

## Market Manipulation with Outside Incentives

**Yiling Chen**

Harvard SEAS

yiling@eecs.harvard.edu

**Xi Alice Gao**

Harvard SEAS

xagao@seas.harvard.edu

**Rick Goldstein**

Harvard SEAS

rgoldst@fas.harvard.edu

**Ian A. Kash**

Harvard CRCS

kash@seas.harvard.edu

### Abstract

Much evidence has shown that prediction markets, when used in isolation, can effectively aggregate dispersed information about uncertain future events and produce remarkably accurate forecasts. However, if the market prediction will be used for decision making, a strategic participant with a vested interest in the decision outcome may want to manipulate the market prediction in order to influence the resulting decision. The presence of such incentives outside of the market would seem to damage information aggregation because of the potential distrust among market participants. While this is true under some conditions, we find that, if the existence of such incentives is certain and common knowledge, then in many cases, there exists a separating equilibrium for the market where information is fully aggregated. This equilibrium also maximizes social welfare for convex outside payoff functions. At this equilibrium, the participant with outside incentives makes a costly move to gain the trust of other participants. When the existence of outside incentives is uncertain, however, trust cannot be established between players if the outside incentive is sufficiently large and we lose the separability in equilibrium.

### Introduction

Prediction markets are powerful tools created to aggregate information from individuals about uncertain events of interest. As a betting intermediary, a prediction market allows traders to express their private information through trading shares of contracts and rewards their contributions based on the realized outcome. The reward scheme in a prediction market is designed to offer incentives for traders to reveal their private information. For instance, Hanson's market scoring rule (Hanson 2007) incentivizes risk-neutral, myopic traders to truthfully reveal their probabilistic estimates by ensuring that truthful betting maximizes their expected payoffs. Substantial empirical work has shown that prediction markets produce remarkably accurate forecasts (Berg et al. 2001; Wolfers and Zitzewitz 2004; Goel et al. 2010).

However, in many cases, the ultimate purpose of information aggregation mechanisms is to inform decision making. If the forecast of a prediction market is used to make a decision, some market participants may stand to benefit if a

particular decision outcome is reached. This creates strong incentives from outside of the market for these participants to strategically manipulate the market probability and deceive other participants, especially when the outside incentive is relatively more attractive than the payoff from inside the market.

As a motivating example, suppose the US Centers for Disease Control and Prevention (CDC) wants to accurately predict the flu activity level for the next flu season in order to purchase an appropriate amount of flu vaccine in advance. To accomplish this, the CDC could run a prediction market and base its purchasing decision on the final market forecast. In this case, suppliers of flu vaccines, such as pharmaceutical companies, may have conflicting incentives inside and outside of the market. A pharmaceutical company can profit either by truthfully reporting their information in the market or by driving up the final market probability to increase their profit from selling flu vaccines. This outside incentive may cause the pharmaceutical company to manipulate the market probability in order to mislead the CDC about the expected flu activity level.

When participants have outside incentives to manipulate the market probability, it is questionable whether information can be fully aggregated. In this paper, we investigate information aggregation in prediction markets when such outside incentives exist. We study a simple model of prediction markets with two participants. Following a predefined sequence, each participant makes a single trade. With some probability, the first participant has an outside payoff which is an increasing function of the final market probability. We analyze two cases: (1) the first participant has the outside payoff with probability 1, and (2) the probability for the first participant to have the outside payoff is less than 1. Our main results are:

- For case (1), we give a necessary and sufficient condition for the existence of a separating equilibrium under which information is fully aggregated despite the outside incentive. We characterize a separating equilibrium where the first participant makes a costly move to gain trust of the second participant.
- For case (2), we prove that there exists no separating or semi-separating equilibrium where information is fully aggregated if the outside incentive is sufficiently large.

Information loss is inevitable since the first participant can benefit by pretending to not have the outside payoff when she actually does.

**Related Work** An emerging line of research has studied incentive issues that arise when using prediction markets as a decision tool. Once incorporated into the decision process, prediction markets often unintentionally create incentives for participants to manipulate the market probability. These incentives could take the form of the potential to profit in a subsequent market (Dimitrov and Sami 2010) or the ability to influence the decision being made in a decision market (Chen and Kash 2011; Othman and Sandholm 2010) to make more profit within the market. Other types of manipulation in a prediction market include influencing the market outcome through alternative means other than trading in the market (Shi, Conitzer, and Guo 2009), and taking advantage of the opportunity to participate multiple times and mislead other traders (Chen et al. 2010)

Of this line of research, Dimitrov and Sami's (2010) work is the closest to our own. The main differences between their work and ours are (1) the outside incentives in their model take the form of the potential profit in a second market whereas ours take the general form of any monotone function of the final market probability, and (2) they show some properties of players' payoffs at the equilibria without the explicit characterization of any equilibrium whereas we characterize the equilibria of our game play.

## Model

Consider a binary random variable  $X$ . We run a prediction market to predict its realization  $x \in \{0, 1\}$ . Our market uses a logarithmic market scoring rule (LMSR) (Hanson 2007), which is a sequential shared version of the logarithmic scoring rule

$$s(x, p) = \begin{cases} b \log(p), & \text{if } x = 1 \\ b \log(1 - p), & \text{if } x = 0 \end{cases} \quad (1)$$

where  $b$  is a positive parameter and  $p$  is the reported probability for  $x = 1$ . We assume  $b = 1$  without loss of generality. Starting with an initial market probability  $f_0$ , the LMSR market sequentially interacts with each trader to collect his probability assessment<sup>1</sup>. When a trader changes the market probability from  $p$  to  $q$ , he is paid the scoring rule difference,  $s(x, q) - s(x, p)$ . It is known that LMSR incentivizes risk-neutral, myopic traders to truthfully reveal their probabilistic assessment.

Alice and Bob are two rational, risk-neutral participants in the market. They receive private signals described by the random variables  $S_A$  and  $S_B$  with realizations  $s_A, s_B \in \{H, T\}$  respectively. Let  $\pi$  denote a joint prior probability distribution over  $X$ ,  $S_A$  and  $S_B$ . We assume  $\pi$  is common knowledge and omit it in our notation for brevity.

<sup>1</sup>Even though we describe the LMSR market in terms of updating probabilities, it can be implemented as a market where participants trade shares of contracts (Hanson 2007; Chen and Pennock 2007).

We define  $f_{s_A,0} = P(x = 1|S_A = s_A)$  and  $f_{0,s_B} = P(x = 1|S_B = s_B)$  to represent the posterior probability for  $x = 1$  given the individual signal of Alice and Bob respectively. Similarly,  $f_{s_A,s_B} = P(x = 1|S_A = s_A, S_B = s_B)$  represents the posterior probability of  $x = 1$  given both signals. We assume that the  $H$  signal indicates a strictly higher probability for  $x = 1$  than the  $T$  signal. That is, we assume  $f_{H,s_B} > f_{0,s_B} > f_{T,s_B}$  for any  $s_B$  and  $f_{s_A,H} > f_{s_A,0} > f_{s_A,T}$  for any  $s_A$ , which imply  $f_{H,0} > f_{T,0}$  and  $f_{0,H} > f_{0,T}$ .

In the context of our flu prediction example, we can interpret the realization  $x = 1$  as the event that the flu is widespread and  $x = 0$  as the event that it is not. Then the two private signals can be any information acquired by the participants about the flu activity, such as the number of people catching the flu in a local area or the person's own health condition.

## Sequence of Play

Our game has two stages. In stage 1, Alice observes her signal  $s_A$  and changes the market probability from  $f_0$  to  $r_A$ . In stage 2, Bob observes Alice's report  $r_A$  in stage 1 and his private signal  $s_B$ , and changes the market probability from  $r_A$  to  $r_B$ . The market closes after Bob's report. The sequence of play is common knowledge.

## Player Payoffs

In our model, both Alice and Bob can profit from the LMSR market. Moreover, with a fixed probability  $\alpha \in (0, 1]$ , Alice is of a type which has an outside payoff  $Q(r_B)$ , a continuous and (weakly) increasing function of the final market probability  $r_B$ . In the flu prediction example, this outside payoff may correspond to the pharmaceutical company's profit from selling flu vaccines. The outside payoff function  $Q(\cdot)$  and the value of  $\alpha$  are common knowledge.

## Solution Concept

Our solution concept is the Perfect Bayesian Equilibrium (PBE) (Fudenberg and Tirole 1991), which is a refinement of Bayesian Nash equilibrium. Informally, a strategy-belief pair is a PBE if the players' strategies are optimal given their beliefs and the players' beliefs can be derived from their strategies using Bayes' rule whenever possible.

We use the notion of separating and pooling equilibrium (Spence 1973) in our analysis. In our model, if two types of Alice separate at a PBE, then these types of Alice must report different values. Otherwise, these two types of Alice pool at the PBE and report the same value. An equilibrium can be semi-separating, in which case some types separate and other types pool. If all types of Alice separate at a PBE, then information can be fully aggregated since Bob can distinguish Alice's signals and always make the optimal report. Note that, in our model, when  $\alpha \in (0, 1)$  Alice has 4 types based on whether she has the outside payoff and her realized signal. However, if  $\alpha = 1$ , then Alice only has 2 types distinguished by her signal.

## Strategies and Beliefs

In stage 1, The market starts with the probability  $f_0$ . For a given signal  $s_A$ , Alice moves the market probability from  $f_0$  to  $r_A \in [0, 1]$ . When Alice does not have an outside

payoff, since she only participates once, her optimal strategy facing the market scoring rule is to report  $f_{s_A,0}$  with probability 1 after receiving the  $s_A$  signal. If Alice has an outside payoff, we denote her strategy as a mapping  $\sigma : \{H, T\} \rightarrow \Delta([0, 1])$ , where  $\Delta(S)$  denotes the space of distributions over a set  $S$ . We further assume that the support of Alice's strategy is finite<sup>2</sup>. We use  $\sigma_{s_A}(r_A)$  to denote the probability for Alice to report  $r_A$  after receiving the  $s_A$  signal.

In stage 2, Bob moves the market probability from  $r_A$  to  $r_B$ . We denote Bob's belief function as a mapping  $\mu : [0, 1] \times \{H, T\} \rightarrow \Delta(\{H, T\})$ , and we use  $\mu_{s_B, r_A}(s_A)$  to denote the probability that Bob assigns to Alice having received the  $s_A$  signal given that she reported  $r_A$ . Since Bob participates last in our game, his optimal strategy is uniquely determined by Alice's report  $r_A$ , his realized signal  $s_B$  and his belief  $\mu$ ; he will report  $r_B = \mu_{s_B, r_A}(H)f_{H, s_B} + \mu_{s_B, r_A}(T)f_{T, s_B}$ .

In the rest of the paper, we simplify our PBE representation by only describing Alice's strategy and Bob's belief since Alice plays first and Bob has a dominant strategy.

### Known Outside Incentive

In this section, we analyze the special case of our model when  $\alpha = 1$ , that is, Alice's incentive is common knowledge. Due to the presence of the outside payoff, Alice has an incentive to mislead Bob in order to drive up the final market probability. In equilibrium, Bob recognizes this incentive, and discounts Alice's report accordingly. Therefore, we naturally expect information loss in equilibrium due to Alice's manipulation. However, from another perspective, Alice's welfare is also hurt by her manipulation since she cannot fully convince Bob when she has a favorable signal  $H$ . In the following analysis, we characterize a necessary and sufficient condition under which there exists a separating equilibrium that achieves full information aggregation and maximizes social welfare with a convex  $Q(\cdot)$ . At this separating equilibrium, Alice makes a costly statement, in the form of a loss in the market scoring rule payoff, in order to convince Bob that she is revealing her signals, despite the incentive to manipulate. If the condition is violated, we show that there does not exist any separating equilibrium and information loss is inevitable.

### Non-Existence of Truthful Equilibrium

Before we begin our equilibrium analysis, we present a simple argument that when certain types of outside incentives are present Alice's truthful strategy cannot be part of any PBE. By Alice's truthful strategy, we mean the strategy of reporting  $f_{s_A,0}$  with probability 1 after receiving the  $s_A$  signal, i.e.

$$\sigma_H(f_{H,0}) = 1, \sigma_T(f_{T,0}) = 1. \quad (2)$$

Suppose that Alice uses the truthful strategy at some PBE. Then Bob's belief on the equilibrium path must be derived

<sup>2</sup>This assumption is often used to avoid the technical difficulties that PBE has for games with a continuum of strategies, e.g. in (Cho and Kreps 1987).

from Alice's strategy using Bayes' rule, that is,

$$\mu_{s_B, f_{H,0}}(H) = 1, \mu_{s_B, f_{T,0}}(T) = 1. \quad (3)$$

Given Bob's belief, Alice can reason about her potential loss and gain by misreporting the  $T$  signal. In general, if Alice reports  $r_A$  with positive probability after receiving the  $s_A$  signal, her expected loss in market scoring rule payoff is

$$L_{s_A}(r_A) = f_{s_A,0} \log \frac{f_{s_A,0}}{r_A} + (1 - f_{s_A,0}) \log \frac{1 - f_{s_A,0}}{1 - r_A} \quad (4)$$

So her loss by reporting  $f_{H,0}$  with probability 1 after receiving the  $T$  signal is  $L_T(f_{H,0})$ . Based on Bob's belief, her gain by such misreporting is  $E_{s_B}[Q(f_{H, s_B}) - Q(f_{T, s_B}) \mid s_A = T]$ . As a result, if the outside payoff function  $Q(\cdot)$  satisfies  $L_T(f_{H,0}) < E_{s_B}[Q(f_{H, s_B}) - Q(f_{T, s_B}) \mid s_A = T]$ , then Alice can derive positive net payoff by deviating to the strategy  $\sigma_H(f_{H,0}) = 1, \sigma_T(f_{H,0}) = 1$ . Therefore, Lemma 1 follows.

**Lemma 1.** *There exists an outside payoff function  $Q(\cdot)$  such that Alice's truthful strategy (2) is not part of any PBE.*

### A Condition for Separation

In this part, we derive a condition that, as we will show, is necessary and sufficient for the existence of a separating equilibrium. This allows us to divide our subsequent equilibrium analysis into two cases. This condition involves  $Y_H$ , the unique value in  $[f_{H,0}, 1]$  satisfying equation (5), and  $Y_T$ , the unique value in  $[f_{T,0}, 1]$  satisfying equation (6):

$$L_H(Y_H) = E_{s_B}[Q(f_{H, s_B}) - Q(f_{T, s_B}) \mid s_A = H], \quad (5)$$

$$L_T(Y_T) = E_{s_B}[Q(f_{H, s_B}) - Q(f_{T, s_B}) \mid s_A = T]. \quad (6)$$

The RHS of equations (5) and (6) are nonnegative because  $f_{H, s_B} > f_{T, s_B}$  and  $Q(\cdot)$  is an increasing function.  $L_T(Y_H)$  and  $L_H(Y_T)$  are monotonically increasing for  $Y_H \in [f_{H,0}, 1]$  and  $Y_T \in [f_{T,0}, 1]$ , and has the range  $[0, +\infty]$ . Hence,  $Y_H$  and  $Y_T$  are well defined.

Intuitively,  $Y_H$  and  $Y_T$  are the maximum values that Alice might be willing to report after receiving the  $H$  or  $T$  signal respectively. The RHS of equations (5) or (6) is Alice's maximum possible gain in outside payoff by reporting some value  $r_A$  when she has the  $H$  or  $T$  signal. These maximum gains would be achieved if Bob had the (hypothetical) belief that Alice has the  $H$  signal when she reports  $r_A$  and the  $T$  signal otherwise. Given the  $H$  or  $T$  signal, Alice can not possibly report any value higher than  $Y_H$  or  $Y_T$  because doing so is dominated by reporting  $f_{H,0}$  or  $f_{T,0}$ .

When  $Y_H \geq Y_T$ , if Alice chooses to, it is possible for her to credibly reveal that she has the  $H$  signal by reporting a value that is too high to be profitable given the  $T$  signal. However, when  $Y_H < Y_T$ , this is not possible. While we focus on understanding information aggregation in markets with outside incentives, interestingly, our problem is essentially a signaling game (see (Fudenberg and Tirole 1991) for a definition and examples). In fact, our condition that  $Y_H > Y_T$  is analogous to the requirement in the applicant signaling game (Spence 1973) that education is cheaper for better workers, without which education is not useful as a signal of worker quality.



In what follows, we divide our analysis into these two cases. When  $Y_H \geq Y_T$ , we characterize a separating equilibrium of our game. When  $Y_H < Y_T$ , we prove that no separating equilibrium exists and derive a pooling equilibrium.

### Separating Equilibrium

We assume  $Y_H \geq Y_T$  in this section. Whether this condition is satisfied depends on the prior probability distribution  $\pi$  and the external payoff function  $Q(\cdot)$ . As a special case, if signals  $S_A$  and  $S_B$  are independent, the condition is trivially satisfied.

We will characterize a separating PBE of our game and show that it maximizes social welfare when  $Q(\cdot)$  is convex. In our equilibrium, Bob has the following belief  $\mu^S$ ,

$$\mu_{s_B, r_A}^S(H) = \begin{cases} 1, & \text{if } r_A \in [Y_T, 1] \\ 0, & \text{if } r_A \in [0, Y_T) \end{cases} \quad (7)$$

This belief says that if Alice makes a report that is too high to be consistent with the  $T$  signal ( $r_A \geq Y_T$ ), then Bob believes that she received the  $H$  signal. This is reasonable since Alice has no incentive to report  $Y_T$  or higher when she receives the  $T$  signal by the definition of  $Y_T$ . If Alice reports a value that is low enough such that it is still profitable for her to report the value with a  $T$  signal ( $r_A < Y_T$ ), then Bob believes that she received the  $T$  signal.

We now show that Bob's belief (7) and Alice's strategy

$$\sigma_H^S(r_A) = 1, \sigma_T^S(f_{T,0}) = 1, r_A = \max(Y_T, f_{H,0}) \quad (8)$$

form a PBE of our game. Intuitively, when  $f_{H,0} < Y_T$ , Alice is willing to incur a high enough cost by reporting  $Y_T$  after receiving the  $H$  signal, to convince Bob that she has the  $H$  signal. Since Bob can perfectly infer Alice's signal by observing her report, he would report  $f_{s_A, s_B}$  in stage 2 and information is fully aggregated. So Alice is essentially letting Bob take a larger portion of the market scoring rule payoff in exchange for a larger outside payoff. Theorem 1 describes this separating PBE.

**Theorem 1.** *When  $Y_H \geq Y_T$ , Alice's strategy (8) and Bob's belief (7) form a separating PBE.*

*Proof.* First, we show that if  $Y_H \geq Y_T$ , then Alice's strategy (8) is optimal given Bob's belief.

If  $f_{H,0} < Y_T$ , then it is optimal for Alice to report  $Y_T$  after receiving the  $H$  signal because her gain in outside payoff equals  $L_H(Y_H)$  which is greater than her loss in the market  $L_H(Y_T)$ . Otherwise, if  $f_{H,0} \geq Y_T$ , then it's optimal for Alice to report  $f_{H,0}$  after receiving the  $H$  signal. Therefore, Alice's optimal strategy after receiving the  $H$  signal is to report  $\max(f_{H,0}, Y_T)$ .

When Alice receives the  $T$  signal, Alice would not report any  $r_A > Y_T$  by definition of  $Y_T$ , and furthermore she is indifferent between reporting  $Y_T$  and  $f_{T,0}$ . Any other report is dominated by a report of  $f_{T,0}$ . Therefore, it is optimal for Alice to report  $f_{T,0}$  after receiving the  $T$  signal.

Moreover, we can show that Bob's belief is consistent with Alice's strategy by mechanically applying Bayes' rule (argument omitted). Given the above arguments, Alice's strategy and Bob's belief form a PBE of this game.  $\square$

### Other Equilibria

The separating equilibrium we have derived is not the only one possible when  $Y_H \geq Y_T$ . For example, beliefs can be found for Bob such that there is a separating equilibrium where Alice always reports  $f_{T,0}$  with the  $T$  signal and some value in  $[Y_T, Y_H]$  with the  $H$  signal. However, since all separating equilibria fully aggregate information, they all result in the same social welfare; all that changes is how Alice and Bob split the resulting payoff. As the following theorem shows, the particular equilibrium we have chosen is the one that maximizes Alice's payoff.

**Theorem 2.** *Among all separating PBE of our game, the separating PBE in Theorem 1 where Alice uses the strategy (8) and Bob has the belief (7) maximizes Alice's total expected payoff.*

*Proof.* In all separating PBE, Alice's expected outside payoff is the same. We first show that Alice must report  $f_{T,0}$  after receiving the  $T$  signal at any separating PBE. Suppose not and Alice reports  $r_A \neq f_{T,0}$  after receiving the  $T$  signal. Bob's belief must be  $\mu_{s_B, r_A}(H) = 0$ , and  $\mu_{s_B, f_{T,0}}(H) \geq 0$ . However, given any  $\mu_{s_B, f_{T,0}}(H) \geq 0$ , Alice is strictly better off reporting  $f_{T,0}$ , which is a contradiction. Therefore, Alice's payoff after receiving the  $T$  signal is the same at any separating equilibrium.

In Theorem 1, when  $f_{H,0} \geq Y_T$ , Alice reports  $f_{H,0}$  after receiving the  $H$  signal and this is the maximum payoff she could get after receiving the  $H$  signal. When  $f_{H,0} < Y_T$ , Alice's optimal strategy in Theorem 1 is to report  $Y_T$ . In any equilibrium where she reports a value greater than  $Y_T$ , she is strictly worse off. There does not exist a separating equilibrium in which Alice reports  $r_A \in [f_{T,0}, Y_T)$  after receiving the  $H$  signal. We show this by contradiction. Suppose that there exists a separating equilibrium in which Alice reports  $r_A \in [f_{T,0}, Y_T)$  after receiving the  $H$  signal. Since the PBE is separating,  $r_A$  can not equal  $f_{T,0}$ , which is the report when Alice receives the  $T$  signal. In addition, Bob's belief must be  $\mu_{s_B, r_A}(H) = 1$  to be consistent with Alice's strategy. We can also derive a contradiction or domination by  $Y_T$  for the case when  $r_A < f_{T,0}$  in a similar way (argument omitted). Therefore, when  $f_{H,0} < Y_T$ , reporting  $Y_T$  maximizes Alice's payoff after receiving the  $H$  signal.

Therefore, the separating PBE in Theorem 1 maximizes Alice's expected total payoff among all separating PBE of our game.  $\square$

In addition to other separating equilibria, there may also exist pooling equilibria. In terms of the efficiency of the market, the separating PBE is superior since it achieves the maximum total market scoring rule payoff. Moreover, if we focus on convex  $Q(\cdot)$  functions, we can show that the separating PBE maximizes the social welfare. Situations with a convex  $Q(\cdot)$  function arise, for example, when manufactures have increasing returns to scale, which might be the case in our flu prediction example.

**Theorem 3.** *For any convex  $Q(\cdot)$  function, if  $Y_H \geq Y_T$ , then among all PBE, any separating PBE maximizes social welfare.*

*Proof.* The separating PBE maximizes the total market scoring rule payoff. Next, we show that the separating PBE also maximizes Alice's outside incentive payoff.

Consider an arbitrary PBE of this game. Let  $K$  denote the union of the supports of Alice's strategy after receiving the  $H$  and the  $T$  signals at this PBE. Let  $u_A^G$  denote Alice's expected outside payoff at this PBE and let  $u_A^S$  denote Alice's expected outside payoff at any separating PBE. We prove below that  $u_A^G \leq u_A^S$ . Note that we simplify our notation by using  $P(S_A, S_B) = P(s_A = S_A, s_B = S_B)$ .

$$\begin{aligned}
u_A^G &= \sum_{v \in K} (P(H, H)\sigma_H(v) + P(T, H)\sigma_T(v)) \\
&\quad Q\left(\frac{P(H, H)\sigma_H(v)}{P(H, H)\sigma_H(v) + P(T, H)\sigma_T(v)}f_{HH} \right. \\
&\quad \left. + \frac{P(T, H)\sigma_T(v)}{P(H, H)\sigma_H(v) + P(T, H)\sigma_T(v)}f_{TH}\right) \\
&\quad + (P(H, T)\sigma_H(v) + P(T, T)\sigma_T(v)) \\
&\quad Q\left(\frac{P(H, T)\sigma_H(v)}{P(H, T)\sigma_H(v) + P(T, T)\sigma_T(v)}f_{HT} \right. \\
&\quad \left. + \frac{P(T, T)\sigma_T(v)}{P(H, T)\sigma_H(v) + P(T, T)\sigma_T(v)}f_{TT}\right) \\
&\leq \sum_{v \in K} (P(H, H)\sigma_H(v)Q(f_{HH}) + P(H, T)\sigma_H(v)Q(f_{HT}) \\
&\quad + P(T, H)\sigma_T(v)Q(f_{TH}) + P(T, T)\sigma_T(v)Q(f_{TT})) \quad (10) \\
&= P(H, H)Q(f_{HH}) + P(H, T)Q(f_{HT}) \\
&\quad + P(T, H)Q(f_{TH}) + P(T, T)Q(f_{TT}) \\
&= u_A^S
\end{aligned}$$

where inequality (10) was derived by applying the convexity of the  $Q(\cdot)$  function.

Therefore, among all PBE of this game, the separating PBE maximizes the social welfare.  $\square$

Our model is not unique in suffering from a multiplicity of equilibria; multiple equilibria exist in many signaling games as well (e.g. (Spence 1973)). There has been some work in economics that considers equilibrium refinements stronger than perfect Bayesian equilibrium to try and identify one particular equilibrium as focal. Cho and Kreps (1987) present a number of refinements, including one they call "The Intuitive Criterion." It is easy to show that our separating equilibrium is the unique separating equilibrium consistent with this refinement. However, there can still be pooling equilibria consistent with it.

### Pooling Equilibrium

In the previous parts, we characterized equilibria when  $Y_H \geq Y_T$ . Unfortunately, if  $Y_H < Y_T$ , there no longer exists a separating PBE. Intuitively, even if Alice is willing to make a costly report of  $Y_H$  — which is the maximum value she would be willing to report after receiving the  $H$  signal — she cannot convince Bob that she will report her  $T$  signal truthfully since her costly report is not sufficient to offset her incentive to misreport the  $T$  signal, which is represented by the fact that  $Y_H < Y_T$ . Lemma 2 gives this result.

**Lemma 2.** *If  $Y_H < Y_T$ , then there does not exist a separating PBE.*

*Proof.* Suppose for contradiction that  $Y_H < Y_T$  and there exists a separating PBE. By definition,  $Y_H$  and  $Y_T$  are the maximum values that Alice might be willing to report after receiving the  $H$  or  $T$  signal respectively. At this separating PBE, suppose that Alice reports some  $r_A \in [f_{T,0}, 1]$  with positive probability after receiving the  $H$  signal. We must have  $r_A \leq Y_H$  by definition of  $Y_H$ . Since the PBE is separating, Bob's belief must be that  $\mu_{s_B, r_A}(H) = 1$  to be consistent with Alice's strategy. As we saw in the proof of Theorem 2, in any separating PBE, Bob's belief must be  $\mu_{s_B, f_{T,0}}(H) = 0$  and Alice must report  $f_{T,0}$  after receiving the  $T$  signal. Thus, because  $r_A \leq Y_H < Y_T$ , by the definition of  $Y_T$  Alice would strictly prefer to report  $r_A$  rather than  $f_{T,0}$  after receiving the  $T$  signal, which is a contradiction. We can also derive a contradiction for the case when  $r_A < f_{T,0}$ . The argument for this case is symmetric so we omit it.  $\square$

To further illustrate Alice's incentive to manipulate the market probability, we characterize a pooling equilibrium of this setting given a particular belief for Bob.

#### Bob's belief

For this equilibrium, we define Bob's belief which depends on  $\gamma$  that will be defined shortly.

$$\mu_{s_B, r_A}^P(H) = \begin{cases} g(\gamma), & \text{if } r_A \in [f_{H,0}, 1] \\ 0, & \text{if } r_A \in [0, f_{H,0}) \end{cases} \quad (11)$$

where

$$g(\gamma) = \frac{P(s_A = H|s_B)}{P(s_A = H|s_B) + (1 - P(s_A = H|s_B))\gamma}. \quad (12)$$

For this belief, Bob assumes that Alice received the  $T$  signal if her report is lower than what a truthful Alice would have reported after receiving the  $H$  signal ( $r_A < f_{H,0}$ ). Moreover, if Alice's report is greater than or equal to  $f_{H,0}$ , then Bob believes that Alice reports  $r_A$  with probability  $\gamma$  after receiving the  $T$  signal.  $\gamma$  is defined to be the maximum value within  $[0, 1]$  such that the following inequality is satisfied.

$$\begin{aligned}
L_T(f_{H,0}) &\leq E_{s_B}[Q(g(\gamma)f_{H,s_B} + (1 - g(\gamma))f_{T,s_B}) \\
&\quad - Q(f_{T,s_B}) \mid s_A = T] \quad (13)
\end{aligned}$$

First,  $\gamma$  is well defined. The RHS of equation (13) is strictly monotonically decreasing in  $\gamma$ . When  $\gamma = 0$ , the RHS reduces to  $L_T(Y_T)$ . Because  $f_{H,0} < Y_H < Y_T$ , we know that  $\gamma > 0$ .

#### PBE characterization

We now show in Theorem 4 that Bob's belief (11) and Alice's strategy (14) form a PBE of our game.

$$\sigma_H^P(f_{H,0}) = 1, \sigma_T^P(f_{H,0}) = \gamma, \sigma_T^P(f_{T,0}) = 1 - \gamma \quad (14)$$

**Theorem 4.** *Alice's strategy (14) and Bob's belief (11) form a pooling PBE.*

*Proof.* First, we show that Alice’s strategy (14) is optimal given Bob’s belief (11).

Given Bob’s belief, when Alice receives the  $H$  signal, she would optimally report  $f_{H,0}$ . This is because, reporting any value higher than  $f_{H,0}$  will increase her loss in the market and will not change the outside payoff, while reporting any value lower than  $f_{H,0}$  will reduce her outside payoff and increase her loss in the market.

When Alice receives the  $T$  signal, for any  $r_A \in [0, f_{H,0})$ , Alice maximizes her total payoff by reporting  $f_{T,0}$ . For any  $r_A \in [f_{H,0}, 1]$ , Alice maximizes her total payoff by reporting  $f_{H,0}$ . Thus, the support of Alice’s equilibrium strategy after receiving the  $T$  signal includes at most  $f_{T,0}$  and  $f_{H,0}$ . By the definition of  $\gamma$ , Alice is either indifferent between the two or strictly prefers to report  $f_{H,0}$ . Enforcing the consistency of Bob’s belief, we know that, after receiving the  $T$  signal, Alice’s optimal strategy must be reporting  $f_{H,0}$  with probability  $\gamma$  and reporting  $f_{T,0}$  with probability  $1 - \gamma$ . Note that  $\gamma > 0$  by definition of Bob’s belief (11).

Moreover, we can show that Bob’s belief is consistent with Alice’s strategy by mechanically applying Bayes’ rule (argument omitted). Given the above arguments, Alice’s strategy and Bob’s belief form a PBE of our game.  $\square$

### Uncertain Outside Incentive

In the previous section, we characterized a separating PBE with full information aggregation when  $\alpha = 1$  and  $Y_H \geq Y_T$ . In this section, however, we will show that any uncertainty about Alice’s outside incentive could be detrimental to information aggregation. This distrust arises when we allow  $\alpha \in (0, 1)$ , which introduces uncertainty about Alice’s incentive. In this case, even if the value of  $\alpha$  is common knowledge, information loss in equilibrium is inevitable if Alice has a sufficiently large outside incentive. In particular, when Alice has an outside payoff and has received the  $T$  signal, she can report  $f_{H,0}$  to pretend not to have the outside payoff and to have received the  $H$  signal. This results in these two types pooling, so the overall equilibrium is, at best, semi-separating and there is information loss.

**Theorem 5.** *If  $f_{H,0} < Y_T$  and  $\alpha \in (0, 1)$ , then there does not exist any PBE in which Alice’s type with the  $H$  signal and no outside payoff separates from her type with the  $T$  signal and the outside payoff.*

*Proof.* Proof by contradiction. Suppose that a separating PBE exists. At this separating PBE, with probability  $(1 - \alpha)$ , Alice reports  $f_{H,0}$  after receiving the  $H$  signal and reports  $f_{T,0}$  after receiving the  $T$  signal. To be consistent with Alice’s strategy, Bob’s belief on the equilibrium path must be  $\mu_{s_B, f_{H,0}}(H) = 1$  and  $\mu_{s_B, f_{T,0}}(H) = 0$ . Given this belief, however, when Alice has the outside payoff, she strictly prefers to report  $f_{H,0}$  after receiving the  $T$  signal since  $Y_T > f_{H,0}$ , which is a contradiction.  $\square$

### Conclusion and Future Direction

We study the strategic play of prediction market participants when there exist outside incentives. Our analysis brings out the insight that conflicting incentives inside and outside of

a prediction market do not necessarily damage information aggregation in equilibrium. In particular, under certain conditions, there are equilibria in which full information aggregation can be achieved. However, there are also many situations where information loss is inevitable.

In light of this, one important future direction is to better understand information aggregation mechanisms in the context of decision making, and design mechanisms to minimize or control potential loss in information aggregation and social welfare when there are conflicting incentives within and outside of the mechanism.

### Acknowledgments

This material is based upon work supported by NSF Grant No. CCF-0953516. Gao is partially supported by a NSERC Postgraduate Scholarship. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors alone.

### References

- Berg, J. E.; Forsythe, R.; Nelson, F. D.; and Rietz, T. A. 2001. Results from a dozen years of election futures markets research. In Plott, C. A., and Smith, V., eds., *Handbook of Experimental Economic Results*.
- Chen, Y., and Kash, I. A. 2011. Information elicitation for decision making. *AAMAS’11: Proc. of the 10th Int. Conf. on Autonomous Agents and Multiagent Systems*.
- Chen, Y., and Pennock, D. M. 2007. A utility framework for bounded-loss market makers. *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence* 49–56.
- Chen, Y.; Dimitrov, S.; Sami, R.; Reeves, D.; Pennock, D. M.; Hanson, R.; Fortnow, L.; and Gonen, R. 2010. Gaming prediction markets: Equilibrium strategies with a market maker. *Algorithmica* 58(4):930–969.
- Cho, I.-K., and Kreps, D. M. 1987. Signalling games and stable equilibria. *Quarterly Journal of Economics* 102:179–221.
- Dimitrov, S., and Sami, R. 2010. Composition of markets with conflicting incentives. *EC’10: Proc. of the 11th ACM Conf. on Electronic Commerce* 53–62.
- Fudenberg, D., and Tirole, J. 1991. *Game Theory*. MIT Press.
- Goel, S.; Reeves, D. M.; Watts, D. J.; and Pennock, D. M. 2010. Prediction without markets. In *EC ’10: Proceedings of the 11th ACM Conference on Electronic Commerce*, 357–366. New York, NY, USA: ACM.
- Hanson, R. 2007. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets* 1(1):3–15.
- Othman, A., and Sandholm, T. 2010. Decision rules and decision markets. *Proc. of the 9th Int. Conf. on Autonomous Agents and Multiagent Systems* 625–632.
- Shi, P.; Conitzer, V.; and Guo, M. 2009. Prediction mechanisms that do not incentivize undesirable actions. *WINE’09: Internet and Network Economics* 89–100.
- Spence, A. M. 1973. Job market signalling. *Quarterly Journal of Economics* 87(3):355–374.
- Wolfers, J., and Zitzewitz, E. 2004. Prediction markets. *Journal of Economic Perspective* 18(2):107–126.