What You Jointly Know Determines How You Act — Strategic Interactions in Prediction Markets

XI ALICE GAO, Harvard University
JIE ZHANG, Aarhus University
YILING CHEN, Harvard University

The primary goal of a prediction market is to elicit and aggregate information about some future event of interest. How well this goal is achieved depends on the behavior of self-interested market participants, which are crucially influenced by not only their private information but also their knowledge of others' private information, in other words, the information structure of market participants. In this paper, we model a prediction market using the now-classic *logarithmic market scoring rule* (LMSR) market maker as an extensive-form Bayesian game and aim to understand and characterize the game-theoretic equilibria of the market for different information structures. Prior work has shown that when participants' information is independent conditioned on the realized outcome of the event, the only type of equilibria in this setting has every participant race to honestly reveal their private information as soon as possible, which is the most desirable outcome for the market's goal of information aggregation. This paper considers the remaining two classes of information structures: participants' information being unconditionally independent (the I game) and participants' information being both conditionally and unconditionally dependent (the D game). We characterize the unique family of equilibria for the I game with finite number of participants and finite stages. At any equilibrium in this family, if player i's last stage of participation in the market is after player j's, player i only reveals his information after player j's last stage of participation. This suggests that players race to delay revealing their information, which is probably the least desirable outcome for the market's goal. We consider a special case of the D game and cast insights on possible equilibria if one exists.

Categories and Subject Descriptors: J.4 [**Social and Behavioral Sciences**]: Economics

General Terms: Economics, Theory

Additional Key Words and Phrases: Information aggregation; strategic analysis; prediction market; market scoring rule

1. INTRODUCTION

A prediction market for forecasting a random event allows market participants to express their probability assessments for possible outcomes of the event, typically by trading financial securities, and to be compensated if their assessment is more accurate than the previous market assessment. Participants thus have an economic incentive to improve the accuracy of the market assessment, and hence reveal their information. Moreover, by observing ac-

Authors' addresses: X. A. Gao and Y. Chen, Harvard School of Engineering and Applied Sciences, 33 Oxford Street, Cambridge, MA, USA; J. Zhang, Department of Computer Science, Aarhus University, IT-parken, Aabogade 34, DK 8200 Aarhus N, Denmark.
This work was partially supported by the National Science Foundation, under grant CCF-0953516. Xi Alice Gao was supported by the Canadian NSERC Postgraduate Scholarship. Jie Zhang acknowledges support from the Danish National Research Foundation and The National Science Foundation of China (under grant 61061130540) for the Sino-Danish Center for the Theory of Interactive Computation and from the Center for research in Foundations of Electronic Markets (CFEM), supported by the Danish Strategic Research Council. Any opinions, findings, conclusions, or recommendations expressed here are those of the authors alone.

tivities of other participants, a rational participant can infer some information from their activities and combine such information with his private information when trading in the market. Prediction markets rely on the economic incentives provided by the mechanism and the belief updating of participants to achieve their primary goal of eliciting and aggregating information about uncertain events of interest.

To this end, arguably we desire that participants reveal their private information *truthfully* and *immediately* in prediction markets. However, how well the information elicitation and aggregation goal is achieved depends on the strategic behavior of the self-interested market participants, which in turn is influenced by their private information and their knowledge of others' private information, what we formally call information structure of participants. In this paper, we model a prediction market as an extensive-form Bayesian game where each participant has a private signal and there is a joint distribution of the participants' signals and the event outcome, which is common knowledge to all participants. This joint distribution captures what participants know about each other's private information and is the information structure of the market game. The goal of this work is to understand and characterize game-theoretic equilibria of this market game given different information structures, with the hope to understand how and how quickly information is aggregated in the market.

Our prediction market uses Hanson's logarithmic market scoring rule (LMSR) [Hanson 2007], which is the de facto automated market maker mechanism for prediction markets. Because participants interact with the market maker, which is the mechanism per se and behaves deterministically, we only need to model the participants side of the market. This makes the generally challenging equilibrium analysis for extensive-form Bayesian games tractable for some information structures in our setting.

Prior work [Chen et al. 2007b, 2010] has shown that when participants' information is independent conditioned on the true outcome of the event, the only type of equilibria in this setting has every participant race to truthfully reveal all their information as soon as possible, which is the most desirable outcome for the market's goal. This paper considers the remaining two classes of information structures: participants' information being unconditionally independent (the I game) and participants' information being both conditionally and unconditionally dependent (the D game).

Our technical contributions include: (1) We characterize the unique family of equilibria for the I game with finite number of participants and finite stages. At any equilibrium in this family, if player i's last stage of participation in the market is after player j's, player i only reveals his information after player j's last stage of participation and on or before his own last stage of participation. This suggests that participants race to delay revealing their information, which is probably the least desirable outcome for the market's goal and is in stark contrast to the equilibria when participants' information is conditionally independent. (2) While it is generally challenging to characterize equilibria of extensive-form Bayesian games, we provide a systematic method for finding possible equilibrium strategies in a restricted 3-stage market game. With this method, we examine a restricted D game, where the information structure does not appear to have any characteristics that we can leverage, and are able to cast insights on possible equilibria if one exists. We also show that there exist D games that admit truthful equilibria.

Organization. The rest of the paper is organized as follows. We discuss related work in Section 1.1. Section 2 introduces our formal model of a prediction market game. We focus on a 3-stage market game with general information structures in Section 3, where we show how to more succinctly describe an equilibrium of the market game and provide a systematic method of finding possible equilibrium strategies of the game. Results in this section are building blocks for our subsequent analysis. In Section 4, we characterize all equilibria of the I game. Section 5 provides an exploration on possible equilibria of the D game if one exists and demonstrates the existence of a truthful equilibrium for some information structures in this class. We conclude in Section 6. Due to lack of space, the omitted proofs are provided in an appendix in the version of this paper available on the authors' websites.

1.1. Related Work

We model a prediction market as an extensive-form Bayesian game as in prior work [Chen et al. 2007b; Dimitrov and Sami 2007; Chen et al. 2010; Ostrovsky 2012]. Chen et al. [2010] considered both a finite-stage, finite-player and an infinite-stage, finite-player market game. They showed that when players' information is independent conditioned on the true state of the world, for both the finite- and infinite-stage games, there is a unique type of Perfect Bayesian Equilibria (PBE), where players reveal their information truthfully and as soon as they can. When players' information is (unconditionally) independent, they proved that the truthful play is not an equilibrium for both the finite- and infinite-stage games. An earlier work [Nikolova and Sami 2007] also presented an instance in which the truthful strategy is not optimal in an extensive-form game based on this market. However, whether a PBE exists when players have independent information was left as an open question. In this paper, we characterize all PBE of the finite-stage game with independent information and explore a special case of the setting when players' information is neither conditionally nor unconditionally independent.

Instead of characterizing equilibria, Ostrovsky [2012] studied whether information is fully aggregated *in the limit* at a PBE of an infinite-stage, finite-player market game with risk-neutral players. He characterized a condition under which the market price of a security converges to its expected value conditioned on all information with probability 1 at any PBE. Iyer et al. [2010] extended the setting to risk-averse players and characterized the condition for full information aggregation in the limit at any PBE. However, whether a PBE exists in such market games remains an open question.

The 3-stage version of our prediction market model resembles the ones studied by Dimitrov and Sami [2010] and Chen et al. [2011]: they both study 2-player games and the first player has another chance of participation after the second player's turn in the game. However, both Dimitrov and Sami [2010] and Chen et al. [2011] consider that the first player has utility for some event outside of the current market and the price in the current market influences the outcome of this event. In this paper, players only derive utilities from their trades in the market.

Jian and Sami [2010] studied market scoring rule prediction markets in a laboratory setting. In their experiment, participants may have conditionally or unconditionally independent information and the trading sequence may or may not be structured (a trading sequence is structured if it is pre-specified and is common knowledge to all participants). They confirmed previous theoretical predictions of the strategic behavior by Chen et al. [2010] when the trading sequence is structured. This study suggests that the behavior of participants in a prediction market critically depends on whether they reason about the other participants' private information. Moreover, there are some experimental and empirical studies on price manipulation in prediction markets using double auction mechanisms. The results are mixed, some giving evidence for the success of price manipulation [Hansen et al. 2004] and others showing the robustness of prediction markets to price manipulation [Camerer 1998; Hanson et al. 2007; Rhode and Strumpf 2004, 2007]. In the literature on financial markets, participants have been shown to manipulate market prices [Allen and Gale 1992; Chakraborty and Yilmaz 2004; Kumar and Seppi 1992].

2. MODEL OF THE MARKET GAME

We model a prediction market using an automated market maker mechanism as a Bayesian extensive-form game. Our setting is similar to that of these prior work [Chen et al. 2010, 2007b; Dimitrov and Sami 2007; Ostrovsky 2012].

The prediction market generates forecasts for a binary event with an outcome space $\Omega = \{Y, N\}$. Let $\omega \in \Omega$ denote the realized outcome of this event. Many real-world prediction markets focus on such binary events, for example "whether the UK economy will go into recession in 2013", "whether the movie Lincoln will win the Academy Award for Best Picture", and "whether a Democrat will win the US Presidential election in 2016".

2.1. Logarithmic Market Scoring Rule

The prediction market operates using a logarithmic market scoring rule (LMSR) [Hanson 2007], which is arguably the de facto automated market maker mechanism for prediction markets. In practice, an LMSR market often offers one contract for each outcome that pays off $1 if the corresponding outcome happens. The market maker (i.e. the mechanism) dynamically adjusts the contracts' prices as traders buy and/or sell the contracts. However, it is well known that this implementation is equivalent to a more abstract model where, instead of trading contracts and changing market prices, traders simply report probability estimates of event outcomes to the mechanism. In fact, Hanson [2007] introduced LMSR using this abstract model. In what follows, we will describe LMSR for our setting as a mechanism for changing probability estimates. Abstracting away the contracts makes subsequent analyses more tractable. We refer interested readers to Chen and Pennock [2007] and Abernethy et al. [2013] for more information on the equivalence of the two models.

An LMSR prediction market starts with some initial probability estimate r^0 for event outcome Y. (For a binary event, the probability of outcome N is implicitly $1 - r^0$, and such logic holds in the rest of the paper.) Players participate in the market in sequence and each player can change the current probability estimate to a new one of his choice. The market closes at a predefined time. After that, the realized outcome ω is observed and players receive their payoffs.

If a player reports estimate r^t when the current market estimate is r^{t-1}, his payoff for this report r^t is the scoring rule difference, $s(\omega, r^t) - s(\omega, r^{t-1})$, where $s(\omega, r)$ is the logarithmic scoring rule

$$s(\omega, r) = \begin{cases} b\log(r), & \text{if } \omega = Y \\ b\log(1-r), & \text{if } \omega = N, \end{cases}$$

and b is a parameter. A player may participate in the market multiple times. If T_i denotes the set of stages where player i participates, then player i's total payoff is the sum of the payoff for each of his reports, $\sum_{t\in T_i}(s(\omega, r^t) - s(\omega, r^{t-1}))$. We assume $b = 1$ without loss of generality, since b scales each player's payoff and does not have any effect on the players' strategic behavior in our setting.

The logarithmic scoring rule is one of many strictly proper scoring rules. All strictly proper scoring rules share a nice incentive property: $q = \arg\max_r (q\, s(Y, r) + (1-q)\, s(N, r))$. If a player is paid by a strictly proper scoring rule, then his expected score is uniquely maximized by honestly reporting his probability estimate. As a result, for a single report r^t, a risk-neutral player can maximize his expected payoff in an LMSR market by honestly reporting his probability estimate, because r^{t-1} is fixed for this player. However, if the player can participate multiple times, to maximize his *total* payoff, he may misreport his estimate in order to mislead other players and capitalize on their mistakes later on.

2.2. The Finite-Stage Market Game

The market game we study is an LMSR market with n stages and $m \le n$ players. The players participates in one or more stages of the market game, following a pre-defined sequence, which is common knowledge[1].

[1]It is an interesting future direction to consider a model where players endogenously choose when to participate. However, our equilibrium results for the D game with a pre-defined participation order imply that

Each player i has private information about the event given by a private signal $s_i \in S_i$ with signal space S_i and $|S_i| = n_i$. Each signal is only observed by the intended player. The prior distribution of the event outcomes and the players' private signals, denoted by $\mathcal{P} : \Omega \times S_1 \times \cdots \times S_m \to [0, 1]$, is common knowledge. Before the market starts, nature draws the realized event outcome and the private signals of the players according to $\mathcal{P}$.

The players are risk-neutral Bayesian agents. That is, the belief of the player participating in stage t can depend on the reported estimates in the first $t - 1$ stages as well as on his own private signal.

The 3-Stage Market Game. The simplest version of the market game that admits non-trivial strategic play is a 2-player 3-stage game. The two players are Alice and Bob, and the sequence of participation is Alice, Bob, and then Alice. We denote the signal spaces of Alice and Bob as $S_A = \{a_i : 0 \le i \le n_A - 1, n_A \in \mathbb{Z}^+\}$ and $S_B = \{b_j : 0 \le j \le n_B - 1, n_B \in \mathbb{Z}^+\}$ respectively. The analysis of this 3-stage market game will serve as building blocks for our analysis of the finite-stage market game.

2.3. Information Structure

The prior distribution $\mathcal{P}$ is a critical component of each instance of the market game. It encodes the relationship between the players' private signals and the event outcome, and it enables players with private signals to reason about other players' signals and the realized event outcome. We refer to $\mathcal{P}$ alternatively as the *"information structure"* of the market game. The primary goal of this paper is to characterize the strategic play in a market game in terms of its information structure.

2.3.1. Three Classes of Information Structures. There are three classes of information structures: conditionally independent (*CI game*), unconditionally independent (*I game*), and neither conditionally independent nor unconditionally independent (*D game*). These three classes are mutually exclusive and exhaustive. The first two types impose natural independence assumptions on the prior distribution $\mathcal{P}$, and they were first separately studied by Chen et al. [2007a] and Dimitrov and Sami [2007], and later in their joint work [Chen et al. 2010].

In a CI game, players' signals are independent conditioned on the realized event outcome. Prior work [Chen et al. 2007b, 2010] showed that there is a unique type of perfect Bayesian equilibria (PBE) for the CI game where players honestly report their estimates as early as possible. Thus, in this work, we focus on analyzing the I and D games.

For I games, players' signals are unconditionally independent from one another, but they are not independent of and may stochastically influence the event outcome. Formally, the prior distribution $\mathcal{P}$ for an I game must satisfy: $\Pr(s_i)\Pr(s_j) = \Pr(s_i, s_j), \forall s_i \in S_i, s_j \in S_j$ for any two players i and j. Dimitrov and Sami [2007] and Chen et al. [2010] showed that the I game does not have a truthful PBE where every player honestly reports his estimate as early as he can, but they left the existence of PBE as an open question.

To illustrate the I information structure, consider a stylized setting where each player independently observes a coin flip. The event to be predicted is some aggregate information about all of the independent coin flips, for example, whether more than 1/3 of the coin flips are heads. In this example, the players' signals are independent because the coin flips are independent events. A more realistic example involves a political election prediction market. Each voter independently obtains some private information about the election and decides on a vote, which is arguably independent from each other. The event we are interested in is the election outcome, which is determined by all of the votes. Finally, for an abstract

players will delay revealing their information as much as possible in the D game even with endogenously chosen participation order. We discuss these implications in section 4 after our equilibrium results.

example, each player's private information can be thought of as a single piece of a jigsaw puzzle, and the event being forecasted is related to the completed picture.

Even though the CI and I information structures capture events in some natural settings, they impose strong independence assumptions on the relationship between the players' private signals. Ideally, we would like to understand the players' strategic behavior in the market game without restricting to a particular information structure. For this reason, we study the D information structure consisting of signals that are neither conditionally independent nor unconditionally independent. In other words, the signals in a D game are both conditionally dependent and unconditionally dependent. Formally, a prior distribution $\mathcal{P}$ in a D game satisfies: $\exists s_i \in S_i, s_j \in S_j,$ s.t. $\Pr(s_i)\Pr(s_j) \neq \Pr(s_i, s_j)$ for two players i and j and $\exists s_{i'} \in S_i, s_{j'} \in S_j, \omega \in \Omega,$ s.t. $\Pr(s_{i'}, s_{j'}|\omega) \neq \Pr(s_{i'}|\omega)\Pr(s_{j'}|\omega)$ for two players i' and j'. It would be interesting to explore whether the D information structure could be further divided up into smaller classes with intuitive properties.

2.3.2. The Distinguishability Condition. To avoid degenerate cases in our analysis, we assume that the prior distribution $\mathcal{P}$ satisfies the following *distinguishability* condition, consisting of two parts.

Definition 2.1. The prior distribution $\mathcal{P}$ satisfies the distinguishability condition if for all i it satisfies inequality (1)

$$\Pr(Y|\mathbf{s}_{-i}, s_i) \neq \Pr(Y|\mathbf{s}_{-i}, s_i'), \forall \mathbf{s}_{-i} \in \mathbf{S}_{-i}, \forall s_i, s_i' \in S_i \cup \{\phi\}, s_i \neq s_i' \tag{1}$$

where $s_i = \phi$ means player i's private signal is not observed, and $\mathbf{S}_{-i} = \{S_1 \cup \{\phi\}\} \times \cdots \times \{S_{i-1} \cup \{\phi\}\} \times \{S_{i+1} \cup \{\phi\}\} \times \cdots \times \{S_m \cup \{\phi\}\}$, and inequality (2)

$$\sum_{s_i \in S_i} \mathrm{p}_{s_i} \Pr(Y|s_i, \mathbf{s}) \neq \sum_{s_i \in S_i} \mathrm{p}_{s_i} \Pr(Y|s_i, \mathbf{s}') \tag{2}$$

where $\mathbf{s} \neq \mathbf{s}'$ are any two different vectors of realized signals of any subset of players excluding i, and the vector $(p_{s_i})_{s_i \in S_i}$ is any probability distribution over S_i.

Inequality (1) generalizes the general informativeness condition by Chen et al. [2010]. The inequality is satisfied if different signal realizations of player i always lead to different posterior probabilities of $\omega = Y$, for any vector of realized signals for any subset of the other players (including unobserved signals). In other words, a player's signal always contains some information. Inequality (2) is similar to the distinguishability assumption used by Dimitrov and Sami [2010]. It requires that for any two realizations of signals of a subset of players, they lead to different estimates for outcome Y given any belief about player i's signal. This condition allows other players to infer the signals of the subset of players whenever they reveal their information truthfully.

While the distinguishability condition may be a nontrivial technical restriction, it allows us to focus on interesting strategic decisions in the game play without encountering degenerated cases.

2.4. Solution Concept

We use the *perfect Bayesian equilibrium* (PBE), which is informally a subgame perfect refinement of the Bayesian Nash equilibrium, as our solution concept. A PBE requires specifying each player's strategy given a realized signal at each stage of the game as well as the player's belief about the signals of players participating in all of the previous stages. The strategies and the beliefs of the players form a PBE of the market game if and only if, for each player, his strategy at every stage is optimal given the beliefs, and the beliefs are derived from the strategies using Bayes' rule whenever possible.

2.5. Terminologies for Players' Strategies

By properties of the logarithmic scoring rule, at a player's last chance to participate in the market, the player has the strictly dominant strategy of truthfully revealing his private information. So at any PBE, all private information is fully incorporated into the market estimate at the end of the market game. Thus, the focus of our analysis is on how quickly information gets incorporated into the market estimate throughout the game. In the following paragraphs, we distinguish between truthful and non-truthful strategies for a player in terms of when the player's private information is first revealed in the market game.

We use the term *truthful strategy* (also called truthful betting) to refer to the strategy where at a player's first chance to participate in the market, the player changes the market estimate to his posterior probability of Y given his signal and his belief about other players' signals. The truthful strategy fully reveals a player's private information as early as possible.

In contrast to the truthful strategy, a player may choose to misreport his information and manipulate the market estimate. For instance, a player can play a mixed strategy and reveal a noisy version of his signal to the subsequent players in the game. Alternatively, a player may try to withhold his private information from the other players by not changing the market estimate at all. Such non-truthful strategies hurt information aggregation in the market by causing the market estimate to contain inaccurate information at least temporarily.

3. THE 3-STAGE MARKET GAME WITH ANY INFORMATION STRUCTURE

Before diving into the PBE analysis of the finite-stage market game, we describe some preliminary analysis of the 3-stage market game with any information structure. In section 3.1, we justify that, in order to describe a PBE of the 3-stage market game, it suffices to describe Alice's strategy in the first stage and Bob's belief in the second stage. This allows us to greatly simplify our exposition in later analyses. Next, we prove a theorem in section 3.2, which allows us to systematic identify candidate PBE strategies for the players. This theorem gives us a useful method to make educated guesses about the possible PBE strategies in order to tackle the PBE existence question and to construct a PBE if one exists for the 3-stage market game with a given prior distribution. Finally, in section 3.3, we describe a consistency condition, which must be satisfied by a player's strategy in any PBE of the 3-stage game.

3.1. Describing PBE of the 3-Stage Market Game

We present a preliminary analysis of the 3-stage market game and introduce some notations for our later analyses.

In the 3-stage game, Alice and Bob observe their realized signals a_i and b_j respectively at the beginning of the market. In the first stage, Alice changes the market estimate for outcome Y from the initial market estimate r^0 to r_A of her choice. In the second stage, Bob observes Alice's first-stage report r_A and changes the market estimate to r_B. In the third stage, upon observing Bob's second-stage report r_B, Alice changes the market estimate from r_B to r^f, and then the market closes.

Alice's first-stage strategy is a mapping $\sigma : S_A \to \Delta([0,1])$ where $\Delta([0,1])$ is the set of probability distributions over $[0,1]$. For clarity of analysis and presentation, we assume that the support of Alice's first-stage strategy is finite. The results in this paper however hold even if the support of Alice's first-stage strategy is infinite. Let $\sigma_{a_i}(r_A)$ denote the probability that Alice reports r_A in the first stage after observing the signal a_i according to the strategy σ.

In the second stage, when Bob observes Alice's first-stage report r_A, he forms a belief about Alice's signals. Bob's belief specifies the likelihood that Alice received signal a_i when Alice reported r_A and Bob received signal b_j for any i and j. Let $\mu_{r_A,b_j}(a_i)$ denote the probability that Bob's belief assigns for Alice's a_i signal when Alice reported r_A and Bob received signal b_j. $\mu_{r_A,b_j}(a_i)$ is defined for any $r_A \in [0,1]$. At any PBE, we need to describe Bob's belief both on and off the equilibrium path. When r_A is in the support of Alice's first-stage PBE strategy, the game is on the equilibrium path and $\mu_{r_A,b_j}(a_i)$ is derived from Alice's strategy using Bayes' rule according to the PBE definition. However, when r_A is not in the support of Alice's equilibrium strategy, that is, the game is off the equilibrium path, $\mu_{r_A,b_j}(a_i)$ is still important for a PBE because the belief needs to ensure that Alice does not find it profitable to deviate from her PBE strategy. Off the equilibrium path, there are often more than one set of Bob's beliefs that can satisfy this requirement.

Bob only participates once, in the second stage of this game. By properties of strictly proper scoring rules, Bob has a strictly dominant strategy to report his posterior probability estimate of the event truthfully, given his belief. Thus, at any PBE, Bob must be using a pure strategy, which is fully determined by his belief, his signal, and Alice's first-stage report. Let $x_{b_j}(r_A)$ denote Bob's optimal report given his signal b_j and Alice's first-stage report r_A. At any PBE, Bob's optimal report $x_{b_j}(r_A)$ can be determined from his belief as follows:

$$x_{b_j}(r_A) = \sum_i \mu_{r_A,b_j}(a_i)\Pr(Y|a_i, b_j), \quad \forall\, b_j, 0 \le j \le n_B - 1, r_A \in [0,1].$$

In the third stage, Alice observes Bob's report and may change the market estimate again. At any PBE, knowing Bob's PBE strategy, Alice's belief on the equilibrium path can be derived from Bob's strategy using Bayes' rule. This is Alice's last stage of participation. Thus, by properties of strictly proper scoring rules, Alice has a strictly dominant strategy to report her probability estimate truthfully. Similar to Bob's strategy, Alice's third-stage strategy must be a pure strategy and it is fully determined by her belief, her signal, and Bob's report. We note that Alice's belief off the equilibrium path in the third stage is not important, because Bob has a dominant strategy in the second stage and will not deviate from it no matter what belief Alice has.

The above analysis shows that, to describe a PBE of the 3-stage market game, it suffices to specify Alice's strategy in the first stage and Bob's belief in the second stage. The rest of the strategic play is completely determined given them.

Moreover, for clarity in our analysis, we specify Bob's strategy rather than Bob's belief at a PBE. We can easily derive a belief of Bob such that Bob's strategy is optimal given it, shown as follows. First, Bob's strategy is valid if and only if $x_{b_j}(r_A) \in [\min_i\{\Pr(Y|a_i,b_j)\}, \max_i\{\Pr(Y|a_i,b_j)\}]$ for any b_j, because for any possible belief for Bob, his posterior probability should always fall into this interval. When r_A is in the support of Alice's PBE strategy, Bob's belief is derived from Alice's PBE strategy using Bayes' rule. When r_A is not in the support of Alice's PBE strategy, the PBE definition requires that Bob's belief be derived from a possible strategy for Alice using Bayes' rule. For such an r_A and for any b_j, we know that $\min_{a_i} \Pr(Y|a_i,b_j) \le x_{b_j}(r_A) \le \max_{a_i} \Pr(Y|a_i,b_j)$ holds and one of the two inequalities must be strict due to the distinguishability assumption. For a given b_j, let $a_{i'}$ and $a_{i''}$ be Alice's signal in $\min_{a_i} \Pr(Y|a_i,b_j)$ and $\max_{a_i} \Pr(Y|a_i,b_j)$ respectively. Then consider a possible strategy satisfying $\sigma_{a_{i'}}(r_A) = p$, $\sigma_{a_{i''}}(r_A) = 1-p$ and $\sigma_{a_i}(r_A) = 0$ for any other a_i, where $p = \frac{\Pr(a_{i''}|b_j)(x-\Pr(Y|a_{i''},b_j))}{\Pr(a_{i'}|b_j)(\Pr(Y|a_{i'},b_j)-x)+\Pr(a_{i''}|b_j)(x-\Pr(Y|a_{i''},b_j))}$. This strategy for Alice is valid, and thus we can derive Bob's off the equilibrium path belief for r_A from this strategy using Bayes' rule.

3.2. Systematically Identify Candidate PBE Strategies

To tackle the PBE existence problem and construct a PBE if one exists, it is essential that we make an educated guess of the players' possible PBE strategies. Theorem 3.1 below allows us to pinpoint a possible PBE strategy for Alice in the 3-stage game with any information structure, by comparing Alice's ex-ante expected total payoff (of both the first and the third stages) when using different first-stage strategies assuming that Bob knows and conditions on Alice's strategy.

For Theorem 3.1 below, for any of Alice's strategy σ_1, let $\pi_A(\sigma_1, \sigma_1)$ be Alice's ex-ante expected payoff when Alice uses strategy σ_1 in the first stage, Bob knows Alice's first-stage strategy σ_1 and conditions his belief on this strategy. This means that, for any r in the support of Alice's first-stage strategy σ_1, Bob's belief is derived from strategy σ_1 by using Bayes' rule. For any other r, there is no restriction on Bob's belief as long as it is valid.

In the proof of Theorem 3.1, we make an important distinction between a player's ex-ante and ex-interim expected payoff. A player's ex-ante expected payoff is his expected payoff without observing his signal, whereas his ex-interim expected payoff is his expected payoff given his signal.

THEOREM 3.1. *For the 3-stage market game, if two different first-stage strategies σ_1 and σ_2 for Alice satisfy inequality* (3)*, then strategy σ_2 cannot be part of any PBE of this game.*

$$\pi_A(\sigma_1, \sigma_1) > \pi_A(\sigma_2, \sigma_2) \tag{3}$$

PROOF. We prove this by contradiction. Suppose that two different first-stage strategies σ_1 and σ_2 for Alice satisfy inequality (3), and Alice's first-stage strategy σ_2 is part of a PBE of the 3-stage market game. Let μ_B denote Bob's belief at this PBE. μ_B specifies a distribution over Alice's signals for every possible first-stage report $r \in [0, 1]$ and any of Bob's signals b_j. Alice's ex-ante expected payoff at this PBE is $\pi_A(\sigma_2, \sigma_2)$. This proof holds for any valid belief for Bob at this PBE.

Suppose that Alice deviates from this PBE to play the strategy σ_1 in the first stage and Bob has the same belief μ_B as before. Let $\pi_A(\sigma_1, \sigma_2)$ denote Alice's total ex-ante expected payoff in the game at this deviation. The expression $\pi_A(\sigma_1, \sigma_2)$ is well defined since Alice knows Bob's belief and strategy at the original PBE. Similarly, let $\pi_B(\sigma_1, \sigma_2)$ denote Bob's ex-ante expected payoff in the second stage at this deviation.

At any PBE of this game, in the third stage, Alice can always infer Bob's signal given Bob's report by the distinguishability condition. So Alice always changes the market estimate to $\Pr(Y|a_i, b_j)$ in the third stage given Alice's signal a_i and Bob's signal b_j. Thus, the total expected payoff that Alice and Bob can get at any PBE of the 3-stage market game is

$$\pi_{AB} = \sum_{a_i, b_j} \left\{ \Pr(Y, a_i, b_j) \log \frac{\Pr(Y|a_i, b_j)}{r^0} + \Pr(N, a_i, b_j) \log \frac{\Pr(N|a_i, b_j)}{1 - r^0} \right\}$$

which is fixed given the initial probability r^0 and the prior distribution $\mathcal{P}$. Note that the above result holds not only at a PBE but whenever Bob reveals all of his information and Alice knowing his strategy maximizes her expected payoff. Therefore, by definition of π_{AB}, we must have

$$\pi_{AB} = \pi_A(\sigma_1, \sigma_2) + \pi_B(\sigma_1, \sigma_2), \forall \sigma_1, \sigma_2 \tag{4}$$

Inequality (3) is satisfied by assumption, so we have

$$\pi_A(\sigma_1, \sigma_1) > \pi_A(\sigma_2, \sigma_2)$$
$$\Rightarrow \pi_{AB} - \pi_A(\sigma_1, \sigma_1) < \pi_{AB} - \pi_A(\sigma_2, \sigma_2) \tag{5}$$
$$\Rightarrow \pi_B(\sigma_1, \sigma_1) < \pi_B(\sigma_2, \sigma_2) \tag{6}$$

where equation (5) is due to equation (4).

For a fixed first-stage strategy of Alice and for any belief of Bob, Bob's ex-ante expected payoff is maximized when his belief is derived from Alice's first-stage strategy using Bayes' rule. This can be proven as follows. When Bob's belief is derived from Alice's first-stage strategy by using Bayes' rule, then in the second stage, Bob changes the market estimate to $x_{b_j}(r_A)$ when Alice reports r_A in the first stage and Bob receives the b_j signal. Recall that by definition, $x_{b_j}(r_A) = \Pr(Y|r_A, b_j) = \sum_{a_i} \Pr(a_i|r_A, b_j)\Pr(Y|a_i, b_j)$. In this case, Bob's expected payoff in the second stage is

$$\sum_{b_j, r_A} \Pr(b_j, r_A) \left\{ x_{b_j}(r_A) \log \frac{x_{b_j}(r_A)}{r_A} + (1 - x_{b_j}(r_A)) \log \frac{1 - x_{b_j}(r_A)}{1 - r_A} \right\}. \tag{7}$$

When Bob has another belief, let $\hat{x}$ denote Bob's optimal report with this belief. Then Bob's expected payoff in the second stage is

$$\sum_{b_j, r_A} \Pr(b_j, r_A) \left\{ x_{b_j}(r_A) \log \frac{\hat{x}}{r_A} + (1 - x_{b_j}(r_A)) \log \frac{1 - \hat{x}}{1 - r_A} \right\}. \tag{8}$$

The difference in Bob's ex-ante expected payoff for the two different beliefs for Bob is (7) - (8):

$$\sum_{b_j, r_A} \Pr(b_j, r_A) \left\{ x_{b_j}(r_A) \log \frac{x_{b_j}(r_A)}{\hat{x}} + (1 - x_{b_j}(r_A)) \log \frac{1 - x_{b_j}(r_A)}{1 - \hat{x}} \right\}$$

which is nonnegative by properties of the relative entropy.

Therefore, for any two first-stage strategies σ_1 and σ_2 for Alice, we have shown that

$$\pi_B(\sigma_1, \sigma_2) \leq \pi_B(\sigma_1, \sigma_1) \tag{9}$$

Combining inequalities (6) and (9), we have

$$\begin{aligned} &\pi_B(\sigma_1, \sigma_2) < \pi_B(\sigma_2, \sigma_2) \\ \Rightarrow &\pi_{AB} - \pi_A(\sigma_1, \sigma_2) < \pi_{AB} - \pi_A(\sigma_2, \sigma_2) \\ \Rightarrow &\pi_A(\sigma_1, \sigma_2) > \pi_A(\sigma_2, \sigma_2). \end{aligned} \tag{10}$$

According to inequality (10), if Alice uses the first-stage strategy σ_2 at a PBE, then she can improve her ex-ante expected payoff by deviating to using the strategy σ_1. Then there must exist at least one realized signal for Alice, say a_i, such that Alice's ex-interim expected payoff after receiving the a_i signal is higher when she deviates to the strategy σ_1 than when she follows the strategy σ_2. (Otherwise, if Alice's ex-interim expected payoff for every realized signal is lower when she deviates to using the strategy σ_1 than when she follows the strategy σ_2, then her ex-ante expected payoff must also be lower when she deviates to using the strategy σ_1 than when she follows the strategy σ_2, contradicting inequality (10).) As a result, when Alice receives the a_i signal, she can improve her ex-interim expected payoff by deviating to using the strategy σ_1 and this contradicts with our assumption that Alice's first-stage strategy σ_2 is part of a PBE of the 3-stage market game. □

According to Theorem 3.1, to find Alice's possible PBE strategies for the 3-stage market game, it suffices to compare Alice's ex-ante expected payoffs for all possible first-stage strategies assuming Bob knows Alice's strategy, and only the strategies maximizing Alice's ex-ante expected payoff can possible be Alice's PBE strategy. This gives us a systematic way to identify possible PBE strategies without worrying about constructing Bob's off-equilibrium path beliefs.

3.3. The Consistency Condition

Our analyses of the 3-stage game frequently make use of a consistency condition described in Theorem 2 by Chen et al. [2010]. For completeness, we re-state this condition as a lemma below. The consistency condition requires that, at a PBE of the 3-stage game, for any r_A in the support of Alice's first-stage strategy σ, the posterior probability of Y given σ and

r_A should be equal to r_A. Intuitively, this requires that, Alice's first-stage strategy must not leave free payoff for Bob to claim in the second stage. If Alice's first-stage strategy does not satisfy the consistency condition, then Bob can get positive expected payoff simply by changing the market estimate to a value satisfying the consistency condition, and Bob can claim this positive expected payoff without having any private information about the event being predicted. This is contrary to Alice's goal of minimizing Bob's expected payoff since the 3-stage market game is a constant-sum game in expectation at any PBE.

LEMMA 3.2 (CONSISTENCY CONDITION FOR 3-STAGE MARKET GAME). *At a PBE of the 3-stage market game, if σ is Alice's first-stage strategy and r_A is in the support of strategy σ (i.e. $\exists a_i, \sigma_{a_i}(r_A) > 0$), then σ must satisfy the following consistency condition:*

$$Pr(Y|\sigma, r_A) = r_A$$

4. PBE OF THE FINITE-STAGE I GAME

We characterize all PBE of the finite-stage I game in this section. Our analysis begins with the 3-stage I game. Alice participates twice in the game, so she may have incentives to manipulate the market estimate in the first stage. We first identify a unique candidate PBE strategy for Alice by showing that if a PBE exists for the 3-stage I game, then Alice's first-stage strategy must be changing the market estimate to the prior probability of the event. This is equivalent to Alice delaying her participation until the third stage if the market starts with the prior probability of the event. We refer to this strategy as Alice's *delaying* strategy for the 3-stage I game. Alice's *delaying* strategy reveals absolutely no information to Bob about her signal. Next, we explicitly construct a PBE of the 3-stage I game in which Alice uses the *delaying* strategy in the first stage. These two results together imply that, the *delaying* PBE is unique for this game, in the sense that Alice must use the delaying strategy in every PBE of this game, even though Bob's belief can be different off the equilibrium path.

Given the *delaying* PBE of the 3-stage I game, we construct a family of PBE for the finite-stage I game using backward induction. Suppose that the players in the finite-stage I game are ordered by their last stages of participation. Then at every PBE of the finite-stage I game, each player i withholds his private information until after player $i-1$ finishes participating in the game, and then player i may truthfully reveal his private information in any of the subsequent stages in which he participates. In particular, there exists a particular PBE in this family where each player does not reveal any private information until his last stage of participation, and this is arguably the worst PBE of this game for the goal of information aggregation.

4.1. Delaying PBE of 3-stage I Game

We argue below that the delaying strategy is the only candidate PBE strategy for Alice in the 3-stage I game. Theorem 4.1 essentially proves that the delaying PBE of the 3-stage I game is unique with respect to Alice's strategy, if a PBE exists for this game. Part of the proof of Theorem 4.1 uses the argument in the proof of Theorem 2 in Chen et al. [2010].

THEOREM 4.1. *If the 3-stage I game has a PBE, then Alice's strategy at the PBE must be the delaying strategy, i.e. changing the market estimate to the prior probability of the event in the first stage.*

PROOF SKETCH. We first argue that if a PBE exists for the 3-stage I game, then Alice's first-stage strategy at this PBE must be a deterministic strategy. We show this by contradiction by assuming that there are at least two points in the support of Alice's first-stage PBE strategy. Then we construct another first-stage strategy achieving a better expected payoff for Alice, which means that the original strategy cannot be a PBE strategy by Theorem 3.1. By the consistency condition, if Alice's first-stage strategy is deterministic, it must be the strategy of changing the market estimate to the prior probability of the event. □

While the delaying strategy is the only possible PBE strategy for the 3-stage I game, we still don't know whether a PBE exists. In order for a PBE to exist, there must exist a belief of Bob to ensure that Alice does not find it profitable to deviate from the delaying strategy to any other strategy. Identifying such a belief for Bob can be challenging because essentially we need to specify what Bob will do upon observing every possible report of Alice in $[0, 1]$. In Theorem 4.2, we give an explicit construction of a PBE of the 3-stage I game in which Alice uses the *delaying* strategy in the first stage. At this PBE, Alice's first-stage strategy reveals no information to Bob about her private signal, and Bob's belief makes this delaying strategy the optimal choice for Alice.

THEOREM 4.2. *There exists a PBE of the 3-stage I game where Alice's first-stage strategy is*

$$\sigma_{a_i}(Pr(Y)) = 1, \quad \forall i = 0, ..., n_A - 1$$

and Bob's second-stage strategy is

$$x_{b_j}(r_A) = \begin{cases} f_j(\alpha_j^{\min}), & r_A \in [0, \alpha_j^{\min}) \\ f_j(r_A), & r_A \in [\alpha_j^{\min}, \alpha_j^{\max}] \\ f_j(\alpha_j^{\max}), & r_A \in (\alpha_j^{\max}, 1] \end{cases}, \quad \forall j = 0, ..., n_B - 1$$

where

$$\begin{aligned} f_j(r_A) &= \frac{Pr(Y|b_j)Pr(N)r_A}{Pr(Y)Pr(N|b_j) + (Pr(Y|b_j) - Pr(Y))r_A} \\ \beta_j^{\min} &= \min_{a_i}\{Pr(Y|a_i, b_j)\}, \beta_j^{\max} = \max_{a_i}\{Pr(Y|a_i, b_j)\} \\ \alpha_j^{\min} &= f_j^{-1}(\beta_j^{\min}), \alpha_j^{\max} = f_j^{-1}(\beta_j^{\max}) \end{aligned}$$

PROOF SKETCH. We describe the first part of the proof below showing that Bob's strategy is a valid PBE strategy.

First, Bob's belief on the equilibrium path is derived from Alice's first-stage strategy using Bayes' rule since $x_{b_j}(\Pr(Y)) = \Pr(Y|b_j)$. Moreover, for Bob's strategy to be a valid PBE strategy, it must satisfy $x_{b_j}(r_A) \in [\min_{a_i}\{\Pr(Y|a_i, b_j)\}, \max_{a_i}\{\Pr(Y|a_i, b_j)\}], \forall b_j, r_A \in [0, 1]$. To show this, note that by definition, $\beta_j^{\min} < \beta_j^{\max}$, $\alpha_j^{\min} < \alpha_j^{\max}$, and $f_j(r_A)$ is monotonically increasing in $r_A \in [0, 1]$ since

$$\frac{df_j(r_A)}{dr_A} = \frac{\Pr(Y)(1 - \Pr(Y))\Pr(Y|b_j)(1 - \Pr(Y|b_j))}{\{\Pr(Y)\Pr(N|b_j) + (\Pr(Y|b_j) - \Pr(Y))r_A\}^2} > 0$$

Hence the domain of $x_{b_j}(r_A)$ is well-defined. In addition, we have

$$\beta_j^{\min} = f_j(\alpha_j^{\min}) \le x_{b_j}(r_A) \le f_j(\alpha_j^{\max}) = \beta_j^{\max}, \forall r_A \in [0, 1].$$

Thus, Bob's strategy is valid. The rest of the proof then proves that Alice's delaying strategy is a best response to Bob's strategy. □

Based on Theorems 4.1 and 4.2 above, we have established both the existence and the uniqueness (with respect to Alice's first-stage strategy) of the PBE for the 3-stage I game.

4.2. A Family of PBE for the Finite-Stage I Game

We are ready to characterize the PBE of the finite-stage I game. By using backward induction and the delaying PBE of the 3-stage I game, we characterize a family of PBE of the finite-stage I game in Theorem 4.5. At any PBE in this family, players delay revealing their private information as much as possible.

We first generalize the consistency condition for the 3-stage game to the finite-stage game in Lemma 4.3. This consistency condition dictates that, for any stage k, the posterior probability of $\omega = Y$ given the participants' strategies and reports in the first k stages must be equal to the report of the participant in stage k at any PBE of this game.

LEMMA 4.3 (CONSISTENCY CONDITION FOR FINITE-STAGE MARKET GAME). *At a PBE of the finite-stage I game, suppose that σ^k and r^k are the strategy and the report for the participant of stage k respectively, then for every k, the participants' strategies and reports must satisfy equation* (11).

$$Pr(Y|r^1, \dots, r^k, \sigma^1, \dots, \sigma^k) = r^k \tag{11}$$

In Lemma 4.4 below, we analyze the *tail* of the finite-stage I game starting from the second-to-last stage of participation for the last player to the last stage of the game. The theorem shows that, in terms of strategic play, this portion of the finite-stage I game essentially reduces to a 3-stage I game. Thus, at any PBE, the last player chooses to not participate in the game in his second-to-last stage of participation. This key argument will be used repeatedly in the proof of the PBE of the finite-stage I game.

For Lemma 4.4 and Theorem 4.5, let the m players of the finite-stage I game be ordered by their last stages of participation. That is, for any $1 \leq i \leq m$, let t_i denote player i's last stage of participation, such that $t_i < t_j$ for any $1 \leq i < j \leq m$. Without loss of generality, we assume that player m has more than one stages of participations.

LEMMA 4.4. *Let stage k be the second to last stage of participation for player m ($k < t_m$). At any PBE of the finite-stage I game, player m does not change the market estimate in stage k.*

Finally, in Theorem 4.5, we prove the existence of a family of PBE of the finite-stage I game.

THEOREM 4.5. *At any PBE of the finite-stage I game, the players use the following strategies:*

— *From stage 1 to stage $t_1 - 1$, player 1 uses any strategy that satisfies the consistency condition. In stage t_1, player 1 truthfully reveals his signal.*
— *For any $2 \leq i \leq m-1$, from stage 1 to stage $t_{i-1} - 1$, player i does not participate in the game. From stage $t_{i-1} + 1$ to stage $t_i - 1$, player i uses any strategy that satisfies the consistency condition. In stage t_i, player i truthfully reveals his signal.*
— *From stage 1 to stage $t_m - 1$, player m does not participate in the game. In stage t_m, player m truthfully reveals his signal.*

PROOF SKETCH. We describe the argument for player m and $m-1$ here.

By properties of LMSR, player m truthfully reveals his signal in stage t_m, which is the last stage of the game. If stage t^* denotes the second to last stage of participation for player m, then the game from stage t^* to t_m can be reduced to a 3-stage I game (where player m is Alice and other players participating between t^* and t_m are a composite Bob). By Lemma 4.4, player m does not participate in stage t^*. Now remove this stage and let t^* be the *new* second to last stage of participation for player m, and the game from stage t^* to t_m again reduces to a 3-stage I game. Applying Lemma 4.4 again, we know that player m does not participate in stage t^* either. Inferring recursively, player m does not participate in any stage from 1 to $t_m - 1$.

For player $m-1$, he truthfully reveals his signal in stage t_{m-1} by properties of LMSR. From stage $t_{m-2} + 1$ to $t_{m-1} - 1$, player $m-1$ is the only participant because players 1 to $m-2$ already finished participating and player m does not participate by our earlier argument. Thus, player $m-1$ uses any strategy satisfying the consistency condition from stage $t_{m-2}+1$ to $t_{m-1}-1$. We combine the stages from $t_{m-2}+1$ to $t_{m-1}-1$ (denoted t^{**}) as the new last stage for player $m-1$. Let t^* be the new second to last stage of participation for player $m-1$, and note that $t^* < t_{m-2}$. Again, the game from stage t^* to t^{**} reduces to a 3-stage I game (where player $m-1$ is Alice). By Lemma 4.4, player $m-1$ does not participate in stage t^*. Inferring recursively, player $m-1$ does not participate in any stage from 1 to $t_{m-2}-1$. $\square$

To understand Theorem 4.5, consider dividing the finite-stage I game into m segments with player i being the *owner* of the segment from stage $t_{i-1}+1$ to stage t_i. At any PBE, each player does not participate in any stage before his segment, uses a strategy satisfying the consistency condition within his segment, and truthfully reveals his private signal at the last stage of his segment.

Figure 1 illustrates a particular PBE of a finite-stage I game. The letters A, B, and C denote the three players and their sequence of participation. A black letter means that the player truthfully reveals his signal in that stage. If the letter is gray, then the player uses a strategy satisfying the consistency condition. Note that the strategy of not changing the market estimate satisfies the consistency condition. A white letter means that the player is scheduled to participate but does not change the market estimate in that stage. The thick vertical bars mark the boundaries of the players' segments in the game.

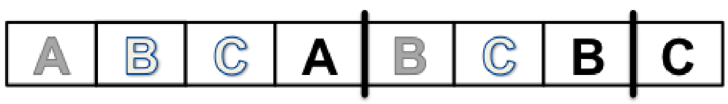

Fig. 1. A PBE of a Finite-Stage I Game with 3 players

The multiple PBE of the finite-stage I game differ by how early each player chooses to truthfully reveal his signal within his segment of the game. For the purpose of information aggregation, the best case is when every player chooses to truthfully reveal his signal in the first stage of his own segment. However, there exists a PBE where every player waits until the last stage of his segment to truthfully reveal his information, and this is arguably the worst PBE for the goal of information aggregation.

Although our model assumes a pre-specified participation order, our results still provide useful insights for the I game if the players endogenously choose when to participate in the game. Consider the I game with n stages and $m < n$ players where each player endogenously chooses in which stage to participate in the game. Our results for the I game suggest that, at any PBE all players will choose to delay their participation and no information is reveal in the first $n-m$ stages. The exact characterization of PBE would critically depend on how multiple trades submitted in the same stage are executed. This dependency is generally undesirable. Our assumption of pre-specified participation order circumvents this dependency and we believe our results still provide useful insights for players' behavior in this setting.

When comparing the PBE of the finite-stage I game with the truthful PBE of the finite-stage CI game [Chen et al. 2010], it is interesting to note how two different information structures can induce equilibrium behavior at the opposite ends of the spectrum: The players in the CI game race to reveal their private information as early as possible, whereas the players in the I game delay as much as possible to reveal their private information.

This difference is spiritually consistent with the concepts of complementarity and substitution of private signals defined by Chen et al. [2010]. Consider the ex ante expected payoffs of players. In the I games, players' private signals can be intuitively considered as complements. When the current market prediction is the prior probability, the sum of players' expected payoffs when each player reports a posterior probability conditioned only on his own private signal is strictly less than the total expected payoff that can be earned by reporting a posterior probability conditioned on all of the available private signals in any I game. This means that, every player in the I game prefers to wait for other players to make their reports first since observing more reports and thus inferring more signals improve the player's expected payoff. In contrast, in the CI games, players' private signals are substitutes. For any current market prediction, the sum of players' expected payoffs when each player reports a posterior probability conditioned only on his own private signal is strictly greater than the total expected payoff that can be earned by reporting a posterior probability conditioned on all of the available private signals. Thus, players prefer to race to capitalize on their private information early in the game.

5. THE 3-STAGE D GAME

The CI and I games admit two families of PBE that seem to lie at the two extremes of the spectrum: players race to reveal information early in the CI game, but race to withhold information in the I game. It is interesting to ask whether some instances of the D game may give rise to one of these two types of equilibria too. Yet, it is challenging to perform equilibrium analysis for the D game, because the dependency among the players' signals does not provide precise mathematical conditions that we can leverage.

Our goal in this section is moderate. We would like to explore a restricted 3-stage D game and obtain insights on what the players' PBE strategies may look like for this game if a PBE exists. We do not prove the existence of a PBE for this class. Nevertheless, we provide a sufficient condition for the prior distribution, which guarantees the existence of a truthful PBE for the D game. We also provide an example distribution that satisfies this condition.

In this section, we consider the 3-stage D game where Alice's private signal has only 2 realizations ($n_A = 2$).

5.1. An Expression for Alice's Ex-Interim Expected Payoff

We derive an expression for Alice's ex-interim expected payoff at any PBE of the 3-stage market game (denoted $u_{a_i}(r)$), for a given signal a_i and a particular first-stage report r. The purpose of discussing this expression is two fold. First, given $u_{a_i}(r)$, Alice's ex-ante expected payoff by using a particular strategy can be easily calculated and used to identify Alice's candidate PBE strategies by Theorem 3.1. Second, to construct a PBE of the market game, it suffices to check that the requirements of a PBE are satisfied using $u_{a_i}(r)$. Thus our discussion of this expression prepares us for developing the results in the following two subsections.

When deriving the expression of $u_{a_i}(r)$, we assume that Alice's first-stage payoff satisfies the consistency condition, Alice and Bob know each other's strategies and beliefs, and mostly importantly Bob's belief for any Alice's report r is derived as if the belief is on the equilibrium path for any given r. That is, for any Alice's report r, Bob's belief for r is derived from Alice's strategy using Bayes' rule as if the report r is in the support of Alice's first-stage strategy. The expression of $u_{a_i}(r)$ is given below. The complete derivation is included in the Appendix.

$$
\begin{aligned}
u_{a_i}(r) =& \Pr(Y|a_i)\log\frac{r}{\Pr(Y)} + \Pr(N|a_i)\log\frac{1-r}{1-\Pr(Y)} \\
&+ \sum_j \left\{ \Pr(Y, b_j|a_i)\log\frac{\Pr(Y|a_i, b_j)}{x_{b_j}(r)} + \Pr(N, b_j|a_i)\log\frac{\Pr(N|a_i, b_j)}{1-x_{b_j}(r)} \right\}
\end{aligned} \tag{12}
$$

where $x_{b_j}(r)$ is

$$x_{b_j}(r) = \frac{\Pr(Y, b_j|a_0)(\Pr(Y|a_1) - r) + \Pr(Y, b_j|a_1)(r - \Pr(Y|a_0))}{\Pr(b_j|a_0)(\Pr(Y|a_1) - r) + \Pr(b_j|a_1)(r - \Pr(Y|a_0))}$$

5.2. Three Candidate PBE Strategies for Alice

We identify three candidate PBE strategies for Alice in the 3-stage D game. These three strategies are the truthful strategy, the delaying strategy, and a mixed strategy in which Alice makes a deterministic report r for one realized signal and she mixes between reporting r and reporting her true posterior probability estimate for the other realized signal.

THEOREM 5.1. *If there exists a PBE of the 3-stage D game, then Alice must play one of the following three strategies at the PBE* [2]*:*

— *the truthful strategy:* $\sigma_{a_i}(Pr(Y|a_i)) = 1, \forall i = 0, 1$
— *the delaying strategy:* $\sigma_{a_i}(Pr(Y)) = 1, \forall i = 0, 1$
— *the mixed strategy:*

$$\sigma_{a_i}(Pr(Y|a_i)) = 1 - p, \sigma_{a_i}(r) = p, \sigma_{a_{1-i}}(r) = 1 \tag{13}$$

where $p = \frac{Pr(a_{1-i})(r - Pr(Y|a_{1-i}))}{Pr(a_i)(Pr(Y|a_i) - r)}$ *and* $u_{a_i}(Pr(Y|a_i)) = u_{a_i}(r)$ *is satisfied for some* $r \in (\min_i Pr(Y|a_i), Pr(Y)) \cup (Pr(Y), \max_i Pr(Y|a_i))$*,* $\forall i = 0, 1$.

5.3. A Sufficient Condition for the Truthful PBE

When the information structure of a 3-stage D game satisfies a monotonicity condition, we show in Theorem 5.2 that there is a truthful PBE of this game. This monotonicity condition requires that, for a fixed $i = 0, 1$, Alice's ex-interim expected payoff $u_{a_i}(r)$ is monotonically decreasing as the value of r changes from $\Pr(Y|a_i)$ to $\Pr(Y|a_{1-i})$.

THEOREM 5.2. *If for any* $i = 0, 1$, $u_{a_i}(r)$ *is monotonically decreasing as the value of* r *changes from* $Pr(Y|a_i)$ *to* $Pr(Y|a_{1-i})$*, then there exists a PBE of the 3-stage D game where Alice's first-stage strategy is*

$$\sigma_{a_i}(Pr(Y|a_i)) = 1, \forall i = 0, 1$$

and Bob's second-stage strategy is

$$x_{b_j}(r) = \frac{Pr(Y, b_j|a_0)(Pr(Y|a_1) - r) + Pr(Y, b_j|a_1)(r - Pr(Y|a_0))}{Pr(b_j|a_0)(Pr(Y|a_1) - r) + Pr(b_j|a_1)(r - Pr(Y|a_0))}, \forall j = 0, ..., n_B - 1 \tag{14}$$

Next, we give an example of a D information structure satisfying the monotonicity condition above.

Example 5.3. Consider an instance of the 3-stage D game where the prior distribution $\mathcal{P}$ is given by the following table. In Table I below, each cell gives the value of $\Pr(\omega, s_A, s_B)$ for the corresponding realizations of ω, s_A, and s_B. This prior distribution satisfies the monotonicity condition specified in Theorem 5.2 because, as r increases from $\Pr(Y|a_0)$ to $\Pr(Y|a_1)$, $u_{a_0}(r)$ decreases and $u_{a_1}(r)$ increases.

[2] Technically, Alice's PBE strategy could be of the form $\sigma_{a_i}(\Pr(Y|a_i)) = 1 - p, \sigma_{a_i}(r) = p, \sigma_{a_{1-i}}(\Pr(Y|a_{1-i})) = 1 - q, \sigma_{a_{1-i}}(r) = q$, for some $p, q \in [0, 1]$, $r \in [\min_{a_i} \Pr(Y|a_i), \max_{a_i} \Pr(Y|a_i)]$. However, if there exists a PBE of a 3-stage D game where Alice plays this mixed strategy, then there also exists a truthful PBE for this game. So we include this strategy as a special case when the 3-stage D game has a truthful PBE.

Table I. An example prior distribution.

$\omega = Y$			$\omega = N$		
	a_0	a_1		a_0	a_1
b_0	0.15	0.2	b_0	0.2	0.05
b_1	0.05	0.05	b_1	0.25	0.05

6. CONCLUSION AND FUTURE WORK

We analyze how the dependency among the participants' private information affect their strategic behavior when trading in a prediction market. We model the logarithmic market scoring rule prediction market as an extensive-form Bayesian game, and characterize PBE of this game for different information structures of the market participants. When the participants have unconditionally independent private information (I game), we show that there exists a family of PBE for the market game with a finite number of players and a finite number of stages. At any PBE in this family, assuming that the players are ordered by their last stages of participation, each player does not participate in the game before the previous player's last stage of participation. There exists a PBE where every player waits until their last stage of participation to truthful reveal their information, and this is arguably the worst outcome with respect to information aggregation. A future research question is to determine whether a PBE exists for the I game with a finite number of players but an infinite number of stages.

We also study a restricted version of the market game with 2 players and 3 stages when the players' private information is neither conditionally independent nor unconditionally independent (D game). Our result narrows down the possible PBE strategies to three simple strategies if a PBE exists. We conjecture that, for any instance of the D game, there exists a PBE where the first participant plays one of these three strategies. For future work, we are interested in proving the existence of the PBE of the D game for any prior distribution, characterizing sufficient and necessary conditions for each type of PBE to exist, and exploring whether the PBE of the 3-stage game extends to the game with a finite or an infinite number of stages.

REFERENCES

Abernethy, J., Chen, Y., and Vaughan, J. W. 2013. Efficient market making via convex optimization, and a connection to online learning. In *ACM Transactions on Economics and Computation*. Vol. 1. ACM, New York, NY, USA.

Allen, F. and Gale, D. 1992. Stock-price manipulation. *The Review of Financial Studies 5,* 3, 503–529.

Camerer, C. F. 1998. Can asset markets be manipulated? A field experiment with racetrack betting. *Journal of Political Economy 106,* 3, 457–482.

Chakraborty, A. and Yilmaz, B. 2004. Manipulation in market order models. *Journal of Financial Markets 7,* 2, 187–206.

Chen, Y., Dimitrov, S., Sami, R., Reeves, D. M., Pennock, D. M., Hanson, R. D., Fortnow, L., and Gonen, R. 2010. Gaming prediction markets: Equilibrium strategies with a market maker. *Algorithmica 58*, 930–969.

Chen, Y., Fortnow, L., Nikolova, E., and Pennock, D. M. 2007a. Betting on permutations. In *EC '07: Proceedings of the 8th ACM conference on Electronic commerce.* ACM, New York, NY, USA, 326–335.

Chen, Y., Gao, X. A., Goldstein, R., and Kash, I. A. 2011. Market Manipulation with Outside Incentives. In *Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI'11).*

Chen, Y. and Pennock, D. M. 2007. A utility framework for bounded-loss market makers. In *UAI '07: Proceedings of the 23rd Conference on Uncertainty in Artificial*

Intelligence. 49–56.

CHEN, Y., REEVES, D. M., PENNOCK, D. M., HANSON, R. D., FORTNOW, L., AND GONEN, R. 2007b. Bluffing and strategic reticence in prediction markets. In *WINE'07: Proceedings of the 3rd international conference on Internet and network economics*. Springer-Verlag, Berlin, Heidelberg, 70–81.

DIMITROV, S. AND SAMI, R. 2007. Non-myopic strategies in prediction markets. In *The 2nd Workshop on Prediction Markets*.

DIMITROV, S. AND SAMI, R. 2010. Composition of markets with conflicting incentives. In *Proceedings of the 11th ACM conference on Electronic commerce (EC '10)*. ACM, New York, 53–62.

HANSEN, J., SCHMIDT, C., AND STROBEL, M. 2004. Manipulation in political stock markets - preconditions and evidence. *Applied Economics Letters 11,* 7, 459–463.

HANSON, R. D. 2007. Logarithmic market scoring rules for modular combinatorial information aggregation. *Journal of Prediction Markets 1,* 1, 1–15.

HANSON, R. D., OPREA, R., AND PORTER, D. 2007. Information aggregation and manipulation in an experimental market. *Journal of Economic Behavior and Organization 60,* 4, 449–459.

IYER, K., JOHARI, R., AND MOALLEMI, C. C. 2010. Information aggregation in smooth markets. In *EC '10: Proceedings of the 11th ACM Conference on Electronic Commerce*. ACM, New York, NY, USA, 199–205.

JIAN, L. AND SAMI, R. 2010. Aggregation and manipulation in prediction markets: effects of trading mechanism and information distribution. *EC'10: Proc. of the 11th ACM Conf. on Electronic Commerce*, 207–208.

KUMAR, P. AND SEPPI, D. J. 1992. Futures manipulation with "cash settlement". *Journal of Finance 47,* 4, 1485–1502.

NIKOLOVA, E. AND SAMI, R. 2007. A strategic model for information markets. In *EC '07: Proceedings of the 8th ACM conference on Electronic commerce*. ACM, New York, NY, USA, 316–325.

OSTROVSKY, M. 2012. Information aggregation in dynamic markets with strategic traders. *Econometrica 80,* 6, 2595–2647.

RHODE, P. W. AND STRUMPF, K. S. 2004. Historical presidential betting markets. *Journal of Economic Perspectives 18,* 2, 127–142.

RHODE, P. W. AND STRUMPF, K. S. 2007. Manipulating political stock markets: A field experiment and a century of observational data. Working Paper.