potential conflict of interest, we derive conditions under which we can incentivize arbiters to vote truthfully by using funds raised from market fees to implement a peer prediction mechanism. Finally, we investigate what parameter values could be used in a real-world implementation of our mechanism.

1 Introduction

Prediction markets are commonly used to elicit information about some future event. The market operates by allowing participants to buy and sell securities which pay off according to the outcome of the event, and participants with an informational edge are able to place profitable trades when the market price disagrees with their own forecast. Through this trading process, the market price can be construed as a consensus forecast of the underlying event probability. Prediction markets have proven effective at forecasting events in a variety of domains, including business and politics (Spann and Skiera 2003; Berg and Rietz 2006).

A key challenge in implementing and scaling prediction markets is the question of outcome determination (i.e., closing markets for events). Traditional prediction markets are centralized, in the sense that there exists a trusted center that creates markets, oversees transactions, and closes the market appropriately. The trusted center is a bottleneck for defining and closing markets, limiting the scope of what can be predicted. However, there has recently been a rise of interest in *decentralized* prediction markets, where any user may create a market for an event. The markets are closed by consensus among a group of arbiters rather than by a center.

A decentralized platform removes the requirement for a highly trusted center, but it also allows each arbiter to directly influence the outcome of the market, in much the same way that agents may deliberately manipulate an event due

to their own stake in the market; this is known as *outcome manipulation* (Shi, Conitzer, and Guo 2009; Berg and Rietz 2006; Chakraborty and Das 2016). Additionally, by allowing anyone to create a market, there is no longer any guarantee that all questions will be sensible, or even have a well-defined outcome. In this paper, we propose a specific prediction market mechanism with crowdsourced outcome determination that addresses several challenges faced by decentralized markets of this sort.

First is the issue of *outcome ambiguity*. At the time the market closes, it might be unreasonable to assign a binary value to the event outcome due to lack of clarity in the outcome. In a centralized market, it may be possible to postpone the market closing date to allow for rare cases of ambiguity, but it is not clear who should make such decisions in a decentralized marketplace. Therefore, it may be more fitting to allow outcomes to be non-binary, reflecting some level of disagreement. Outcomes in our mechanism are determined by the fraction of arbiters that report an event to have occurred. This also guarantees that every market is well-defined, by having traders explicitly trade on their expectations of the arbiter reports, not the actual event.

Second is the problem of *peer prediction*. For the credibility of the market, it is essential that arbiters are incentivized to truthfully report their opinion as to the realized outcome. If, for instance, we reward arbiters for agreeing with the majority opinion, then they are forced to anticipate the reports of other arbiters, not their independent opinion. We address this problem by making a technical change to an existing peer prediction mechanism, the 1/prior mechanism.

Third is the inherent *conflict of interest* that arises by combining prediction markets and peer prediction mechanisms. Even if arbiters can be incentivized to report truthfully in isolation, there is no way to prevent them also having a stake in the market. An arbiter holding securities that pay off in a particular event will be incentivized to report that the event has occurred, even if they do not genuinely believe it to be the case, as long as they have a chance of affecting the market outcome. We tackle this issue by charging a trading fee that is later used to pay the arbiters. We show that, as long as each agent is responsible for a limited fraction of trading, and questions are clear enough, realistic trading fees can fully subsidize truthful reporting on the part of the arbiters.

Related Work. This work is inspired by decentralized prediction markets based on crypto-currencies, including Truthcoin, Gnosis, and especially Augur (Peterson and Krug 2015). As in Augur, our mechanism consists of a prediction market stage and an arbitration stage, with trading fees from the market stage subsidizing the arbitration. The details of the mechanisms differ in both stages, however, and Augur includes additional layers of complexity such as a reputation system. While this complexity does provide additional protection against attack, Peterson and Krug do not obtain any theoretical guarantees or justification for their chosen parameters. Clark et al. (2014) also discuss outcome determination in crypto-based prediction markets, among several other implementation aspects.

Our work is most closely related to that of Chakraborty and Das (2016), who consider a model where two agents participate in a prediction market whose outcome is determined by a vote among the two agents. Our model extends theirs by allowing an arbitrary number of traders, and not requiring that all traders are arbiters. Further, we take a mechanism design approach, obtaining an incentive compatible mechanism, while Chakraborty and Das focus on analyzing the equilibrium of a simple (to play) trading-voting game, with no peer prediction mechanism in the voting phase to incentivize truthful voting. Recent work by Witkowski et al. (2017) also looks at a combination of forecasting and peer prediction, although the forecasts in their paper are elicited via proper scoring rules, rather than prediction markets.

The work of Bacon et al. (2012) is similar in spirit to ours, as is the literature on outcome manipulation mentioned previously, but in all cases the concrete setting is quite different. We also draw heavily on existing literature in prediction markets (Hanson 2003; Chen and Pennock 2007; Chen and Vaughan 2010) and peer prediction (Miller, Resnick, and Zeckhauser 2005; Prelec 2004; Witkowski and Parkes 2012; Chen and Pennock (2010) survey these topics.

2 Preliminaries

Let N be a set of agents and let $A \subset N$ be a small set of distinct and verifiable *arbiters*. Let $m = |A|$ denote the number of arbiters. The agents are anonymous in the sense that one cannot verify whether a trade is placed by an arbiter or non-arbiter, and whether several trades are placed by the same agent. Let X be a binary event with some realized outcome in $\{0, 1\}$. We are interested in setting up a prediction market for the outcome of X , with the resolution of the market decided upon by the arbiters.

Prediction markets. We consider prediction markets implemented via a *Market Scoring Rule (MSR)*, where the underlying scoring rule is strictly proper (Hanson 2003; Chen and Pennock 2007). Every strictly proper MSR can be implemented as a market maker based on a *convex cost function*. Under this market structure, agents trade shares of a security with a centralized market maker, who commits to quoting a buy and sell price for the security at any time. The security payout is contingent on the outcome of X . In the usual implementation, one share of the security pays out \$1

in the event that $X = 1$, and \$0 otherwise.

Let q denote the total number of outstanding shares owned by the agents (note that q can be negative, in the case that more shares have been sold than bought). The market maker charges trades according to a convex, differentiable, and monotonically increasing function C . An agent wishing to buy $q' - q$ securities pays $C(q') - C(q)$. Negative payments encode a transaction where securities are sold back to the market maker. The instantaneous price of the security is given by $p = \frac{dC}{dq}$. Because the market maker always commits to trading, it may run a loss once the outcome is realized and the securities pay out, but the loss is bounded.

In practice, the cost function is also tuned using a *liquidity parameter* b , via the transformation $C_b(q) \equiv bC(q/b)$. A higher setting of b results in lower price responsiveness, in the sense that the price will change less for a fixed dollar trading amount. It also results in a higher worst-case loss bound for the market maker. Unless otherwise stated, our results assume that each agent participates in the market only once. The mechanism and results extend to situations in which agents can participate more than once, and we highlight these extensions where relevant throughout the paper.

Peer prediction. Peer prediction mechanisms are designed to truthfully elicit private information from a pool of agents via a reward structure that takes advantage of information correlation across agents. After the realization of X , each arbiter i receives either a positive or negative signal x_i , which we denote by $x_i = 1$ and $x_i = 0$ respectively. Let μ be the prior probability that an agent receives a positive signal. Let μ_1 (resp. μ_0) be the probability that, given agent i receives a positive (resp. negative) signal, another randomly chosen agent also receives a positive signal.¹ In a standard peer prediction belief model, the quantities μ_1 and μ_0 can be calculated given μ and the signal beliefs $P(x_i = 1|X = 1)$ and $P(x_i = 1|X = 0)$; Witkowski (2014) provides an overview. Assuming common information is not always reasonable, but it is natural in our setting to assume that agents take the closing price of the prediction market as their prior (if not, then they can profit in expectation by participating in the market). The probabilities μ_1 and μ_0 are specific to the nature of the event X .

The peer prediction mechanism of interest in this work is the 1/prior (“one over prior”) mechanism (Witkowski 2014; Jurca and Faltings 2008; 2011). The 1/prior mechanism first asks each agent for their signal report \hat{x}_i . Then, every agent i is randomly paired with a peer agent $j \neq i$, and paid

$$u(\hat{x}_i, \hat{x}_j) = \begin{cases} k\mu & \text{if } \hat{x}_i = \hat{x}_j = 0 \\ k(1 - \mu) & \text{if } \hat{x}_i = \hat{x}_j = 1 \\ 0 & \text{if } \hat{x}_i \neq \hat{x}_j, \end{cases}$$

where k is some positive constant that can be freely adjusted to scale the payments received by the arbiters. Truthfully reporting $\hat{x}_i = x_i$ is an equilibrium provided that

¹Our analysis will assume that μ_1 and μ_0 are common across agents, but this is not a strict requirement. If we allow each agent to have distinct updates μ_1^i, μ_0^i , then we can let $\mu_1 = \min_i \mu_1^i$, corresponding to the minimum update given $\hat{x}_i = 1$, and similarly $\mu_0 = \max_i \mu_0^i$.

1. Market stage.
 - (a) A prediction market is set up for an event X using a market scoring rule.
 - (b) Agents trade in the market. For a security bought at price p , a trading fee of fp is charged, and for a security sold at price p , a fee of $f(1-p)$ is charged.
 - (c) The market closes, trading stops.
2. Arbitration stage.
 - (a) Each arbiter i receives a signal $x_i \in \{0, 1\}$ and reports an outcome $\hat{x}_i \in \{0, 1\}$.
 - (b) Each arbiter i is assigned a peer arbiter $j \neq i$ and paid according to the 1/prior with midpoint mechanism.
 - (c) The outcome of the market, and the payoff of each share sold, is set to the fraction of arbiters that report $\hat{x}_i = 1$.

Figure 1: Prediction market with outcome determined using peer prediction.

$\mu_1 \geq \mu \geq \mu_0$ (Frongillo and Witkowski 2016). This is a natural condition that we will assume throughout the paper—receiving signal $x_i = 1$ should not *decrease* i 's estimate that another agent j also receives signal $\hat{x}_j = 1$. We also assume that at least one of the inequalities is strict, so that $\mu_1 > \mu_0$; this condition is known as *stochastic relevance*. Via a simple adaptation of the corresponding proof for the 1/prior mechanism, it can be shown that truthful reporting remains an equilibrium if μ is replaced by some other constant c with $\mu_0 < c < \mu_1$ in the payments; we will exploit this fact to adapt the 1/prior mechanism for our purpose.

We call the quantity $\delta = \mu_1 - \mu_0$ the *update strength*. This quantity is specific to the instance and, roughly speaking, measures how strongly positively correlated the signals are across arbiters. The update strength is high if, after receiving a positive (negative) signal, an arbiter believes that another given arbiter receives a positive (negative) signal with high probability. For instance, if event X is “Will the Cleveland Cavaliers win the 2016 NBA playoffs?” then we would expect $\delta \approx 1$, since any arbiter reaching a conclusion about the outcome of the series (by watching it live, reading in the news, etc.) would strongly expect any other arbiter to reach the same conclusion. On the other hand, a question like “Will a Presidential candidate tell a lie in the televised debate?” is considerably more open to interpretation, and we would expect it to have a smaller value of δ . If an arbiter believes a candidate to have lied, it is not necessarily the case that another arbiter believes the same.

3 Mechanism

A step by step description of our mechanism is given in Figure 1. The mechanism runs a prediction market where the outcome is determined by a vote among arbiters. The arbiters' signals should be interpreted as the information they receive regarding the outcome of X : checking news sources,

observing events, their own opinions, etc. To ensure that arbiters truthfully report their information, we incentivize them via a peer prediction mechanism.² In both stages we implement non-standard versions of existing mechanisms, which we detail in the following.

Market stage We make use of an MSR with non-binary outcome. The outcome takes a value $\hat{X} \in [0, 1]$ corresponding to the fraction of arbiters that report $\hat{x}_i = 1$. Each share sold pays off \hat{X} . Observe that this fundamentally changes the value of a security to a market participant: in a standard prediction market, an agent's value for a security is his subjective probability that event X occurs, while in our market his value is the fraction of arbiters that he expects to report $\hat{x}_i = 1$. However, given the agent's valuation for a security, his incentives in both markets are similar. A risk-neutral, non-arbiter agent will trade shares until the market price matches the security's expected payoff, or the agent's budget is exhausted.

This change to the payoff structure has two advantages. First, it ensures that any question has a well-defined and unambiguous outcome, avoiding problems with badly worded questions. This is important in any situation where users are allowed to generate markets. Second, any market with a binary outcome that relies on arbitration must have a point of ‘discontinuity’, where a change in report from a single arbiter results in the value of a security changing by \$1.³ There will therefore always be situations where, given the reports of the other arbiters, a single arbiter completely controls the market outcome. If this arbiter also has a significant stake in the market, this creates a very large incentive problem. By utilizing non-binary outcomes, a single arbiter can only change the value of each security by at most $\$1/m$.

Our mechanism imposes trading fees. Theoretical models of prediction markets do not typically incorporate trading fees (an exception is the work of Othman et al. (2013), where a fee in the form of a bid-ask spread is used to allow liquidity to increase over time), but they are standard in real-world implementations. To understand how the fee is implemented, it is important to distinguish between transactions (buy or sell) where an agent increases its position (in terms of risk), and transactions where it liquidates its position. The trading fee that we implement can be seen as a fee on the worst-case loss incurred by an agent: the fee is on p when a new security is bought, and $1 - p$ when a security is sold short (because it may pay out \$1). However, no fee is levied when an agent sells back a share that it holds, or buys back a share that was previously sold short—these are liquidation transactions.

The trading fee serves two distinct purposes in our mechanism. First, it allows us to raise funds which can then be used to pay arbiters. Even assuming that arbiters behave honestly

²Each arbiter makes his report without knowledge of the report of any other arbiter; for instance, the reports could be made simultaneously.

³To see this, consider the case where all arbiters report $\hat{x}_i = 1$, and flip one report at a time to $\hat{x}_i = 0$. One of these flips must change the outcome from $\hat{X} = 1$ to $\hat{X} = 0$.

(in the absence of a sophisticated peer prediction mechanism), they still need to be compensated for the time spent looking up the outcome of X and reporting it to the mechanism. This can, in principle, be achieved by any of a number of fee structures.

Second, the fee provides natural bounds on the value of any given security. Even if an event is certain to occur, with a fee of $f = 2\%$ an agent who moves the market price to (say) 99¢ actually pays a marginal cost of $\$0.99 \cdot 1.02 > \1 (see Lemma 2 for an exact bound). The multiplicative fee effectively bounds the price of the security away from 0 and 1. Thus, it is impossible for an agent to buy securities at an arbitrarily cheap price, which allows us to bound the number of securities, and therefore maximum payout, of any agent with a fixed budget B . We note that there are other reasonable fee structures which do not provide such a lower bound on the price. For example, if the agents only pay a fee on any profit they gain from the market, then the price of an event that is certain to happen will still converge to 1.

Arbitration stage The main challenge in our setting is to incentivize arbiters to truthfully report their signal regarding the realized value of X . In the absence of any conflict of interest, this is a simple peer prediction problem. Since the closing price of the market gives us a natural common prior on the probability that a given arbiter receives signal $x_i = 1$, it is natural to use the 1/prior mechanism. For prior signal probability μ , the 1/prior mechanism uses the fact that $\mu_1 \geq \mu \geq \mu_0$ to guarantee that truthful reporting achieves higher payoff than misreporting. However, as μ_1 approaches μ , the payoff for truthfully reporting signal $\hat{x}_i = 1$ approaches the payoff for misreporting $\hat{x}_i = 0$. In isolation, there is still no reason to misreport, but if the arbiter has some stake in the market then it may be worthwhile to incur a small misreporting loss to achieve other gains. The following example illustrates this issue.

Example 1. Consider a prediction market for the event “Will the Democratic presidential candidate be leading the Republican presidential candidate in the polls at the end of the month?” Suppose it is known that 10% of arbiters only check conservative news sources, which always report that the Republican candidate is ahead, and another 10% only check liberal news sources, which always report the opposite. Suppose the market closes at $\mu = 0.89$. Consider an arbiter j who checks a (moderate) news source and finds that the Democratic candidate is ahead (i.e., $x_j = 1$). Since it is still the case that 10% of the arbiters will certainly receive signal $x_i = 0$, the updated belief μ_1 remains no higher than 0.9. That is, the update is very small, and the expected profit from reporting $\hat{x}_j = 1$ is also small. If j has bet against the outcome (i.e., sold some securities to the market maker), it could be in his interest to lie and report $\hat{x}_j = 0$.

However, if the moderate news site had reported that the Republican candidate was leading (i.e., $\hat{x}_j = 0$), the updated belief μ_0 could be quite small, even in the range of 0.1 (since most arbiters check moderate sources). Now j has a lot to gain from reporting $\hat{x}_j = 0$. Therefore, j would have to hold a relatively large number of shares for misreporting to outweigh the expected profit from the 1/prior mechanism.

Example 1 stems from an asymmetry in update strength, leading to potentially different incentives for arbiters depending on which signal they receive. We modify the mechanism, making the update strength symmetric. Given that we know the updated beliefs μ_1 and μ_0 , we can pay arbiters according to the 1/prior mechanism but use the value $(\mu_1 + \mu_0)/2$ instead of the prior, μ . We call this the 1/prior with midpoint mechanism. Using the midpoint guarantees that the incentives for arbiters are the same regardless of the signal they receive. For the arbiter with the greatest incentive to misreport, using the 1/prior with midpoint mechanism (weakly) decreases his incentive to misreport over the standard 1/prior mechanism, allowing us to achieve better bounds in our worst-case analysis.

Analysis

In this section, we derive conditions for truthful reporting ($\hat{x}_i = x_i$) to be a best response, given that all other arbiters report truthfully. The main restriction we require is an upper bound B on the total budget any given arbiter spends in the market—without such a bound, an arbiter could have an arbitrarily large incentive to manipulate the market’s outcome. Thus, B appears as a parameter in our analysis.

Arguably, an arbiter confident in their ability to manipulate a market outcome could procure enough funds as to have a very large budget, especially relative to a small market. However, in current decentralized prediction markets, each arbiter arbitrates only a small fraction of markets. As long as the assignment of arbiters to markets is done *after* the market closes, there is no way for manipulators to target a specific market. For this reason, we believe that manipulations are most likely to be of a form where arbiters participate honestly in the first stage, but, if they happen to be assigned to arbitrate a market that they also participated in, may be able to gain by not reporting truthfully, rather than arbiters mounting deliberate high-budget attacks in the market stage. Of course, our analysis is not specific to that particular interpretation, but we do consider it a compelling argument in favor of using a budget bound in our analysis.

Intuitively, we need to scale the payments made to arbiters in the arbitration stage by a sufficiently large k so that the increased payoff for truthful reporting in this stage overwhelms the gains from manipulating the outcome.

Lemma 1. Let n_i be the number of securities held by arbiter i . Then truthfully reporting $\hat{x}_i = x_i$ is a best response for arbiter i , given that all other arbiters report truthfully, if and only if

$$k \geq \frac{2|n_i|}{m\delta}.$$

Proof. We prove the case where $n_i > 0$; the case for $n_i < 0$ is symmetric. The total payoff for arbiter i is the sum of the payoffs from the market phase and the arbitration phase. Fixing the reports of the other arbiters, the market payout for i is higher when i reports $\hat{x}_i = 1$. And, in expectation, the payoff for i in the arbitration phase is higher for truthful reporting than for lying. Thus, the only problematic case is when $x_i = 0$, but i may wish to report $\hat{x}_i = 1$.

So suppose that $x_i = 0$. The expected payoff for truthfully reporting $\hat{x}_i = 0$, assuming all other arbiters truthfully report their signal, is

$$n_i \mu_0 \frac{m-1}{m} + (1 - \mu_0) k \frac{\mu_0 + \mu_1}{2}. \quad (1)$$

Here $\mu_0(m-1)$ is the expected number of arbiters that report signal 1, and therefore $n_i \mu_0(m-1)/m$ is i 's expected payoff from the market, while the remaining term is $1 - \mu_0$, the probability of i 's peer agent also reporting 0, multiplied by the payment i receives in this case. On the other hand, the expected payoff for misreporting $\hat{x}_i = 1$ is

$$n_i \left(\mu_0 \frac{m-1}{m} + \frac{1}{m} \right) + \mu_0 k \left(1 - \frac{\mu_0 + \mu_1}{2} \right), \quad (2)$$

where the extra $1/m$ in the first term is due to the additional market payoff from i reporting $\hat{x}_i = 1$, and the latter term is now the probability of i 's peer agent reporting 1, multiplied by the payoff i receives when this happens.

We require that the expected payoff for reporting $\hat{x}_i = 1$ be at most the expected payoff for truthfully reporting $\hat{x}_i = 0$. Setting term (2) to be at most term (1) and simplifying yields the result. \square

This characterization requires an upper bound on the number of securities that any single agent owns. In itself this is an unsatisfying restriction; however, we can think about it in terms of the size of the fee, f , and the amount of money that any single arbiter spends in the market, B . For fixed fee f , let q^- and q^+ be the number of outstanding securities such that the market prices are $p(q^-) = f/(1+f)$ and $p(q^+) = 1/(1+f)$ respectively. Note that these quantities depend on the liquidity parameter b used in the cost function.

Lemma 2. *For fixed percentage fee f , the number of outstanding securities lies in the interval $[q^-, q^+]$.*

Proof. Suppose that some agent sells a security when there are already q^- outstanding. Then the marginal price is exactly $f/(1+f)$. When selling a security at this price, the agent receives $f/(1+f)$ from the mechanism but must pay a trading fee of

$$f \left(1 - \frac{f}{1+f} \right) = \frac{f}{1+f}.$$

Thus the agent's net revenue from the sale is 0 (and the possibility remains that he must pay the mechanism in the event that X occurs). Therefore no agent makes such a sale, and the number of outstanding securities never drops below q^- .

A similar argument shows that q never exceeds q^+ . To buy a security when there are already q^+ outstanding, an agent must pay a price of at least \$1, when the fee is included. \square

Lemma 2 provides us with the minimum and maximum number of outstanding securities at any time. As a corollary, we can derive the maximum number of securities that a single agent with budget B is able to purchase or short sell. We interpret the budget as an upper bound on the worst-case loss that the agent is able to incur. When buying a security for price p , the worst-case loss is p , under outcome $X = 0$.

When selling a security for price p , the worst-case loss is $1 - p$, under outcome $X = 1$. Let $\phi_b^-(B) = C_b^{-1}(B + C_b(q^+)) - q^+$ and $\phi_b^+(B) = C_b^{-1}(B + C_b(q^-)) - q^-$.

Corollary 1. *At the end of the market stage, an agent i with budget B holds $n_i \in [\phi_b^-(B), \phi_b^+(B)]$ securities.*

The proof, along with all other omitted proofs, can be found in the full version of the paper. An interesting special case is the limit as $b \rightarrow \infty$. This corresponds to the market having zero price responsiveness, meaning that all securities are purchased at a fixed price. Conceptually, it is equivalent to the situation where agents participate in the market more than once. In that setting, an agent could wait until the market price reaches $\frac{f}{1+f}$, buy a small number of securities, then wait again until the price drops. An agent spending all their budget in this way can, in the extreme case, buy as if the market has infinite liquidity.

Corollary 2. *For an agent that spends at most B dollars in a market with trading fee f and infinite liquidity, n_i lies in the range $\left[-\frac{B(1+f)}{f}, \frac{B(1+f)}{f} \right]$.*

If every agent has budget at most B in the market stage, we can combine the bounds from Corollaries 1 and 2 and Lemma 1 to determine the minimum payment that guarantees truthful reporting in the arbitration phase.

Theorem 1. *Given that all other arbiters report truthfully, truthful reporting is a best response for arbiter i if*

$$k \geq \frac{2 \max\{|\phi_b^-(B)|, |\phi_b^+(B)|\}}{m\delta}.$$

In the case that agents may participate in the market many times, truthful reporting requires that

$$k \geq \frac{2B(1+f)}{f m \delta}.$$

Therefore, fixing an agent budget B and a trading fee f , we know how large one needs to make the payments in the arbitration phase in order to incentivize truthful reporting. We now take a global view, and examine the total funds required to incentivize all arbiters to report truthfully.

Lemma 3. *The total payment made to the arbiters is at most mk . We can implement a truthful equilibrium with total payment at most*

$$\frac{2 \max\{|\phi_b^-(B)|, |\phi_b^+(B)|\}}{\delta}.$$

In the case that agents may participate in the market many times, we require total payment at most

$$\frac{2B(1+f)}{f\delta}.$$

Proof. As $0 \leq \mu_0, \mu_1 \leq 1$, their mean also lies between 0 and 1, and therefore each arbiter's payment in the 1/prior with midpoint mechanism is at most k . Thus the total payment to the arbiters is at most mk , which proves the first part. Combining this with the bounds on k from Theorem 1 yields the second part. \square

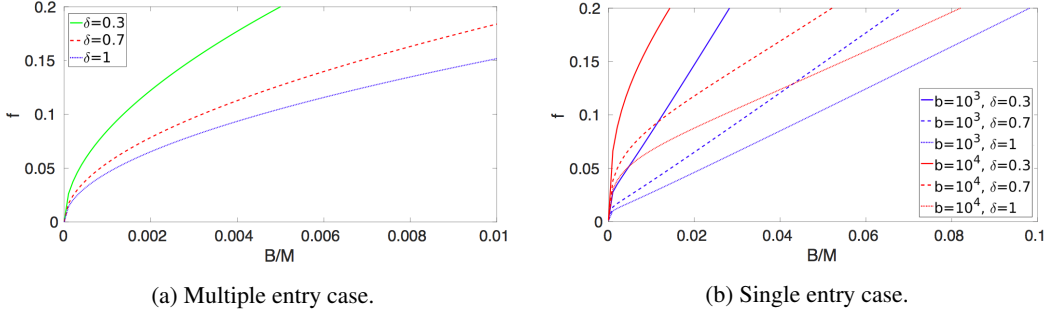


Figure 2: Minimum fee f required to adequately incentivize arbiters, plotted as a function of $\frac{B}{M}$. In both cases, $M = 10^6$ is fixed. Relationships are shown for selected values of update strength δ and, in the right-hand plot, liquidity b .

Now that we have an expression for the total amount needed to pay the arbiters, we can determine a suitable value for the fee f so that the mechanism does not need any outside subsidy to finance these payments. Let c_i denote the total cost paid by agent i to the mechanism (so c_i is negative if agent i sells securities). Define M by

$$M = \sum_{i:n_i>0} c_i + \sum_{i:n_i<0} (n_i + c_i).$$

M can be interpreted as the sum of the worst-case losses of the agents. By definition, the total fee revenue collected by the mechanism is fM . The mechanism is guaranteed to generate enough fees to incentivize truthful reporting if the revenue is at least as large as the total payment required for the arbiters. We state this result as a theorem.

Theorem 2. *The mechanism generates enough fee revenue to pay the arbiters without requiring any outside subsidy if*

$$fM \geq \frac{2 \max\{|\phi_b^-(B)|, |\phi_b^+(B)|\}}{\delta}. \quad (3)$$

If agents may participate in the market many times, then we require that

$$fM \geq \frac{2B(1+f)}{f\delta}. \quad (4)$$

Observe that inequality (4) aligns with intuition. An increase in total trader spend M , or the trading fee f , makes it easier to incentivize the arbiters to report truthfully since the market collects more revenue. Likewise, an increase in δ helps us satisfy the inequality, since a large update strength increases the incentive for arbiters to report truthfully to the peer prediction mechanism. However, a large value of B increases the stake that any single arbiter can have in the market, which in turn increases their payoff for misreporting.

An interesting feature of inequalities (3) and (4) is the lack of any dependence on the number of arbiters m . One might expect that increasing the number of arbiters would be beneficial, since this reduces the influence that any one of them has on the market outcome. However, this is canceled out by the fact that as we add arbiters, the payment per arbiter decreases, so that we cannot incentivize them as strongly.

4 Parameter Calibration

In this section we investigate the constraints imposed by inequalities (3) and (4). The purpose of the exercise is to illustrate how Theorem 2 can be used to inform the choice of fee f , and to confirm that realistic fees could be charged in practice to subsidize truthful arbitration. We consider the logarithmic market scoring rule (LMSR), which is the most common MSR used in practice. For the LMSR, the cost and price functions are

$$C_b(q) = b \log(1 + e^{q/b}), \quad p(q) = \frac{e^{q/b}}{1 + e^{q/b}}.$$

By the symmetry of LMSR, $q^- = -q^+$ and $\phi_b^-(B) = -\phi_b^+(B)$. We will therefore solve for $\phi_b^+(B)$. To find q^- , we set $p(q) = f/(1+f)$ and solve for q , which gives $q^- = b \log f$. Now, substituting the relevant components into the expression $\phi_b^+(B) = C_b^{-1}(B + C_b(q^-)) - q^-$ leads to the following expression for inequality (3):

$$fM \geq \frac{2b(\log((1+f)e^{B/b} - 1) - \log f)}{\delta}. \quad (5)$$

In the case where we allow agents to participate multiple times, inequality (4) remains unchanged.

We plot (3) and (4) in Figure 2, considering their tight versions as equalities. First consider Figure 2a, which represents the worst-case scenario in which agents can enter multiple times and potentially spend their entire budget buying securities at minimum price p^- . Suppose that some entity is creating a prediction market for event X . Having decided on a question, the main decision is what value to set for f , typically in the 2-5% range. To do so, the market creator needs to first estimate a value for δ , which will be determined by question clarity, whether the arbiters have reliable sources to check the outcome, and other such factors. Each line in the graph represents a specific value of δ . With δ fixed, the market creator can estimate a value for $\frac{B}{M}$. This is the maximum proportion of money that any single arbiter will contribute to the market. We would expect $\frac{B}{M}$ to be small for markets that generate a lot of interest, while niche markets would be vulnerable to having a single agent contribute a large percentage of the total trade. Given these values, the creator can arrive at the smallest f that is guaranteed to subsidize truthful reporting. From the graph, we see that in the case of a

question where $\delta = 1$ and $\frac{B}{M} = 0.001$, we can subsidize the arbiter payment with a fee of approximately 4%. This may seem large for a clear question with high participation, but we stress that this fee is based on a severe worst case where an agent is able to spend its entire budget purchasing securities at the minimum price.

Now consider Figure 2b, which returns to the case where an agent only enters once, where liquidity now plays a role and we have to consider different values for parameter b . Figure 2b includes two reasonable values for b , as well as three different values for δ . We note that the situation looks considerably better for the market creator; indeed, the horizontal axis is now ten times larger indicating that we can now handle much smaller markets. When $\delta = 1$, we can handle situations where a single agent can contribute as much as 2% of the total trade with a fee of less than 5%. Even for questions with δ as low as 0.3, in a market with $b = 1000$ and $\frac{B}{M} = 0.005$ the fee can be set to approximately 5%.

5 Conclusion

This paper proposed and analyzed a mechanism where the outcome of an MSR prediction market is determined via a peer prediction mechanism among a set of arbiters. The mechanism relies on two key adaptations to incentivize truthful arbitration: market shares pay out according to the proportion of arbiters who vote affirmatively, instead of a binary payout, and peer prediction payments are based on the midpoint of the two possible posteriors, rather than the prior. We showed that, with this combination of adaptations, it is possible to charge a trading fee that fully subsidizes truthful arbitration. Calibration based on plausible values of question clarity and trading volume suggests that realistic fees of 5% should be sufficient in practice.

While we have derived conditions under which truthful reporting is an equilibrium, there remains the possibility of the arbiters reporting according to uninformative equilibria that achieve higher payoff. This problem has recently been addressed in the peer prediction literature in situations where reporters complete several tasks instead of just one (Dasgupta and Ghosh 2013; Shnayder et al. 2016); it may be worthwhile to apply these techniques to our setting. In practice, arbiters vote on many questions across time, which opens the possibility of using a reputation system to incentivize them to vote truthfully and accurately (Peterson and Krug 2015). The interplay of the incentives from all these mechanisms is fertile ground for future research.

Acknowledgments

This work was done in part while Freeman was at Microsoft Research, New York City. He also thanks NSF IIS-1527434 and ARO W911NF-12-1-0550 for support.

References

Bacon, D. F.; Parkes, D. C.; Chen, Y.; Rao, M.; Kash, I.; and Sridharan, M. 2012. Predicting your own effort. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 695–702.

Berg, J. E., and Rietz, T. A. 2006. The Iowa Electronic Markets: Stylized facts and open issues. *Information Markets: A New Way of Making Decisions* 142–169.

Chakraborty, M., and Das, S. 2016. Trading on a rigged game: Outcome manipulation in prediction markets. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*.

Chen, Y., and Pennock, D. 2007. A utility framework for bounded-loss market makers. In *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence (UAI)*.

Chen, Y., and Pennock, D. M. 2010. Designing markets for prediction. *AI Magazine* 31(4):42–52.

Chen, Y., and Vaughan, J. W. 2010. A new understanding of prediction markets via no-regret learning. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, 189–198. ACM.

Clark, J.; Bonneau, J.; Felten, E. W.; Kroll, J. A.; Miller, A.; and Narayanan, A. 2014. On decentralizing prediction markets and order books. In *Workshop on the Economics of Information Security, State College, Pennsylvania*.

Dasgupta, A., and Ghosh, A. 2013. Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd International Conference on World Wide Web*, 319–330. ACM.

Frongillo, R., and Witkowski, J. 2016. A geometric method to construct minimal peer prediction mechanisms. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*.

Hanson, R. 2003. Combinatorial information market design. *Information Systems Frontiers* 5(1):107–119.

Jurca, R., and Faltings, B. 2008. Incentives for expressing opinions in online polls. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, 119–128. ACM.

Jurca, R., and Faltings, B. 2011. Incentives for answering hypothetical questions. In *Workshop on Social Computing and User Generated Content, EC-11*.

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

Othman, A.; Pennock, D. M.; Reeves, D. M.; and Sandholm, T. 2013. A practical liquidity-sensitive automated market maker. *ACM Transactions on Economics and Computation* 1(3):14.

Peterson, J., and Krug, J. 2015. Augur: A decentralized, open-source platform for prediction markets. *arXiv preprint arXiv:1501.01042*.

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

Shi, P.; Conitzer, V.; and Guo, M. 2009. Prediction mechanisms that do not incentivize undesirable actions. In *International Workshop on Internet and Network Economics (WINE)*, 89–100. Springer.

Shnayder, V.; Agarwal, A.; Frongillo, R.; and Parkes, D. C. 2016. Informed truthfulness in multi-task peer prediction. *arXiv preprint arXiv:1603.03151*.

Spann, M., and Skiera, B. 2003. Internet-based virtual stock markets for business forecasting. *Management Science* 49(10):1310–1326.

Witkowski, J., and Parkes, D. C. 2012. A robust Bayesian truth serum for small populations. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI)*.

Witkowski, J.; Atanasov, P.; Ungar, L. H.; and Krause, A. 2017. Proper proxy scoring rules. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*. Forthcoming.

Witkowski, J. 2014. *Robust Peer Prediction Mechanisms*. Ph.D. Dissertation, Department of Computer Science, Albert-Ludwigs-Universität Freiburg.