

Trading on a Rigged Game: Outcome Manipulation In Prediction Markets

Mithun Chakraborty, Sanmay Das
Washington University in St. Louis
{mithunchakraborty,sanmay}@wustl.edu

Abstract

Prediction markets are popular mechanisms for aggregating information about a future event. In situations where market participants may significantly influence the outcome, running the prediction market could change the incentives of participants in the process that creates the outcome. We propose a new game-theoretic model that captures two aspects of real-world prediction markets: (1) agents directly affect the outcome the market is predicting, (2) some outcome-deciders may not participate in the market. We show that this game has two types of equilibria: When some outcome-deciders are unlikely to participate in the market, equilibrium prices reveal expected market outcomes conditional on participants' private information, whereas when all outcome-deciders are likely to participate, equilibria are collusive – agents effectively coordinate in an uninformative and untruthful way.

1 Introduction

Prediction markets are widely used tools for aggregating and disseminating private information on some future event, dispersed among a potentially diverse crowd. However, attention is seldom paid in the literature to the possibility that market participants might have some degree of control on the outcome of the forecast event, and hence the presence of a prediction market may make agents affecting the outcome act differently than they otherwise would. In fact, sometimes it is this very power to affect outcomes that gives agents the informational edge that such markets get their value from.

Consider three real-world examples where prediction (or betting) markets have demonstrated their forecasting ability: politics [Berg *et al.*, 2008], sporting events [Wolfers and Zitzewitz, 2004], and software product releases [Cowgill and Zitzewitz, 2015]. In each of these cases, it is easy to see how a prediction market may distort outcomes: A congressional staffer or member of congress may know more about the probable result of a key vote than the general public, but she is also in a position to influence this result. A referee or player has substantial ability to influence the outcome of a sporting event. A software engineer has the potential to delay (or speed up) the release of a product. When such outcome

manipulation incentives are present, it is natural to ask two questions: (1) Are the actions of the outcome-deciders still *truthful*, i.e. do they take the same actions that they would in the absence of the prediction market? (2) Are market prices still *informative*, i.e. how much do they still tell us about the realized outcome?

While it is acknowledged that prediction markets have value as aids in making business and policy decisions, they have gone through cycles of hype and bust for reasons that include regulatory concerns about manipulation. The emblematic anecdote about this problem is the failure of DARPA's proposed Policy Analysis Markets which were caricatured in the media as "terrorism futures" [Hanson, 2007b; Stiglitz, July 31 2003]. There are obviously markets that will not work but stock and futures markets have been used for a long time as forecasting tools, and prediction markets are similar in essence. The key is to understand when markets may be prone to manipulation and how much to trust them.

1.1 Contributions

A model for outcome manipulation: We propose a new model for studying manipulative behavior that captures two aspects of real-world prediction markets: (1) agents directly affect the forecast event, (2) some of the agents who have influence over the outcome may not participate in the prediction market (e.g. employees who have an impact on the outcome of a product launch typically would not all take part in the company's in-house prediction market for its release date). In markets where an individual has a small effect on the outcome, agents' incentives for manipulation are likely to be weak. With this in mind, we mainly focus on a two-stage game-theoretic model of a market with two players, Alice and Bob, who affect the outcome and can also trade on it (Sections 2 and 3), and then discuss how our insights extend to models with more outcome-deciders in Section 4.

In the first stage of the game, Alice and Bob each receive a private signal about the underlying entity, and then they have the opportunity to participate sequentially, once each, in a prediction market mediated by a *market scoring rule* or MSR [Hanson, 2007a]. However, depending on his *type*, there is a probability that Bob may not participate in trading. In the second stage, the two players simultaneously take actions which we term "votes" for convenience, although in general they

model each participant’s role in determining the outcome.¹ The payoffs from the first-stage prediction market are determined by a simple function of the stage two votes. If Bob has not traded, his vote is consistent with his private signal, otherwise he is strategic; the first mover Alice is always strategic.

Our model directly captures the experiments of Chakraborty *et al.* [2013] where prediction markets with student participants were used to forecast the fraction of “up” (vs “down”) ratings given by students to course instructors. Moreover, *Augur*, a recently launched “decentralized, open-source platform for prediction markets” [Peterson and Krug, 2015], is a real-world mechanism with manipulation incentives similar to those in our model: A consensus, computed from votes cast by participants called “reporters”, serves as a proxy for the payoff-deciding ground truth of a market on which these reporters can also wager.

A dichotomy of equilibria: Our main result is that the equilibria of our game-theoretic model can be cleanly categorized into two types, depending on Bob’s probability of participation in the trading stage. Below a threshold on this probability, say \tilde{p} (a function of the MSR used and the signal structure), we call the equilibrium a low participation probability equilibrium (LPPE), and above \tilde{p} , we call it a high participation probability equilibrium (HPPE). In a LPPE, Alice essentially predicts Bob’s vote, and then bases her trading on the optimal combination of her own and Bob’s votes, and the prediction market price is reflective of the expected outcome. In a HPPE, on the contrary, Alice effectively expects Bob to enter and collude with her, and she chooses a prediction market price that allows Bob and her to “split the profit”. We summarize the qualitative implications of our result in Section 3.1.

1.2 Related work

The market microstructure we use here is a *market scoring rule* (MSR), introduced by Hanson [2003] and expanded upon by many researchers ([Chen and Pennock, 2007; Chen *et al.*, 2009; Gao *et al.*, 2013] etc.). Incentives for manipulation in prediction markets may arise in a number of ways. There is a plethora of literature on *price manipulation* – tampering with the market price owing to some extra-market incentives ([Hanson *et al.*, 2006; Rhode and Strumpf, 2006; Hanson and Oprea, 2009; Boutilier, 2012; Dimitrov and Sami, 2010; Chen *et al.*, 2011a; Huang and Shoham, 2014]), e.g. a politically motivated manipulator might make a large investment in an election prediction market to make one of the candidates appear stronger [Rothschild and Sethi, 2015]; a related body of work pertains to *decision markets* – a collection of contingent markets set up to predict the outcomes of different decisions such that only markets contingent on decisions that are taken pay off, the rest being voided [Hanson, 1999; Othman and Sandholm, 2010; Chen *et al.*, 2011b]. The type of manipulation that we are concerned with is *outcome manipulation* where an agent can take an action that partially

¹For example, for a product release date prediction market, a (binary) private signal could stand for whether an agent knows / believes she is capable of contributing her share in making sure that the launch is on time; her (binary) vote in this case would indicate whether she actually plays her part honestly.

influences the outcome to be predicted (e.g. [Berg and Rietz, 2006; Shi *et al.*, 2009]), and we study it in a new model.

The second major body of related literature comes from theoretical finance and market microstructure. Kyle [1985], followed by Holden and Subrahmanyam [1992], Foster and Viswanathan [1996] etc., studied the effect of informed “insider(s)” on market price; Glosten and Milgrom [1985] presented another view of how asymmetric information affects price formation, and their model has been adapted for market making in prediction markets [Das, 2008; Brahma *et al.*, 2012]. Ostrovsky [2012] recently examined information aggregation with differentially informed traders under both Kyle’s pricing model and market scoring rules. Again, in all of these models, the market outcome is assumed to be exogenously determined, unlike in ours.

2 Model and definitions

Let $\tau \in T$ denote the unobservable “true value” of the random variable on which the voting system and its associated prediction market are predicated. At $t = 0$, the two agents, Alice and Bob (A and B in subscripts), receive private signals $s_A, s_B \in \Omega = \{0, 1\}$ respectively. The *signal structure*, comprising the prior distribution $\Pr(\tau)$ on the true value and the conditional joint distribution $\Pr(s_A, s_B|\tau)$ of the private signals given the true value, is common knowledge.

Let $q_0(\cdot)$ denote Alice’s posterior probability that Bob received the signal $s_B = 0$, given her own signal and common knowledge, i.e. $q_0(s) \triangleq \Pr(s_B = 0|s_A = s) \forall s \in \{0, 1\}$. We ignore the uninteresting special cases $q_0 \in \{0, 1\}$ corresponding to Alice having no uncertainty about her peer’s private signal. Although we need no further assumptions on the signal structure for our main result (Theorem 1), it is worthwhile to define here the property of *stochastic relevance* [Miller *et al.*, 2005] which is a necessary assumption for one of our important corollaries.

Definition 1. For binary random variables $s_i, s_j \in \{0, 1\}$, s_j is said to be *stochastically relevant* for s_i if and only if the posterior distribution of s_i given s_j is different for different realizations of s_j , i.e. if and only if $\Pr(s_i = 0|s_j = 0) \neq \Pr(s_i = 0|s_j = 1)$.

We now describe the rules of the two-stage game comprising the market and voting mechanisms. We will call this the *trading-voting game*.

Stage 1 (market stage): The market price at any time-step t is public, the starting price at $t = 0$ being p_0 which is the market designer’s baseline estimate of the outcome (equal to the final gross payoff per unit of the security).

The prediction market is implemented using a market scoring rule (MSR) [Hanson, 2007a] with the underlying *strictly proper* scoring rule $s(r, \omega)$, where ω is the outcome to be forecast – in our model, determined by **Stage 2** (see below) – and r is an agent’s forecast / report on it; strict propriety, by definition, implies that if a forecaster is promised an ex post compensation of $s(r, \omega)$, then the only way she can maximize her subjective expected compensation is by reporting her expectation of the random variable ω . For a clean analysis, we shall focus on strictly proper rules for eliciting the expectation

of the random variable $\omega \in [0, 1]$, satisfying some regularity and smoothness conditions [Gneiting and Raftery, 2007; Abernethy and Frongillo, 2012]:

$$s(r, \omega) = f(r) + f'(r)(\omega - r), \quad \omega, r \in [0, 1] \quad (1)$$

where $f(\cdot)$ is a continuous, finite, strictly convex function on $[0, 1]$; its first derivative $f'(\cdot)$ is continuous, monotonically increasing, and finite on $[0, 1]$ except possibly that $f'(0) = -\infty$ or $f'(1) = \infty$, with $\lim_{r \rightarrow 0} f'(r)r = \lim_{r \rightarrow 1} f'(r)(1 - r) = 0$; the second derivative $f''(\cdot)$ is positive on $[0, 1]$ and finite in $(0, 1)$. Additionally, we need the function to have the following symmetry:

$$f\left(\frac{1+y}{2}\right) - f\left(\frac{1-y}{2}\right) = yf'\left(\frac{1}{2}\right) \quad \forall y \in [0, 1]. \quad (2)$$

Henceforth, we shall refer to (market) scoring rules possessing all the above properties as *symmetric well-behaved* (market) scoring rules. This covers a large family of MSRs including three of the most widely used and studied – logarithmic (LMSR), quadratic (QMSR), and spherical (SMSR) – defined in terms of their respective convex functions as:

$$\text{LMSR: } f(r) = r \ln r + (1 - r) \ln(1 - r),$$

$$\text{QMSR: } f(r) = r^2,$$

$$\text{SMSR: } f(r) = \sqrt{r^2 + (1 - r)^2}.$$

At $t = 1$, Alice interacts with the market maker and changes the price to p_A . At $t = 2$, Bob has an opportunity to trade but may not show up with a (common-knowledge) probability $\pi \in [0, 1]$ called Bob’s *non-participation probability*; if he does trade, he changes the price to p_B . Regardless of whether Bob trades, the market terminates after $t = 2$.

Stage 2 (voting stage): In this stage, Alice and Bob simultaneously declare their “votes” $v_A, v_B \in \Omega = \{0, 1\}$ respectively. Taking part in **Stage 2** is mandatory for both agents. We define *truthful voting* as declaring one’s private signal as one’s vote, *i.e.* $v_k = s_k$, $k \in \{A, B\}$. We assume that, if Bob did not trade in **Stage 1**, he votes truthfully, and we call such a Bob *HONEST*. Any agent participating in the prediction market is Bayesian, strategic, and risk-neutral. Hence, if Bob trades, we refer to him as *STRATEGIC Bob*.

The market outcome is given by the average vote $v = (v_A + v_B) / 2 \in \{0, \frac{1}{2}, 1\}$. Alice and Bob’s ex post net payoffs of are respectively

$$\mathcal{R}_i(p_i, p_j, v_A, v_B) = s\left(p_i, \frac{v_A + v_B}{2}\right) - s\left(p_j, \frac{v_A + v_B}{2}\right), \quad (3)$$

where $i \in \{A, B\}$, $j = 0$ for $i = A$, $j = A$ for $i = B$.

Bob does not strategically decide whether to take part in the market, it is determined extraneously – the proclivity to trade can be viewed as one of the components of Bob’s *type*, the other being his private signal s_B . Here, *HONEST Bob* models agents whose outcome-deciding actions remain unaltered by the introduction of the prediction market.

3 Equilibrium analysis

The solution concept we use for the trading-voting game described in Section 2 is the *perfect Bayesian equilibrium* (PBE), a refinement of Nash equilibria for Bayesian games [Fudenberg and Tirole, 1991]. In this section, we present the

main theorem, its supporting lemmas and main corollaries, and defer full proofs to a longer version of the paper.

Our first result formalizes the intuition that, if Alice pulls the market price down (resp. up) from its initial value, she is “forecasting” that the final outcome will be lower (resp. higher) than the market’s initial estimate and hence it is in her best interest to do everything in her power to ensure a low (resp. high) average vote – this is because her payoff is higher for a prediction closer to the realized outcome.

Lemma 1. *For the trading-voting game described in Section 2, if the prediction market has a starting price $p_0 \in (0, 1)$ at $t = 0$, and (p_A, v_A) denotes Alice’s combined action in the two-stage game, *i.e.* her report-vote pair, then*

- for any $p_A < p_0$, $(p_A, 0)$ strictly dominates $(p_A, 1)$;
- for any $p_A > p_0$, $(p_A, 1)$ strictly dominates $(p_A, 0)$; and
- she is indifferent between the actions $(p_0, 0)$ and $(p_0, 1)$.

This result holds regardless of Bob’s report-vote pair (p_B, v_B) .

Lemma 1 implies that immediately after Alice has traded, anyone can infer that $v_A = 0$ deterministically if $p_A < p_0$, $v_A = 1$ deterministically if $p_A > p_0$, and that she is equally likely to vote 0 or 1 if $p_A = p_0$, which is equivalent to not trading with the market maker. As soon as *STRATEGIC Bob* arrives to trade, he acquires all the information relevant to his decision making procedure that the rules of the game allow him to have $(p_A$ and $v_A)$. Thus, *STRATEGIC Bob* makes his trading and voting decisions (p_B, v_B) simultaneously.

The next step is to determine Bob’s best response to Alice’s **Stage 1** action. For this, we need to define quantities that we call the lower and upper thresholds p^L and p^H :

$$p^L \triangleq (f')^{-1}(2(f(1/2) - f(0))) \in (0, 1/2),$$

$$p^H \triangleq (f')^{-1}(2(f(1) - f(1/2))) \in (1/2, 1).$$

For any symmetric, well-behaved scoring rule, we can show that $p^H = 1 - p^L$. For LMSR, QMSR, and SMSR, p^L is 0.2, 0.25, and 0.2725 respectively.

Lemma 2. *For the trading-voting game described in Section 2, where the market is implemented with a symmetric well-behaved market scoring rule with lower and upper thresholds p^L, p^H , and has a starting price $p_0 \in (0, 1)$,*

- if $p_A < p_0$, then *STRATEGIC Bob’s best-response vote is $v_B = 1$ (resp. $v_B = 0$) if $p_A < p^L$ (resp. $p_A > p^L$) but he is indifferent between the two possible voting choices if $p_A = p^L$; his accompanying price-report is $p_B = \frac{v_B}{2}$;*
- if $p_A > p_0$, then *STRATEGIC Bob’s best-response vote is $v_B = 1$ (resp. $v_B = 0$) if $p_A < p^H$ (resp. $p_A > p^H$) but he is indifferent between the two possible voting choices if $p_A = p^H$; his accompanying p_B is $\frac{1+v_B}{2}$;*
- if $p_A = p_0$, then *STRATEGIC Bob’s best-response vote is $v_B = 0$ (resp. $v_B = 1$) if $p_0 > 1/2$ (resp. $p_0 < 1/2$) but he is indifferent if $p_0 = 1/2$, and his accompanying price-report is $p_B = \frac{1+v_B}{2}$.*

This result is independent of Bob’s private signal s_B .

p^L , p^H , and p_0 are points of transition in agent behavior: Bob’s best response is to “disagree with” Alice’s voting choice (revealed through p_A) in **Stage 2** if Alice’s price-report lies in either of the “outer” sub-intervals $[0, \min(p_0, p^L))$ or $(\max(p_0, p^H), 1]$, and to “agree with” her otherwise; for $p_A \neq p_0$, STRATEGIC Bob knows what the market outcome v is going to be, so he can make a perfect forecast $p_B = v$.

This fully characterizes how the game unfolds *after* Alice has taken her **Stage 1** action. Now, the final step towards completing the equilibrium specification is to figure out Alice’s best-response price-report p_A , based on q_0 , π , and her knowledge of Lemmas 1 and 2. In the extreme case $\pi = 0$, when Bob’s (strategic) participation is certain, it is natural to conjecture that Alice as the first mover will invite Bob to create a “fake world” with little connection to their private signals; at the other end of the spectrum ($\pi = 1$) where Alice is sure that Bob is HONEST, we expect her action to shed some light on her posterior belief about Bob’s signal. Theorem 1 tells us that there exists some critical value of Bob’s non-participation probability at which a switch between partially revealing and collusive equilibria occur.²

Theorem 1. *For any value of Bob’s non-participation probability $\pi \in (0, 1)$ and Alice’s posterior belief $q_0 \in (0, 1)$, the trading-voting game described in Section 2 has a perfect Bayesian equilibrium with the following attributes: For every q_0 , there exists a fixed value of Bob’s non-participation probability, say $\pi_c(q_0)$, which we call the “crossover” probability (dependent on the MSR), on either side of which the equilibria are qualitatively different. We call the sub-interval $\pi < \pi_c$ the high participation probability (HPP) equilibrium domain, and the sub-interval $\pi > \pi_c$ the low participation probability (LPP) equilibrium domain.*

• **In a HPP equilibrium:**

– In **Stage 1**, Alice moves the market price to $p_A = p^L$ if $q_0 > \frac{1}{2}$, and to $p_A = p^H$ if $q_0 < \frac{1}{2}$; STRATEGIC Bob’s price-update is $p_B = 0$ if $p_A = p^L$, and $p_B = 1$ if $p_A = p^H$.

– In **Stage 2**, Alice votes $v_A = 0$ if she set $p_A = p^L$, $v_A = 1$ if $p_A = p^H$; STRATEGIC Bob votes $v_B = 0$ if he set $p_B = 0$, and $v_B = 1$ if he set $p_B = 1$.

• **In a LPP equilibrium:**

– In **Stage 1**, Alice’s price-report p_A^{LPP} is equal to her posterior expectation of the market liquidation value (average vote) given the parameters π, q_0 and her report p_A^{LPP} , i.e. $p_A^{LPP} = \mathbb{E}[v|\pi, q_0, p_A = p_A^{LPP}]$. Moreover, $p_A^{LPP} < \frac{1}{2}$ if $q_0 > \frac{1}{2}$, $p_A^{LPP} > \frac{1}{2}$ if $q_0 < \frac{1}{2}$. STRATEGIC Bob’s price-update is $p_B = 0$ if $p^L \leq p_A \leq \frac{1}{2}$, $p_B = 1$ if $\frac{1}{2} < p_A \leq p^H$, and $p_B = \frac{1}{2}$ otherwise.

– In **Stage 2**, Alice votes $v_A = 0$ if $p_A > \frac{1}{2}$, $v_A = 1$ if $p_A < \frac{1}{2}$; STRATEGIC Bob votes $v_B = 0$ if $p_A \in [p^L, \frac{1}{2}] \cup (p^H, 1]$, $v_B = 1$ otherwise.

More specifically, p_A^{LPP} is one of $\mu_{0,0} = \frac{\pi(1-q_0)}{2}$, $\mu_{0,1} = \frac{1-\pi q_0}{2}$, $\mu_{1,0} = \frac{1+\pi(1-q_0)}{2}$, $\mu_{1,1} = 1 - \frac{\pi q_0}{2}$, where $0 < \mu_{0,0} < \mu_{0,1} < \frac{1}{2} < \mu_{1,0} < \mu_{1,1} < 1$. Table 1 presents the crossover

²We present our results for $p_0 = \frac{1}{2}$, which corresponds to starting the market at a uniform “prior” – a standard practice in prediction markets.

q_0	$\pi_c(q_0)$		p_A^{PBE} for LPP	
	Range of p^L		Range of p^L	
	$< \frac{1}{4}$	$\geq \frac{1}{4}$	$< \frac{1}{4}$	$\geq \frac{1}{4}$
$0 < q_0 < 2p^L$	$\pi_1^*(q_0)$	$\pi_1^*(q_0)$	$\mu_{1,0}$	$\mu_{1,0}$
$q_0 = 2p^L$	1	$\pi_1^*(q_0)$	NA	$\mu_{1,0}$
$2p^L < q_0 < \frac{1}{2}$	$\frac{2p^L}{q_0}$	$\pi_1^*(q_0)$	$\mu_{1,1}$	$\mu_{1,0}$
$q_0 = \frac{1}{2}$	$4p^L$	π^*	$\mu_{1,1} / \mu_{0,0}$	$\mu_{1,0}$

Table 1: The crossover probability π_c and as a function of q_0 over the sub-intervals into which p^L splits the entire possible range $(0, 1)$ of q_0 , for symmetric well-behaved MSRs. NA indicates that the LPP domain is never attained for that q_0 . LMSR is an example with $p^L < 1/4$, QMSR with $p^L = 1/4$ and SMSR with $p^L > 1/4$. $\pi_1^*(q_0)$ is the unique root in $(0, 1)$ of the following equation in π : $f(\mu_{1,0}) - f(p^H) - (\mu_{1,1} - p^H) f'(p^H) + (\frac{1-\pi}{2}) f'(\frac{1}{2}) = 0$; π^* is 1 for $p^L = 1/4$, and for $p^L > 1/4$, it is the unique root of $f(\frac{1}{2} - \frac{\pi}{4}) - f(p^L) - (\frac{\pi}{4} - p^L) f'(p^L) - (\frac{1-\pi}{2}) f'(\frac{1}{2}) = 0$.

probability π_c as a function of q_0 for sub-classes of symmetric well-behaved MSRs defined on the basis of the value of p^L , and also the values of Alice’s equilibrium price-reports in the LPP domain, for $q_0 \leq 1/2$. For $q_0 > 1/2$, the results are symmetric, hence omitted. In the HPP domain, p_A^{PBE} is p^H whenever $q_0 < 1/2$, and p^L whenever $q_0 > 1/2$. At $\pi = \pi_c$, then Alice is indifferent between her LPP and HPP price-reports although the values of these reports are, in general, distinct. Below, we outline the proof of Theorem 1.

Proof sketch Owing to linearity, Alice’s expected payoff with respect to her uncertainty in Bob’s participation and signal is equal to her payoff function evaluated at her expected outcome. Hence, using Lemmas 1 and 2, we can show that this expected payoff function is a piecewise continuous function of p_A consisting of segments of the “component” functions $\mathcal{R}_A(p_A, \frac{1}{2}, 0, 1)$, $\mathcal{R}_A(p_A, \frac{1}{2}, 0, 0)$, $\mathcal{R}_A(p_A, \frac{1}{2}, 1, 1)$, and $\mathcal{R}_A(p_A, \frac{1}{2}, 1, 0)$ over the sub-intervals $(0, p^L)$, $(p^L, \frac{1}{2}]$, $(\frac{1}{2}, p^H)$, $(p^H, 1)$ respectively, with jump discontinuities at p^L, p^H . The global maxima of these components in $[0, 1]$ are located respectively at $\mu_{0,1}$, $\mu_{0,0}$, $\mu_{1,1}$, and $\mu_{1,0}$, which depend on π and q_0 . However, for any given q_0 (and p^L, p^H), depending on the value of π , the global maximum of one of these components might lie outside the sub-interval in which that component is applicable so that the local suprema of some of the segments may lie at a threshold p^L or p^H . Taking these issues into account, we can determine the local suprema of the four segments and compare them to find the p_A that maximizes Alice’s expected payoff; the rest of the theorem follows from Lemmas 1 and 2, with the restriction that $v_B = 0$ if $p_A = p^L$ and $v_B = 1$ if $p_A = p^H$. \square

Figure 1 depicts the crossover probability π_c as a function of q_0 for each of the three selected MSRs. An interesting point is that for LMSR, when $0.4 < q_0 < 0.6$, which is a region of “high” uncertainty in Alice’s posterior about Bob’s private signal, π_c actually decreases with Alice’s increasing

uncertainty, i.e. the partially revealing LPP domain is realized for lower values of Bob’s non-participation probability than for the other MSRs. This is a peculiarity of any MSR with $p^L < 1/4$ as opposed those with $p^L \geq 1/4$.

3.1 Implications

Private signal revelation: STRATEGIC Bob’s report-vote pair is fully determined by Alice’s report and does not depend on s_B ; there is no guarantee that Alice’s vote will be truthful either, even in a LPP equilibrium (in general, she is likely to guess which way Bob will vote and vote the same way). However, if we invoke the assumption of stochastic relevance (Definition 1), then we can use p_A to uncover s_A .

Corollary 1. *If Alice’s signal s_A is stochastically relevant for Bob’s signal s_B , then the value of s_A can be recovered from Alice’s price-report in a LPP equilibrium $p_A^{LPP} = \mu_{u,v}(\pi, q_0)$, $u, v \in \{0, 1\}$, regardless of whether $v_A = s_A$.*

In other words, the very possibility of Bob not trading but voting truthfully engenders a situation (LPP domain) in which Alice, though strategic, is forced to indirectly divulge her private information! This has major implications for the value of prediction markets in situations where the “voting” or outcome-deciding actions of individuals are already assumed to be truthful and the designer would prefer not to introduce incentives that make them untruthful just because they added a trading platform. Corollary 1 suggests that, for an MSR-based prediction market, manipulators run the risk of having their cover blown when some fraction of the outcome-deciders refrain from market participation and are truthful in their outcome-affecting actions.

In a HPP equilibrium, p_A can only tell us whether $q_0 > \frac{1}{2}$ (if $p_A = p^L$) or $q_0 < \frac{1}{2}$ (if $p_A = p^H$), and is insufficient for recovering s_A without further assumptions.

HPP profit sharing: The HPP equilibria are a world where collusion appears with Alice as the “leader” picking the vote that both will coordinate on, and moving the price to just the level where it makes sense for STRATEGIC Bob to push the price all the way to 0 or 1 and vote the same way as Alice. In this way, they extract the maximum profit from the market maker, and split it between the two of them in a ratio that is dependent on the functional form of the MSR. In particular, for the three major MSRs considered, Alice makes more profit than Bob in a collusive equilibrium, with the discrepancy being the least for LMSR – we omit the straightforward calculations, and present the results in the following table:

Share in total HPP profit if Bob is STRATEGIC	LMSR	QMSR	SMSR
Alice’s share	67.81%	75%	78.32%
Bob’s share	32.19%	25%	21.68%

Corollary 2. *In a trading-voting game where the prediction market is implemented by any symmetric well-behaved MSR with lower threshold $p^L \geq \frac{1}{4}$, Alice’s ex post net profit in a HPP equilibrium is greater than that of STRATEGIC Bob.*

If Bob is HONEST, Alice’s payoff is obviously a function of his private signal faithfully announced in the outcome-deciding voting stage. Corollary 2 tells us that, even if Bob is

STRATEGIC and hence ends up colluding with the manipulator Alice, her profit share in a collusive equilibrium depends strongly on the MSR used – an insight that can potentially inform the choice of an MSR for market design.

Informativeness of market prices about final outcome: Finally, we can put all our results together into the following table (recall that the final price $p_B = p_A$ for HONEST Bob, and $p_B = v$ for STRATEGIC Bob), where “Bayes. Est.” (Bayesian estimate) is Alice’s expectation of the average vote before Bob trades; “Pre.” (Predetermined) signifies $p_A \in \{p^L, p^H\}$; “Correct” denotes the actual outcome.

	STRATEGIC Bob		HONEST Bob	
	LPP	HPP	LPP	HPP
p_A	Bayes. Est.	Pre.	Bayes. Est.	Pre.
p_B	Correct	Correct	Bayes. Est.	Pre.

3.2 A specific signal structure

We now consider a concrete example scenario to illustrate our findings: The underlying random variable takes values in the signal space itself, i.e. $T = \Omega = \{0, 1\}$, the prior probability of $\tau = 0$ being $\rho_0 \in (0, 1)$. Given τ , the agents’ signals are independently and identically distributed: for any “true” $\tau \in \{0, 1\}$, each participant gets the “correct” signal (identical to the true τ) with probability $(1 - \rho_e)$, otherwise gets the wrong signal; the error probability $\rho_e \in (0, 1) \setminus \{\frac{1}{2}\}$. Note that if and only if $\rho_e = \frac{1}{2}$, we have $q_0(0) = q_0(1) = \frac{1}{2}$ regardless of ρ_0 , hence signals are not *informative* [Chen *et al.*, 2009]. Then, $q_0(0) = \frac{(1-\rho_e)^2\rho_0 + \rho_e^2(1-\rho_0)}{(1-\rho_e)\rho_0 + \rho_e(1-\rho_0)}$, and $q_0(1) = \frac{(1-\rho_e)\rho_e}{\rho_e\rho_0 + (1-\rho_e)(1-\rho_0)}$. This signal structure has multiple interesting information-revealing characteristics: First, we have $q_0(0) \neq q_0(1)$, i.e. Alice’s signal is stochastically relevant for that of Bob. Hence, Corollary 1 applies. Second, it is easy to show that, if $\rho_0 = \frac{1}{2}$ (a uniform common prior), then Alice’s vote is always truthful since, for any $\rho_e \in (0, 1)$, $s_A = 0 \Leftrightarrow q_0 > \frac{1}{2} \Leftrightarrow v_A = 0$.

Figure 2 shows Alice’s equilibrium report in a LMSR market and her expected market outcome vs. π , for $s_A = 0$ and fixed ρ_0, ρ_e (hence, a fixed q_0). Note the HPP and LPP regions to the left and right of the cross-over probability, where Alice’s price-report (the dashed curve) is distinct from and coincides with her expectation of the average vote (the continuous curve) respectively. The corresponding plots for the other two MSRs are qualitatively similar.

4 Discussion

Our model is stylized, but the framework and methodology can be applied to more complex scenarios. Below, we sketch two specific lines of generalization.

Additional outcome-deciders who do not trade: Consider a scenario in which Alice and Bob are the only traders but jointly decide less than 100% of the outcome, say, $v = \frac{v_A + v_B + \sum_{i=1}^n v_i}{n+2}$, where $\{v_i\}_{i=1}^n$ are the votes (and also the private signals) of n non-strategic agents. To solve for equilibria, we now need, in addition to Bob’s non-participation probability π and Alice’s posterior belief about others’ signals given s_A , Bob’s posterior belief given s_B and Alice’s trading action; but we can use the same methodology as in

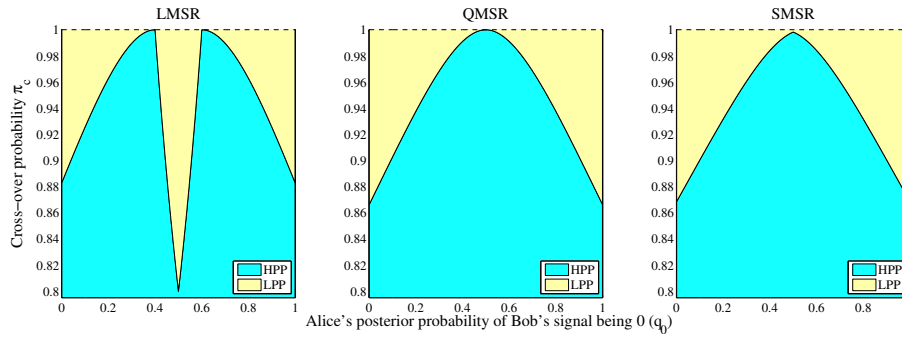


Figure 1: Dependence of crossover probability on Alice’s posterior belief about Bob for the three MSRs; e.g. for QMSR, if $q_0 = 0.25$. then $\pi_c \approx 0.9537$, so we have a LPP equilibrium with $p_A = (1 + \pi(1 - q_0))/2$ for $\pi > 0.9537$, and a HPP equilibrium with $p_A = p^H = 0.75$ for $\pi < 0.9537$.

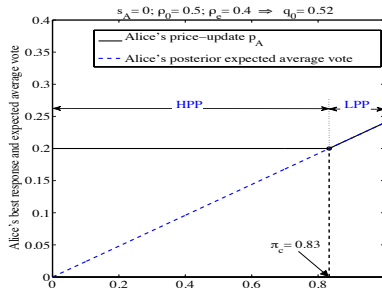


Figure 2: Crossover from HPP to LPP equilibria regions for LMSR over $0 < \pi < 1$ for a particular instantiation of the signal structure described in Section 3.2.

Section 3 to show that the PBE is still of two types (with some additional characteristics) depending on model parameters. For example, if all agents receive independent and identically distributed signals conditional on the type of the underlying entity, and the signal structure is such that $v_A = 0$ if and only if $s_A = 0$ (as in Section 3.2), then the main deviation from the analysis in Section 3 is that Bob’s thresholds p^L and p^H now become functions of $s_B \in \{0, 1\}$ so that each has two possible values. For a low enough π , Alice’s best response is to set p_A at an uninformative value ($p_{s_B=0}^L, p_{s_B=1}^L, p_{s_B=0}^H$, or $p_{s_B=1}^H$) but now STRATEGIC Bob decides whether or not to “agree with” Alice depending on s_B , hence this equilibrium reveals Bob’s signal; for a high enough π , p_A coincides with Alice’s expected outcome but now her action fully determines STRATEGIC Bob’s action, so Bob’s signal cannot be inferred from his action, as in the two-player game; in either case, STRATEGIC Bob moves the price to his posterior expectation of v , which is no longer in $\{0, 1\}$.

Additional traders who do not affect the outcome: Agents with no control over the outcome who trade before Alice only matter in the level to which they move the price seen by Alice but, from Alice’s perspective, this is equivalent to a ‘starting’ price of a general $p_0 \in (0, 1)$; if they all trade after Bob, Alice and Bob’s equilibrium actions remain unchanged because, in an MSR, an agent’s payoff depends on the actions of her predecessors and not on those of her successors, by de-

sign (as long as these successors are not outcome-deciders). The game becomes more complex for Alice when there are intermediate traders between Alice and Bob, but we believe that our model can serve as a foundation for analyzing this extension as well.

5 Conclusion

This paper is a first step in exploring the crucial incentive issues that have the potential to derail the effectiveness of prediction markets for various forecasting tasks. We have introduced a new formal model for studying the incentives for and the impact of manipulation in prediction markets whose participants can affect the outcome by taking actions external to the market but there is some uncertainty about the market participation of some outcome-deciders. We have characterized the equilibria of the induced game, discussed their properties, and outlined important extensions. Interesting avenues for future work include generalizing our results to markets with other price-setting mechanisms, richer signal structures, outcome functions other than the mean vote (such as non-linear and / or noisy functions of the agents’ second-stage actions), and agents who also strategically pick the time-points at which they trade.

Acknowledgments

We would like to thank Yiling Chen for helpful discussions. We are grateful for support from NSF IIS awards 1414452 and 1527037.

References

- [Abernethy and Frongillo, 2012] J. Abernethy and R. Frongillo. A characterization of scoring rules for linear properties. *COLT*, 2012.
- [Berg and Rietz, 2006] J. Berg and T. Rietz. The Iowa electronic markets: Stylized facts and open issues. *Information Markets: A New Way of Making Decisions*, pages 142–169, 2006.
- [Berg et al., 2008] J. Berg, R. Forsythe, F. Nelson, and T. Rietz. Results from a dozen years of election futures markets research. *Handbook Exp. Econ. Res.*, 1:742–751, 2008.

- [Boutilier, 2012] C. Boutilier. Eliciting forecasts from self-interested experts: scoring rules for decision makers. *Proc. AAMAS*, 2:737–744, 2012.
- [Brahma *et al.*, 2012] A. Brahma, M. Chakraborty, S. Das, A. Lavoie, and M. Magdon-Ismail. A Bayesian market maker. In *Proc. ACM EC*, pages 215–232, 2012.
- [Chakraborty *et al.*, 2013] M. Chakraborty, S. Das, A. Lavoie, M. Magdon-Ismail, and Y. Naamad. Instructor rating markets. In *Proc. AAAI*, pages 159–165, 2013.
- [Chen and Pennock, 2007] Y. Chen and D. Pennock. A utility framework for bounded-loss market makers. In *Proc. UAI*, pages 49–56, 2007.
- [Chen *et al.*, 2009] Y. Chen, S. Dimitrov, R. Sami, D. Reeves, D. Pennock, R. Hanson, L. Fortnow, and R. Gonen. Gaming prediction markets: Equilibrium strategies with a market maker. *Algorithmica*, pages 1–40, 2009.
- [Chen *et al.*, 2011a] Y. Chen, X. Gao, R. Goldstein, and I. Kash. Market manipulation with outside incentives. In *Proc. AAAI*, 2011.
- [Chen *et al.*, 2011b] Y. Chen, I. Kash, M. Ruberry, and V. Shnayder. Decision markets with good incentives. In *Proc. WINE*, pages 89–100, 2011.
- [Cowgill and Zitzewitz, 2015] B. Cowgill and E. Zitzewitz. Corporate prediction markets: Evidence from Google, Ford, and Firm X. *Review Econ. Studies*, 82(4):1309–1341, 2015.
- [Das, 2008] S. Das. The effects of market-making on price dynamics. In *Proc. AAMAS*, May 2008.
- [Dimitrov and Sami, 2010] S. Dimitrov and R. Sami. Composition of markets with conflicting incentives. In *Proc. ACM EC*, pages 53–62, 2010.
- [Foster and Viswanathan, 1996] F. Foster and S. Viswanathan. Strategic trading when agents forecast the forecasts of others. *J. Fin.*, 51:1437–1478, 1996.
- [Fudenberg and Tirole, 1991] D. Fudenberg and J. Tirole. Perfect Bayesian equilibrium and sequential equilibrium. *J. Econ. Theory*, 53(2):236–260, 1991.
- [Gao *et al.*, 2013] X. Gao, J. Zhang, and Y. Chen. What you jointly know determines how you act: Strategic interactions in prediction markets. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 489–506. ACM, 2013.
- [Glosten and Milgrom, 1985] L. Glosten and P. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *J. Fin. Econ.*, 14:71–100, 1985.
- [Gneiting and Raftery, 2007] T. Gneiting and A. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.
- [Hanson and Oprea, 2009] R. Hanson and R. Oprea. A manipulator can aid prediction market accuracy. *Economica*, 76(302):304–314, 2009.
- [Hanson *et al.*, 2006] R. Hanson, R. Oprea, and D. Porter. Information aggregation and manipulation in an experimental market. *J. Econ. Beh. and Org.*, 60:449–459, 2006.
- [Hanson, 1999] R. Hanson. Decision markets. *IEEE Intelligent Systems*, 14(3):16–19, 1999.
- [Hanson, 2003] R. Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5(1):107–119, 2003.
- [Hanson, 2007a] R. Hanson. Logarithmic market scoring rules for modular combinatorial information aggregation. *J. Prediction Markets*, 1(1):3–15, 2007.
- [Hanson, 2007b] R. Hanson. The Policy Analysis Market (a thwarted experiment in the use of prediction markets for public policy). *Innovations*, 2(3):73–88, 2007.
- [Holden and Subrahmanyam, 1992] C. Holden and A. Subrahmanyam. Long-lived private information and imperfect competition. *J. Fin.*, 47:247–270, 1992.
- [Huang and Shoham, 2014] E. Huang and Y. Shoham. Price manipulation in prediction markets: analysis and mitigation. In *Proc. AAMAS*, pages 213–220, 2014.
- [Kyle, 1985] A. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1336, 1985.
- [Miller *et al.*, 2005] N. Miller, P. Resnick, and R. Zeckhauser. Eliciting Informative Feedback: The Peer-Prediction Method. *Management Science*, 5(9):1359–1373, Sept. 2005.
- [Ostrovsky, 2012] M. Ostrovsky. Information aggregation in dynamic markets with strategic traders. *Econometrica*, 80(6):2595–2647, 2012.
- [Othman and Sandholm, 2010] A. Othman and T. Sandholm. Decision rules and decision markets. In *Proc. AAMAS*, pages 625–632, 2010.
- [Peterson and Krug, 2015] J. Peterson and J. Krug. Augur: a decentralized, open-source platform for prediction markets. *arXiv preprint arXiv:1501.01042*, 2015.
- [Rhode and Strumpf, 2006] P. Rhode and K. Strumpf. Manipulating political stock markets: A field experiment and a century of observational data. *University of Arizona, mimeo*, 2006.
- [Rothschild and Sethi, 2015] D. Rothschild and R. Sethi. Trading strategies and market microstructure: Evidence from a prediction market. *Available at SSRN 2322420*, 2015.
- [Shi *et al.*, 2009] P. Shi, V. Conitzer, and M. Guo. Prediction mechanisms that do not incentivize undesirable actions. In *Proc. WINE*, 2009.
- [Stiglitz, July 31 2003] J. Stiglitz. Terrorism: There’s no futures in it. *Los Angeles Times*, July 31, 2003.
- [Wolfers and Zitzewitz, 2004] J. Wolfers and E. Zitzewitz. Prediction markets. *J. Econ. Perspectives*, 18(2):107–126, 2004.