Decision Support Systems 104 (2017) 13–25

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

On a participation structure that ensures representative prices in prediction markets

Arthur Carvalho

Department of Information Systems and Analytics Farmer School of Business, Miami University, United States

ARTICLE INFO

Article history:
Received 22 May 2017
Received in revised form 24 September 2017
Accepted 27 September 2017
Available online 1 October 2017

Keywords:
Prediction markets
Market scoring rules
The wisdom of crowds
Equilibrium analysis

ABSTRACT

The logarithmic market scoring rule (LMSR) is now the *de facto* market-maker mechanism for prediction markets. We show how LMSR can have more representative final prices by simply imposing a participation structure where the market proceeds in rounds and, in each round, traders can only trade up to a fixed number of contracts. Focusing on markets over binary outcomes, we prove that under such a participation structure, the market price converges after a finite number of rounds to the median of traders' private information for an odd number of traders, and to a point in the median interval for an even number of traders. Thus, the final market price effectively represents all agents' private information since those equilibria are measures of central tendency. We also show that when traders use market price data to revise their private information, the aforementioned equilibrium prices do not change for a broad class of learning methods.

1. Introduction

We consider the task of eliciting information from crowd members (agents) so as to form a collective forecast, thus leveraging the so called wisdom of crowds. Following standard equilibrium analysis, we assume that agents hold potentially different private information (beliefs) regarding the outcomes of a future event. The goal is to design a *mechanism* that extracts and aggregates agents' private information and provides a collective forecast. Ideally, the resulting forecast should be equivalent to the omniscient forecast that has direct access to all agents' beliefs.

Prediction markets are mechanisms designed to elicit and aggregate agents' private information. A prediction market offers financial securities whose future payoffs are tied to outcomes of a future event. A canonical security used in prediction markets is known as *Arrow-Debreu contracts*. An Arrow-Debreu contract associated with a particular outcome pays off $1 if that outcome occurs, and $0 otherwise. Hence, a risk-neutral agent who believes that the probability that a particular outcome will occur is, say, p should be willing to buy (respectively, sell) the underlying Arrow-Debreu contract at any price below (respectively, above) $\$p$.

In standard prediction-market models, a group of agents is invited to trade Arrow-Debreu contracts however they please, and the final market prices are taken as the group's estimate of how likely it is that each future outcome will occur. At the time of writing, it is estimated that over 100 organizations have run internal prediction markets, including companies such as Microsoft, General Electric, and Google [1]. Prediction markets have been successfully used to predict the result of presidential elections [2], sales forecasting [3], film awards [4], among many other applications. A reason for such a popularity is that prediction markets seem to yield equal or better forecasts than other methods of information aggregation [2,3].

The classic way of organizing a prediction market is by means of a *continuous double auction* (CDA), where the market maintains a list of *bid* and *ask* prices. Bids represent offers to buy securities, whereas asks represent offers to sell securities. Agents place their orders asynchronously, and a trade takes place when there is a match between agents' offers. There are three major problems with CDA markets. First, agents may not be willing to wait indefinitely for partners to trade with because offers that wait too long before being accepted might suffer adverse selection from new information. Second, the spread between the lowest ask price and the highest bid price can be wide. Consequently, substantial trading activity might only be expected for a limited set of popular securities. Finally, an informed agent might not want to reveal his willingness to trade at a given price since this might tip off other agents and, consequently, prevent the former from making a profit. This phenomenon is related to the so called *no-trade theorems* [5].

To increase trading volume, a market designer can add *market makers* to the market, *i.e.*, traders who are always ready to trade. Even a single agent can trade with a market maker as long as the former accepts the latter's offers. Algorithmic traders, called *automated market makers*, provide a way to automate the trading and

E-mail address: arthur.carvalho@miamioh.edu.

https://doi.org/10.1016/j.dss.2017.09.008
0167-9236/

pricing process. We argue that an automated market maker should satisfy the following properties: 1) *Bounded loss*: the automated market maker should run on a bounded loss; 2) *Tractability*: computing new prices should be a computationally tractable problem; and 3) *Representativeness*: the final market prices should represent agents' private information.

The first property is economically intuitive. The automated market maker is essentially subsidizing the market by always being ready to trade. Hence, it is possible to lose money. For the sake of limiting its loss, it is essential for an automated market maker to adjust its bid and ask prices after every trade. The second property is related to computational complexity. Since prediction markets are dynamic environments, the automated market maker must not spend a considerable amount of time computing the market prices. Finally, the last property reflects the main goal of a prediction market, namely to provide a forecast that accurately represents agents' private information.

Arguably, the most famous family of automated market makers is the *market scoring rules* [6]. Its most prominent member, the *logarithmic market scoring rule* (LMSR), has become the canonical automated market maker for prediction markets, being used by companies such as Inkling Markets, Consensus Point, Yahoo!, and Microsoft [7]. LMSR has bounded loss, and it computes market prices in polynomial time, hence satisfying two of the previously mentioned properties. However, LMSR alone provides no theoretical guarantees regarding representativeness. Concerning that point, Robin Hanson, the inventor of MSRs, said:

> "*Market scoring rules produce consensus estimates in the same way that betting markets produce consensus estimates. While each person is always free to change the current estimate, doing so requires taking on more risk, and eventually everyone reaches a limit where they do not want to make further changes, at least not until they receive further information. At this point the market can be said to be in equilibrium.*" [8, p. 5]

However, no formal characterization of such an equilibrium was proposed. Depending on the frequency with which agents interact with the market and/or their budgets, one can easily show that LMSR's final prices do not necessarily represent all agents' private information. This point raises important concerns regarding the forecast accuracy of LMSR, *e.g.*, that the market prices might be manipulated by wealthy agents with ulterior motives.

We show how LMSR can achieve the representativeness property by imposing a participation structure where the market proceeds in rounds and, in each round, agents can only trade up to a certain number of Arrow-Debreu contracts. Focusing on markets over binary outcomes, we prove that the market price converges after a finite number of rounds to the median of agents' private information for an odd number of agents, and to a point in the median interval for an even number of agents. Thus, the final market price effectively represents all agents' private information since the equilibrium prices (median and a point in the median interval) are measures of central tendency.

The exact number of rounds required for the convergence of the market price in our model is complexly determined by parameters of LMSR. To simplify this, we suggest a binary-search algorithm to reset the market price in the beginning of each trading round. As a consequence, the market price converges to a point inside an interval containing an equilibrium price, and the size of this interval linearly decreases with the number of rounds. This result does not depend on parameters of the market model. Since the final market price is guaranteed to be within a certain distance from an equilibrium price given any number of rounds, the proposed algorithm then provides an answer to an important open question in prediction markets, namely for how long one should run a prediction market in order to obtain representative market prices.

We finally investigate the influence of agents' learning on our convergence results. We show that when agents use market price data to revise their beliefs at the end of each trading round, the equilibrium price does not change for a broad class of learning methods.

To summarize, we suggest a participation structure where agents can only buy/sell up to a pre-defined number of contracts per round from LMSR. Such a simple structure ensures representative prices in that the final market prices from LMSR converge after a finite number of rounds to agents' median belief/interval (Section 3). Moreover, the participation structure allows for a very natural way of estimating for how long one must run the prediction market in order to obtain meaningful results (Section 4). From a budget perspective, our results are desirable in that the market maker is expected to spend less money in a worst-case scenario than it would without imposing the participation structure (Section 5). Finally, our results are robust in that they are still valid even when agents update their beliefs by learning from the market prices (Section 6).

2. Related work

As previously mentioned, the traditional way of organizing a prediction market has been via continuous double auctions (CDA). The theoretical underpinnings of using CDA market prices as reliable forecasts of future events are provided by the theory of rational expectations. Rooted in this theory, the efficient market hypothesis implies that prices in prediction markets can fully aggregate all individual agents' private information. Paradoxically, the theory of rational expectations allows for situations where the final market prices reflect all the agents' private information, but a single agent would never be willing to trade, a phenomenon often referred to as *no-trade theorems* [5]. In detail, consider a rational expectation equilibrium, meaning that agents collectively use all the relevant information when forming expectations regarding market prices. This implies that in equilibrium, the market prices are optimal in that they reflect all possible relevant information (it is noteworthy that our results do not rely on this optimality assumption and that we model agents' beliefs differently). Further assume that agents are risk averse and common knowledge holds true, meaning that all agents agree that agreed-upon trades are feasibly and mutually acceptable. Under these circumstances, one could expect that an agent receiving new information on the event of interest would try to leverage and profit from that information by trading in the market. However, the no-trade theorem asserts that no incentives for such a trading exist. This happens because "*... the mere willingness of the other traders to accept their parts of the bet is evidence to at least one trader that his own part is unfavorable. Hence no trade can be found that is acceptable to all traders*" [5, page 18].

Manski [9] showed that the equilibrium price of a CDA prediction market offering Arrow-Debreu contracts depends on the conjunction of agents' budgets and beliefs, rather than on their beliefs alone. In the same vein, Storkey [10] showed that the prices in prediction markets composed by agents with logarithmic and exponential utility functions produce equilibrium prices corresponding to, respectively, the wealth weighted mean and the geometric mean of agents' beliefs.

Interestingly, the equilibrium price in all the aforementioned work depends on agents' wealth. This point raises the question on whether the market prices can be manipulated by wealthy agents with ulterior motives. Deck et al. [11] presented evidence indicating that well-funded manipulators can in fact destroy a CDA prediction market's ability to aggregate informative prices.

The aforementioned points are all related to CDA prediction markets, whereas our work assumes the existence of an automated market maker. When looking for the existence of equilibrium prices in CDA-like markets, one must consider market clearing conditions, *i.e.*, prices where the supply of contracts matches the quantity demanded. Market clearing conditions make little sense in the presence of market makers since there is always an agent willing to trade for a certain price. Hence, instead of looking for market clearing conditions, one must look at where a potentially infinite sequence of trades leads the market price to. In this regard, our main contribution is the suggestion of a market participation structure that guarantees that the market price in the presence of a certain market maker (LMSR) will always converge to a point that represents all agents' beliefs.

The literature on the convergence of market prices in the presence of market makers is rather scarce. Ostrovsky [12] proved that there exists an equilibrium point where the market price of a prediction market based on a market scoring rule fully aggregates all agents' beliefs. Similar to our model, Ostrovsky's model assumes that agents trade with the maker maker in rounds. We argue, however, that our model is more realistic than Ostrovsky's model. For example, our model does not impose any restrictions on the order of trading, whereas Ostrovsky's model assumes that agents always follow a predefined trading order. Furthermore, Ostrovsky's main result is that an equilibrium price exists, but it is unclear how this price looks like. Our main result, on the other hand, is very constructive in nature, and we fully characterize how the equilibrium price looks like.

Another closely related work is the study by Sethi and Vaughan [13]. These authors investigated the convergence of market prices under market scoring rules for binary securities traded by myopic, risk-averse agents. Different than our work, the authors assumed no strict participation structure, such as rounds. Although they were able to prove that the market price does converge to an equilibrium price, the authors were not able to characterize such an equilibrium. Using very specific market parameters and utility functions, Sethi and Vaughan[13] showed via simulations that the market price in their model is relatively close to the weighted average of agents' beliefs, where the weights represent agents' budgets. From a theoretical perspective, our results are stronger than the results by Sethi and Vaughan [13] in that we formally characterize where the market prices converge to in our model. From a practical perspective, our suggested approximation algorithm determines how far from an equilibrium point the market price at the end of a round is. This can guide practitioners on for how long to run a market. This result is not possible in the model by Sethi and Vaughan [13] since it is not clear what an equilibrium price looks like.

Regarding participation structures, Jian and Sami [14] explored the impact of having structured participating order in prediction markets. The authors found that prediction markets structured to have trading rounds and a fixed sequence of trades exhibit greater accuracy of information aggregation than markets without such a participation structure. In our proposed model, we impose trading rounds, but we do not place any restrictions on trading sequences. Enforcing more structured participation has been shown to foster group-forecasting performance. For example, the experiments by Graefe and Armstrong [15] showed that methods that are based on structured rounds, including prediction markets, are more accurate than methods without such a structure.

3. Modeling markets and agents

We consider prediction markets offering Arrow-Debreu contracts on the occurrence of a binary outcome. For the sake of illustration, suppose that the prediction market is trying to predict the outcome of a presidential election featuring only two candidates, say A and B. If one agent holds one Arrow-Debreu contract associated with a certain outcome when the prediction market closes (say, candidate A will win the election), and that outcome turns out to be the observed outcome (say, candidate A indeed wins the election), then that agent receives one unit of numeraire (say, \$1). Otherwise, the agent receives nothing.

One can denote the prices in prediction markets over binary outcomes by a single value p inside the range $(0,1)$ because no-arbitrage conditions imply the price for the complementary outcome (*i.e.*, $1-p$). The market price should eventually represent a collective prediction of the occurrence of the underlying outcome. A total of n agents hold fixed private information (belief) regarding the occurrence of the outcomes. Similar to the market price, we represent the belief of each agent i by a single real value $f_i \in (0,1)$. One can interpret f_i as agent i's belief (subjective probability) regarding the occurrence of the outcome of interest. Without loss of generality, we assume that agents are indexed according to their beliefs, *i.e.*, $f_1 \leq f_2 \leq \cdots \leq f_n$. Let $\tilde{f}$ be the median of agents' beliefs. For an even number of agents, we denote the interval $\left[f_{\frac{n}{2}}, f_{\frac{n+2}{2}}\right]$ by the *median interval*.

We impose a market structure where the market proceeds in rounds and, in each round, agents can only trade securities with an automated market maker. We call $p_0^{(t)}$ and $p^{(t)}$, respectively, the automated market maker's market price in the beginning and at the end of round t. We use p to generally refer to the market price at any point during any round. In each round, each agent might trade with the market marker multiple times following any order. When doing so, agents compare the current market price to their beliefs and trade accordingly, as we elaborate on in Section 3.2. The major constraint in our model is that agents cannot hold more than a fixed number $y \in \Re_+$ of purchased/sold Arrow-Debreu contracts per round. Formally, let $g_i^{(t)}$ be the number of contracts traded by agent i during round t, where $g_i^{(t)}$ is initially zero. Agent i can then trade in the market during round t whenever he wants to, however he wants to, as long as $-y \leq g_i^{(t)} \leq y$, where negative values encode the number of (short) sold contracts. For example, when $y = 5$, agent i can then initially buy 3 contracts and later sell 2 contracts during the same round t since $g_i^{(t)} = 3 - 2 = 1 < 5$. However, that agent would not be able to initially short sell 6 or more contracts since $g_i^{(t)} = -6 < -5$. Note that short selling is allowed, meaning that an agent does not necessarily need to hold a contract in order to be able to sell it.

Agents interact with the market one at a time, repeatedly, in arbitrary order, and potentially multiple times. In each round, agents will continue to trade until trading is no longer profitable or doable due to the imposed market constraint. We formalize individual trading behavior in Section 3.2. A round is over when all the agents are no longer willing/able to trade since trading is unprofitable/impossible. At that point in time, the market maker starts a new round t, where the number of contracts each agent i holds in that new round, $g_i^{(t)}$, is reset to zero.

When the market closes and the true outcome is disclosed/observed, agents can then cash out their contracts. To do so, one sums the number of contracts each agent i holds at the end of each trading round, *i.e.*, $\sum_t g_i^{(t)}$. For the sake of illustration, consider the above presidential-election example, where the outcome of interest is whether candidate A will win the election. Suppose agent i purchases 3 contracts in the first round, and then 3 more contracts in the second round. Assume that the market closes after two trading rounds. If candidate A wins (respectively, loses) the election, agent i receives 6 (respectively, 0) units of numeraire. If, on the other hand, agent i buys 3 contracts during the first round, and sells 5 contracts during the second round, then agent i holds -2 contracts when the market closes. If candidate A wins (respectively, loses) the election, agent i then must pay the market maker 2 (respectively, 0) units of numeraire.

Our model allows agents to trade fractions of a contract (*e.g.*, 0.3 of an Arrow-Debreu contract), in which case the security pays off the equivalent fraction when the outcome of interest occurs (*e.g.*, 0.3 units of numeraire). Clearly, when an agent buys (respectively, sells) contracts from (respectively, to) the market maker, that agent must pay (respectively, receive) some money during the transaction. It is important to note that we impose no constraints on the number of contracts the market maker must buy/sell in each round. This would go against the definition of a market maker, namely an agent who is always willing to trade. We explain the pricing scheme used by the market maker in the following subsection.

Intuitively, the participation structure imposed on the market gives each agent equal power to drive the market price. In order to simplify our analysis, we assume that agents' budgets are large enough so that they can cover all their trades. We return to this issue when discussing the automated market maker's loss in Section 5. We also make the traditional assumptions that the true outcome will be observed at a certain point in the future, so that contracts can be cashed out, and that agents have no influence on that outcome, *i.e.*, the observed outcome does not depend on agents' actions. In the following subsections, we describe: 1) the logarithmic market scoring rule, which is used as the automated market maker in our model; 2) how we expect agents to trade with the market maker; and 3) our main price convergence result. For the sake of readability, we defer all the mathematical proofs in this paper to the appendix.

3.1. Logarithmic market scoring rule

Hanson [6] proposed a family of automated market makers known as *market scoring rules* (MSRs). A MSR always has a complete probability distribution (market prices) over the entire outcome space. Any agent can at any time change any part of that distribution as long as he agrees to both pay the scoring rule payment associated with the current probability distribution and receive the scoring rule payment associated with the new probability distribution. For example, if outcome x is realized, an agent who changes the probability distribution from $\mathbf{q}$ to $\mathbf{q}'$ pays $R(\mathbf{q}, x)$ and receives $R(\mathbf{q}', x)$, where R is a *proper scoring rule*. Proper scoring rules are functions that provide *ex ante* incentives for truthful reporting and that measure the informativeness of the reported distributions *ex post*, when the true outcome is observed [16].

The most prominent market scoring rule, called the *logarithmic market scoring rule* (LMSR), is defined when one sets $R(\mathbf{q}, x) = b\ln q_x$, where $q_x \in (0, 1]$ represents an estimate that outcome x will occur, and $b \in \mathbb{R}^+$ is a parameter that controls the liquidity of the market [6]. For small values of b (small liquidity), the market prices fluctuate wildly after every trade. Alternatively, the market prices move slowly with larger liquidity. Large liquidity is good for the traders since it stimulates more trades, but it comes at the expense of increasing the market maker's maximum loss exposure. Being an agent who is always ready to trade, any market maker is prone to losses since it is essentially subsidizing a market. Under the assumption that the initial market prices have the same value, *i.e.*, they are uniformly distributed, LMSR has a worst-case loss bounded by $b\ln m$, where m is the size of the outcome space [6]. In our setting, $m = 2$. This discussion illustrates that LMSR satisfies the bounded-loss property we discussed in Section 1.

Berg and Proebsting [17] derived the necessary formulae to implement a corresponding automated market maker in terms of buying and selling Arrow-Debreu contracts. For example, the total cost of buying $x \geq 0$ contracts whose current market price is p is:

$$C(b, p, x) = b\ln(p(e^{x/b} - 1) + 1) \tag{1}$$

The gain from selling contracts can be obtained by using negative values of x in Eq. (1) and, hence, negative costs encode sale proceeds earned by the trader. For any values of $b > 0$, $p \in (0, 1)$, and $x \geq 0$, the following relationships always hold true:

$$-C(b, p, -x) \leq xp \leq C(b, p, x) \tag{2}$$

i.e., the cost of buying x contracts is always greater than or equal to the current market price times x, which in turn is always greater than or equal to the profit made by an agent who sells x contracts. If this was not the case, then arbitrage opportunities would exist. For the sake of illustration, suppose that $b = 2$ and assume that the current market price is $p = 0.5$. Under these circumstances, the cost of purchasing one contract would be $C(2, 0.5, 1) \approx 0.56$, whereas the gain from selling one contract would be $-C(2, 0.5, -1) \approx 0.44$. The inequalities in Eq. (2) capture the fact that LMSR buys low and sells high according to its current "belief", *i.e.*, the current market price.

Naturally, the market price increases (respectively, decreases) whenever an agent buys (respectively, sells) contracts, *i.e.*, the market maker "updates its belief" after each trade. The resulting market price of a security after an agent buys x contracts having the current market price p is:

$$Q(b, p, x) = \frac{1}{1 + \frac{1/p - 1}{e^{x/b}}} \tag{3}$$

The resulting price when an agent sells contracts is obtained by using negative values of x in Eq. (3). Note that the above function can be easily and quickly calculated in real time, which makes LMSR very attractive in practice since it satisfies the tractability property we discussed in Section 1. Continuing with the above example, after an agent buys one contract, the new market price will be $Q(2, 0.5, 1) \approx 0.62$. If that agent instead sells one contract, the new market price will be $Q(2, 0.5, -1) \approx 0.38$. Given Eq. (3), it is not difficult to show that the number of contracts required to move the market price from $p \in (0, 1)$ to $p' \in (0, 1)$ is:

$$S(b, p, p') = b\ln\frac{p'(1 - p)}{p(1 - p')} \tag{4}$$

where positive (respectively, negative) values encode the number of contracts to be purchased (respectively, sold). Clearly, assuming that $b \neq \infty$, the number of contracts required to move the market price from any $p \in (0, 1)$ to any $p' \in (0, 1)$ is always finite.

An important property of LMSR, which is heavily used in this paper, is called *path independence* [6]. In words, it means that the cost of acquiring a bundle of securities is the same no matter how an agent splits up the transaction. Formally, given $x = x_1 + x_2$, then $C(b, p, x) = C(b, p, x_1) + C(b, Q(b, p, x_1), x_2)$. As a result, one can write the price function in Eq. (3) as:

$$Q(b, p, x) = Q(b, Q(b, p, x_1), x_2) \tag{5}$$

A consequence of the path-independence property in Eq. (5) is that the resulting market price after a sequence of trades $(x_1, x_2, \ldots, x_k)$ performed by any number of agents is equivalent to the market price after a single trade performed by a single agent, where $x = x_1 + x_2 + \cdots + x_k$ contracts are traded. For example, the resulting market price after one agent sells 5 contracts followed by a purchase of 7 contracts by another agent is equivalent to the market price after a single trade where a single agent buys $7 - 5 = 2$ contracts. Our theoretical results strongly rely on this observation.

To better understand the path independence property, consider a rather naive automated market maker that updates the current market price p after a purchase of x contracts according to the following function: $J(p, x) = \min\left(1 - p, \frac{x}{10}\right) * p + p$. Assume that an agent purchases a single contract in this market when the market price is $p = 0.5$. The resulting price after the purchase is then $J(0.5, 1) = 0.55$. Now, if that agent purchases another contract, the resulting market price will become $J(0.55, 1) = 0.605$. If, alternatively, the agent did not split up the purchase, the final price would be $J(0.5, 2) = 0.6$. Note that the final price is dependent on the underlying trading sequence/path. Let's now analyze a similar situation under LMSR with $b = 2$. Recall the price function in Eq. (3). When one agent purchases a contract, the resulting market price becomes $Q(2, 0.5, 1) \approx 0.622$. If that agent purchases another contract, then the market price becomes $Q(2, 0.622, 1) \approx 0.731$. Note that this latter value is equivalent to $Q(2, 0.5, 2) \approx 0.731$, *i.e.*, the final market price when the agent does not split up the purchase. The path independence property thus allows one to interpret all the transactions in a market as if they were performed by a single agent, regardless of the original sequence of transactions. We heavily exploit this perspective in our proofs. Clearly, CDA markets, whose final market prices reflect the last matched bid/ask offers, are path dependent.

3.2. Agents' trading behavior

We make two traditional assumptions regarding agents' trading behavior. First, we assume that they are *myopic*, *i.e.*, they do not reason about how their current trades might affect future trades and, consequently, future expected gains. There is some theoretical evidence that myopic behavior is optimal for markets based on LMSR, at least when traders' private information are independent conditional on the observed outcome [18]. We also assume that agents are *risk neutral*, meaning that they behave so as to maximize their expected returns. Without the risk-neutrality assumption, arbitrage opportunities will exist in the market (*e.g.*, see [19], Chapter 1). Together, the above assumptions mean that each agent trades so as to obtain the highest immediate expected return without thinking about future expected rewards. We expect these assumptions to reasonably capture the trading behavior of an average prediction market trader who might not necessarily employ advanced trading strategies, *e.g.*, trading in the opposite direction to his private information aiming at making a bigger profit by correcting the market prices later on.

Using the cost function in Eq. (1), agent i's expected profit in markets over binary outcomes when he buys x contracts whose current market price is p is equal to:

$$f_i \cdot (x - C(b, p, x)) + (1 - f_i) \cdot (-C(b, p, x)) \equiv x f_i - C(b, p, x) \quad (6)$$

i.e., it is equal to the agent's belief regarding the occurrence of the underlying outcome multiplied by the number of purchased contracts minus the cost of buying these contracts.

For such a trade to be profitable, the following relation must hold:

$$x f_i - C(b, p, x) > 0 \Rightarrow x f_i > C(b, p, x) \quad (7)$$

Equivalently, agent i's expected gain when he sells x contracts is equal to:

$$-C(b, p, -x) \cdot (1 - f_i) + (-C(b, p, -x) - x) \cdot f_i \equiv -C(b, p, -x) - x f_i \quad (8)$$

i.e., it is equal to the profit made by agent i minus his subjective probability regarding the occurrence of the underlying outcome multiplied by the number of sold contracts. For such a trade to be profitable, the following relation must hold:

$$-C(b, p, -x) - x f_i > 0 \Rightarrow x f_i < -C(b, p, -x) \quad (9)$$

Whenever agent i's belief is equal to the current market price, *i.e.*, $f_i = p$, the inequalities in Eqs. (7) and (9) become, respectively, $xp > C(b, p, x)$ and $xp < -C(b, p, -x)$. Thus, trading is never profitable when $f_i = p$ because Eqs. (7) and (9) violate the market inequalities in Eq. (2).

Whenever agent i's belief is greater than the current market price, *i.e.*, $f_i > p$, the inequality in Eq. (9) implies that $xp < -C(b, p, -x)$. Thus, selling contracts is never profitable because Eq. (2) is violated. However, buying contracts is always profitable as long as the inequality in Eq. (7) is not violated. Suppose that at certain round t, agent i's belief is greater than the current market price, $f_i > p$, and then agent i decides to purchase x contracts. A question that then arises is: how many contracts should agent i buy in order to maximize his expected profit without violating the imposed market constraint (*i.e.*, $-y \leq g_i^{(t)} \leq y$)? We answer this question in the following proposition.

Proposition 1. *When $f_i > p$, the number of contracts that maximizes agent i's expected gain in* Eq. (6) *without violating the market constraint $-y \leq g_i^{(t)} \leq y$ is $x = \min\left(S(b, p, f_i), y - g_i^{(t)}\right)$.*

Proof of Proposition 1. See Appendix A.

One can derive a similar result when agent i's belief is less than the current market price, *i.e.*, $f_i < p$. In this case, buying contracts is never profitable because the market inequalities in Eq. (2) would be violated. The optimal selling strategy is described in the following proposition.

Proposition 2. *When $f_i < p$, the number of contracts that maximizes agent i's expected gain in* Eq. (8) *without violating the market constraint $-y \leq g_i^{(t)} \leq y$ is $x = \max\left(S(b, p, f_i), -y - g_i^{(t)}\right)$.*

Proof of Proposition 2. See Appendix B.

To summarize, each risk-neutral and myopic agent maximizes his expected gain by driving the market price towards his belief. Buying is only and always profitable when an agent's belief is greater than the current market price. Selling is only and always profitable when an agent's belief is less than the current market price. Hence, in each trading round, each agent will trade until further trading is no longer possible due to a violation of the market constraint or because trading is no longer profitable due to that agent's belief being equal to the current market price.

3.3. Convergence results

We are now ready to present our main price convergence result. In particular, given the imposed market constraint and agents' trading behavior, for any positive liquidity parameter b, the market price converges after a finite number of rounds to the median of agents' beliefs for an odd number of agents, and to a point in the median interval for an even number of agents.

Theorem 1. *Consider a population of risk-neutral and myopic agents of size n. Given the market constraint $-y \leq g_i^{(t)} \leq y$, for any round $t \geq 1$ and every agent i, there exists a finite number of rounds z such that $\lim_{t \to z} p^{(t)} = \tilde{f}$, when $n \bmod 2 = 1$, and $\lim_{t \to z} p^{(t)} = \tilde{f}'$, when $n \bmod 2 = 0$, where $f_{n/2} \leq \tilde{f}' \leq f_{(n/2)+1}$.*

Proof of Theorem 1. See Appendix C.

It is interesting to note that Proposition 1 does not depend on specific values of LMSR's liquidity parameter, the maximum number of contracts an agent can hold per round (y), the initial market price $\left(p_0^{(1)}\right)$, the distribution of agents' beliefs, or any fixed participation order. A straightforward corollary of Proposition 1 is that the market price at the end of any round $t \geq 1$ is no farther from the median of agents' beliefs than the market price in the beginning of t.

Corollary 1. *Given the market constraint $-y \leq g_i^{(t)} \leq y$, the inequality $\left|\tilde{f} - p^{(t)}\right| \leq \left|\tilde{f} - p_0^{(t)}\right|$ holds true for any round $t \geq 1$ and every risk-neutral and myopic agent i.*

For an even number of agents, Theorem 1 means that the market price converges to the belief equal to $f_{n/2}$ (respectively, $f_{(n/2)+1}$) whenever the initial price $p_0^{(1)} \leq f_{n/2}$ (respectively, $p_0^{(1)} \geq f_{(n/2)+1}$). Then, in a worst-case scenario, the market price converges to any arbitrary point within the range $(0, 1)$ for an even number of agents, *e.g.*, when half of the population's belief is arbitrarily close to 0, the other half is arbitrarily close to 1, and the initial market price is $p_0^{(1)} \in (f_{n/2}, f_{(n/2)+1})$. One solution to this issue is to add one dummy agent with extreme private information when the original number of agents is even. If that agent's belief is equal to 0 (respectively, 1), then the final market price converges to $f_{n/2}$ (respectively, $f_{(n/2)+1}$). However, such a solution brings a new problem in that the market maker has to subsidize the new agent's trades. Another solution is to add an agent with belief equal to 0.5. If $f_{n/2} \leq 0.5 \leq f_{(n/2)+1}$, then the market price converges to 0.5. Otherwise, the market price converges to one of the median interval's extreme values.

Corollary 1 implies that one can declare that a market price is in equilibrium when the market prices at the end of two successive rounds are the same. To determine whether a trading round is over, one can simply monitor trading activity. Propositions 1 and 2 imply that an agent is always willing to trade contracts when the market price does not equal the agent's belief. When trading is no longer happening in a round, it then means that either agents can no longer trade because of the imposed market constraint or trading is no longer profitable for an agent because the market price is equal to the agent's belief. In practical applications, one can always predefine the amount of time each round will take. This was the approach implemented by Jian and Sami [14].

3.4. Numerical example

In this subsection, we illustrate Theorem 1 and the influence of different parameters on the market price's convergence speed. First, we run two simulated prediction markets for a total of 100 rounds using two different populations of size 5 and 6. The other parameters are set as follows: the liquidity parameter $b = 100$; the initial market price $p_0^{(1)} = 0.5$; and the maximum number of contracts each agent is allowed to hold per round $y = 5$. Agents' private information are uniformly distributed over the interval $[0.5, 0.99]$. At the end of each round $t \in \{1, \ldots, 100\}$, we plot the difference between the median belief and the market price, *i.e.*, $\tilde{f} - p^{(t)}$. Fig. 1 shows the obtained results.

As expected from Theorem 1, the market price monotonically converges after a finite number of rounds to the median belief for an odd number of agents ($n = 5$), and to the belief $f_{n/2}$ for an even number of agents ($n = 6$) because $p_0^{(1)} \leq f_{n/2}$. The convergence speed is faster during the initial rounds because all agents have beliefs greater than or equal to the market price, which implies that all of them buy contracts. As the number of rounds increases, the convergence speed slows down because a few agents start selling contracts while the majority of agents still buy contracts. Specifically, consider the case where $n = 5$ in Fig. 1. During the first three rounds, four agents have beliefs greater than the initial market price (0.5), whereas one agent has belief equal to 0.5. Hence, whenever one of the former agents buys a contract, the latter agent can make an expected profit by (short) selling a contract. Hence, the trades from two agents cancel out, and the market price is essentially driven by three agents. Recall that due to the path-independence property, the combined effect of two opposing trades on the market price is equivalent to no trade at all, which means that those trades are not effectively helping with the convergence of the market price. Now, consider the rounds 4 to 16. In this case, the market price has moved enough towards the median belief so that, in the beginning of each round, the same is greater than the private information of two agents. These two agents will sell contracts while the other three agents will buy, which implies that only one agent (the one with the median belief) is effectively changing the market price. In round 17, the market price converges to the median belief, and any purchase from an agent to the right of the median belief creates a profitable opportunity for an agent to the left of the median to sell a contract. This causes the market price to remain unchanged at the end of each round.

To understand the influence of the liquidity parameter b on the market price's convergence speed, we repeat the original experiment, but this time we set $b = 500$, five times higher than the original value. We show the obtained results in Fig. 2. Intuitively, the higher the value of b, the slower the market price converges to an

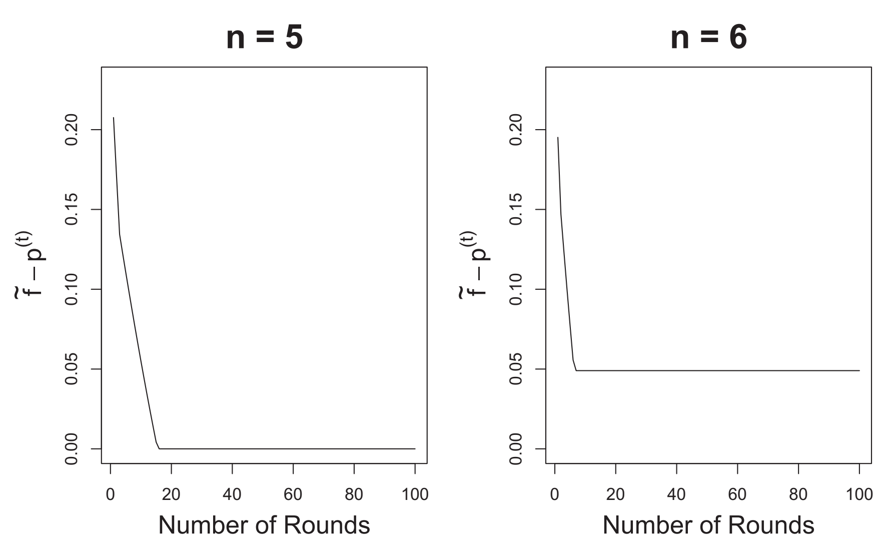

Fig. 1. Convergence of the market price under the proposed market model for $n = 5$ (left), $n = 6$ (right), $b = 100, p_0^{(1)} = 0.5, y = 5$. Agents' private information are uniformly distributed over the interval $[0.5, 0.99]$.

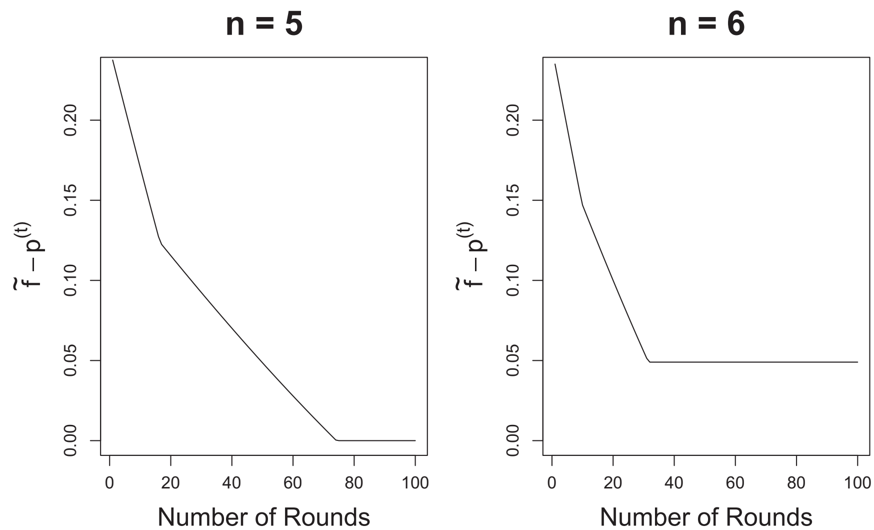

Fig. 2. Convergence of the market price under the proposed market model for $n = 5$ (left), $n = 6$ (right), $b = 500$, $p_0^{(1)} = 0.5$, $y = 5$. Agents' private information are uniformly distributed over the interval $[0.5, 0.99]$.

equilibrium. An opposite phenomenon occurs with y, the maximum number of Arrow-Debreu contracts agents are allowed to hold per round: the higher the value of y, the faster the market price convergences to an equilibrium. We note, however, that this analysis only takes into account the market prices at the end of successive rounds. In practice, increasing the value of y or decreasing the value of b might result in longer rounds, where agents trade more times before trading becomes either unprofitable or impossible due to the market constraint.

We also investigate the influence of the initial market price on the convergence speed. We repeat the original experiment, but this time we set $p_0^{(1)} = 0.1$, five times less than the original value. Naturally, it now takes more rounds for the market price to converge to an equilibrium since the initial price is more distant from the median belief than before. We show this result in Fig. 3.

Finally, to illustrate that our results are indeed robust when it comes to population size and distribution of beliefs, we run a simulated market for 100 rounds with the following parameters: $n = 51$, $b = 100$, and $y = 5$. We set the private information of five agents to the extreme value of 0. On the other end of the spectrum, we set the beliefs of 25 agents to the extreme value of 0.99. Finally, we set the beliefs of 20 agents to 0.2 and the median belief to 0.45. Fig. 4 shows the convergence results when the initial market price is $p_0^{(1)} = 0.1$ and $p_0^{(1)} = 0.9$. As expected from Theorem 1, the market price converges to the median belief (0.45) from above and below after a finite number of rounds, regardless of the distribution of agents' beliefs.

4. Approximate convergence

Section 3.4 illustrates that the market price's convergence speed depends on parameters of the market model. In this section, we propose an algorithm that provides additional theoretical guarantees regarding the price convergence. In particular, for any parameter values, the proposed algorithm guarantees that the final market price converges to a point inside an interval containing an equilibrium price after a number of rounds. As the number of rounds increases, the size of that interval decreases. Thus, the final market price is

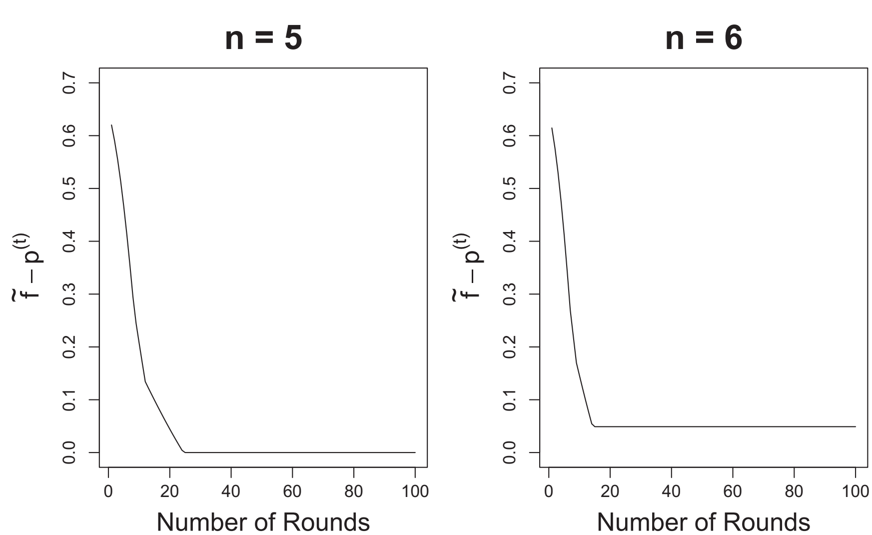

Fig. 3. Convergence of the market price under the proposed market model for $n = 5$ (left), $n = 6$, $b = 100$, $p_0^{(1)} = 0.1$, $y = 5$. Agents' private information are uniformly distributed on the interval $[0.5, 0.99]$.

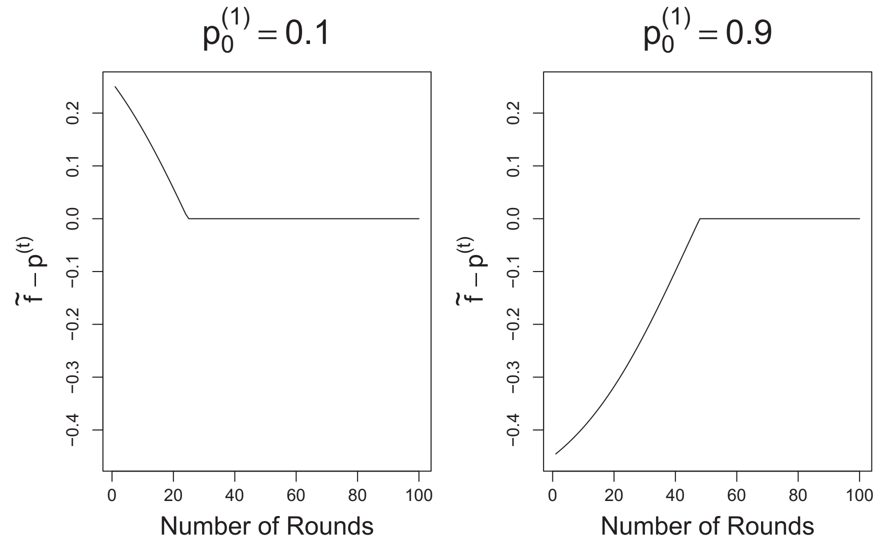

Fig. 4. Convergence of the market price under the proposed market model for $p_0^{(1)} = 0.1$ (left), $p_0^{(1)} = 0.9$, $n = 51, b = 100, y = 5$. The distribution of beliefs is as follows: 5 agents have beliefs equal to 0; 20 agents have beliefs equal to 0.2; 1 agent has belief equal to 0.45; and 25 agents have beliefs equal to 0.99.

guaranteed to be within a certain distance from an equilibrium price given any number of rounds.

The proposed algorithm is a form of binary search. It uses two extra variables throughout its execution: *lb* and *ub*. These variables store, respectively, a lower bound and an upper bound of the interval where an equilibrium price lies. Their initial values are, respectively, 0 and 1. We denote by *range* the interval $[lb, ub]$. Algorithm 1 shows the proposed binary search. First, the algorithm sets the initial market price in the beginning of each round to the range's midpoint. Thereafter, agents are invited to trade Arrow-Debreu contracts in the same way as before, *i.e.*, without violating the market constraint $-y \leq g_i^{(t)} \leq y$. When a round t is over, if the final market price is equal to the initial market price, *i.e.*, $p^{(t)} = p_0^{(t)}$, then the algorithm stops and returns $p^{(t)}$ as the equilibrium market price, which is equal to the median belief for an odd number of agents, and to a point in the median interval for an even number of agents. Otherwise, if $p^{(t)} > p_0^{(t)}$, then the algorithm updates the current lower bound by setting $lb = p_0^{(t)}$. Alternatively, if $p^{(t)} < p_0^{(t)}$, then the algorithm updates the current upper bound by setting $ub = p_0^{(t)}$.

Algorithm 1. A binary-search algorithm for finding a representative market price.

1: set $lb = 0$ and $ub = 1$
2: **for** $t = 1$ **to a number of rounds** T **do**
3: Set $p_0^{(t)} = \frac{ub+lb}{2}$ and invite agents to trade
4: **if** $p^{(t)} > p_0^{(t)}$ **then** $lb = p_0^{(t)}$
5: **else if** $p^{(t)} < p_0^{(t)}$ **then** $ub = p_0^{(t)}$
6: **else return** $p^{(t)}$
7: **return** $\frac{ub+lb}{2}$

4.1. Numerical example

A numerical example may clarify the mechanics of the proposed algorithm. Consider a population of size three ($n = 3$), where agents' private information are: $f_1 = 0.2, f_2 = 0.65$, and $f_3 = 0.7$. Further, assume that $b = 100$ and $y = 5$. It is important to remember that the lower and upper bounds relate to the median private information, as opposed to all agents' beliefs. In the beginning of the first round, the initial market price is $p_0^{(1)} = 0.5$ because $lb = 0$ and $ub = 1$. Agents are invited to trade contracts under the imposed market constraint. The resulting market price at the end of the first round is $p^{(1)} \approx 0.5125$ no matter the participation order since agents will trade until trading is no longer profitable or doable and, due to the path independence property, all the trades in a round can be interpreted as a single trade (see Section 3.1). Since $p^{(1)} > p_0^{(1)}$, we set $lb = p_0^{(1)} = 0.5$. In the beginning of the second round, the initial market price is set to $p_0^{(2)} = (lb + ub)/2 = (0.5 + 1)/2 = 0.75$. The resulting market price at the end of the second round is $p^{(2)} \approx 0.7208$. Since $p^{(2)} < p_0^{(2)}$, we set $ub = p_0^{(2)} = 0.75$. Thus, the difference between the market price after 2 rounds and an equilibrium price is at most $ub - lb = 0.75 - 0.5 = 0.25$. If the market closes after two trading rounds, the final market price is defined as $\frac{ub+lb}{2} = \frac{0.75+0.5}{2} = 0.625$. The absolute difference between the final market price and the median belief is then 0.025.

4.2. Correctness and convergence

As a consequence of the path-independence property in Eq. (5), if the market price at the end of a round t is greater than the market price in the beginning of round t, *i.e.*, $p^{(t)} > p_0^{(t)}$, then more than half of the agents have beliefs greater than $p_0^{(t)}$ and $\tilde{f} > p_0^{(t)}$. Thus, we can set the lower bound of the interval where an equilibrium price lies to $lb = p_0^{(t)}$ (line 4 in Algorithm 1). If the market price at the end of a round t is less than the market price in the beginning of round t, *i.e.*, $p^{(t)} < p_0^{(t)}$, then more than half of the agents have beliefs less than $p_0^{(t)}$ and $\tilde{f} < p_0^{(t)}$. Thus, we can set the upper bound of the interval where an equilibrium price lies to $ub = p_0^{(t)}$ (line 5 in Algorithm 1). Hence, the price movement indicates the direction where the median belief lies and, by construction, $lb \leq \tilde{f} \leq ub$. The market price at the end of a round is only equal to the market price in the beginning of that round when an equilibrium is reached (line 6 in Algorithm 1).

By moving the market price to the range's midpoint in the beginning of each round (line 3 in Algorithm 1), the algorithm guarantees that either *ub* decreases or *lb* increases as well as that the range is halved after each round. Thus, the ratio between the size of the ranges after two consecutive rounds is always 1/2. This observation provides a natural way for computing the range's length after a certain number of rounds T, assuming that an equilibrium price was not reached before: $H(T) = 0.5^T$. For example, the range's length is 0.03125 for $T = 5$, which is equivalent to say that the absolute difference between the market price at the end of the 5th round and an equilibrium price is at most 0.03125. The above function neither depends on the number of agents nor on the parameters b and y. It also implies that the market price linearly converges to an equilibrium under the proposed algorithm with the convergence rate equal to 0.5. By taking the inverse of the above function, one can compute the number of rounds required to obtain a desirable range length L:

$$H^{-1}(L) = \log_{0.5} L \tag{10}$$

The above function provides an answer to an important open question in prediction markets, namely for how long one should run a prediction market in order to obtain representative market prices. For example, if one wants the absolute difference between the final market price and an equilibrium price to be at most 0.05, then a total of $log_{0.5} 0.05 \approx 4.32$ rounds are needed.

Similar to Theorem 1, we show below that when using the proposed algorithm, the market price converges after a finite number of rounds to the median of agents' beliefs for an odd number of agents, and to a point in the median interval for an even number of agents.

Theorem 2. *Consider a population of risk-neutral and myopic agents of size n. Given the market constraint $-y \le g_i^{(t)} \le y$, for any round $t \ge 1$ and agent i, there exists a finite number of rounds z in which Algorithm 1 returns a market price $p^{(t)}$ so that $\lim_{t\to z} p^{(t)} = \tilde{f}$, when $n \bmod 2 = 1$, and $\lim_{t\to z} p^{(t)} = \tilde{f}'$, when $n \bmod 2 = 0$, where $f_{n/2} \le \tilde{f}' \le f_{(n/2)+1}$.*

Proof of Theorem 2. See Appendix D.

To summarize this section's results, the market price resulting from the proposed binary-search algorithm converges to an equilibrium price after a finite number of rounds. Moreover, the function in Eq. (10) provides guidelines on how to set the number of trading rounds so that if an equilibrium price is not reached after a number of rounds, then one can still guarantee that the resulting market price is within a certain distance from an equilibrium price.

5. The automated market maker's loss

In our model and analysis, we have assumed thus far that agents can either hold or (short) sell up to y Arrow-Debreu contracts per round. Hence, we have implicitly assumed that each agent has an available budget of at least $\$yT$, where T is the total number of rounds. Such a budget enables each agent to buy the maximum allowed number of Arrow-Debreu contracts per round as well as to cover the maximum allowed number of short-selling operations. Assuming the availability of such a budget is not necessarily a practical concern since many prediction markets use play money instead of real money due to a series of legal, technical, and ethical reasons [20]. In play-money markets, agents receive an initial amount of money to invest. A few of those agents with the largest net worth when the market closes win some prize. Servan-Schreiber et al. [20] showed that prediction markets using play money are as accurate as markets using real money.

Another practical issue concerns the automated market maker's loss. An automated market maker is expected to lose money since it essentially subsidizes a prediction market by always being ready to trade. Originally, LMSR's loss is bounded by $b\ln 2$ for markets over binary outcomes under the assumption that the initial market price is equal to 0.5 [6]. However, this bound is not valid anymore when using the algorithm proposed in Section 4 because the market maker changes the current market price in the beginning of each round.

We note that trading rounds provide a natural upper bound on the market marker's loss that does not depend on the liquidity parameter b. Assuming the unrealistic worst-case scenario where all the agents always pay nothing for a contract and receive one unit of numeraire for each held contract when the market closes, the market maker loses is then $\$ny$ per round, *i.e.*, the total number of agents times the maximum number of contracts that each agent is allowed to hold per round. Thus, the total cost of running the proposed model for a total of T rounds is at most $\$Tny$.

An interesting question that then arises is: how tight is the above upper bound when compared to the original LMSR's loss bound of $b\ln 2$? We start by noting that our proposed bound is tighter than the original bound whenever $b \ln 2 > Tny \Rightarrow b > \frac{Tny}{\ln 2}$. Unfortunately, there is no universally accepted way of setting the value of b. Berg and Proebsting [17] proposed an interesting approach that works as follows: one should set b in such a way that whenever n agents invest all their available budgets buying a contract associated with an outcome, then the resulting market price of the underlying security should equal a parameter $p_{upper} \in (0.5, 1)$, where p_{upper} represents the maximum price of a security. Now, instead of defining b directly, one can simply define p_{upper}, which is ideally close to one. Mathematically, for markets over binary outcomes, the initial market price $p_0^{(1)} = 0.5$, and a combined budget of K units of numeraire, Berg and Proebsting's equation is:

$$b = \frac{-K}{\ln(2 - 2p_{upper})} \tag{11}$$

For similar settings, *i.e.*, when $K = Tny$, our bound of Tny is tighter than the original LMSR's loss bound of $b\ln 2$ whenever b is set according to Eq. (11) and $p_{upper} > 0.75$. One can derive this result from simple arithmetic manipulations:

$$b \ln 2 > Tny \Rightarrow \frac{-Tny}{\ln(2 - 2p_{upper})} > \frac{Tny}{\ln 2} \Rightarrow -\ln 2 > \ln(2 - 2p_{upper})$$
$$\Rightarrow p_{upper} > 0.75$$

In practice, one should expect the proposed loss bound of Tny to be tighter than the original LMSR's loss bound of $b\ln 2$ with b set according to Eq. (11) since it is desirable to have the final market price, which is a probability value, as close to one as possible when all the available money is used to buy a single security since this indicates that agents' beliefs are all close to one.

6. Using price data to revise beliefs

Another assumption we have made throughout this paper is that agents hold fixed beliefs regarding the likelihood of the underlying outcomes. It is conjectured, however, that agents use price data to revise their beliefs [9]. We next show that our convergence results in Theorems 1 and 2 do not change when agents use market price data to revise their beliefs for any learning method where agents' prior and posterior beliefs have the same ordinal relationship to the market price.

In this section, we slightly change notation by representing the initial belief of each agent i in the beginning of each round t by $f_i^{(t)} \in (0, 1)$. We assume that agents learn when the market price is stable, *i.e.*, they update their beliefs at the end of each round. We make two additional assumptions regarding how agents learn. The first one, referred to as the *directional assumption*, means that each agent only learns when the market price is moving away from his belief. For example, consider the following initial and final market prices during the first round: $p_0^{(1)} = 0.3$ and $p^{(1)} = 0.7$. An agent

i with initial belief $f_i^{(1)} = 0.9$ will not update his belief in such a scenario because the market price is moving towards his belief, *i.e.*, $\left|f_i^{(1)} - p_0^{(1)}\right| > \left|f_i^{(1)} - p^{(1)}\right|$. Thus, $f_i^{(1)} = f_i^{(2)}$. More generally, learning happens when $\left|f_i^{(t)} - p_0^{(t)}\right| < \left|f_i^{(t)} - p^{(t)}\right|$, where $f_i^{(t)}$ is agent i's belief during round t.

The second assumption, referred to as the *magnitude assumption*, concerns how much agents learn. It means that $\text{sgn}\left(f_i^{(t)} - p^{(t)}\right) = \text{sgn}\left(f_i^{(t-1)} - p^{(t)}\right)$, where $\text{sgn}(\cdot)$ is the sign function that extracts the sign of a real number. In words, agents' prior and posterior beliefs have the same ordinal relationship to the market price at the end of a round. In other words, agents do not revise their beliefs beyond the market price, which is a reasonable assumption since we are assuming that the market price is the agents' source of new information.

We show below that our price convergence results in Theorems 1 and 2 remain unchanged under any learning model that satisfies both the directional and the magnitude assumptions.

Proposition 3. *Consider a population of risk-neutral and myopic agents of size n, where each agent updates his belief at the end of each round according to a learning model that satisfies both the directional and the magnitude assumptions. Given the market constraint* $-y \le g_i^{(t)} \le y$ *, for any round* $t \ge 1$ *and agent i, there exists a finite number of rounds z such that* $\lim_{t\to z} p^{(t)} = \tilde{f} = f^{(1)}_{(n+1)/2}$ *, when n mod* $2 = 1$*, and* $\lim_{t\to z} p^{(t)} = \tilde{f}'$ *, when n mod* $2 = 0$*, where* $f^{(1)}_{n/2} \le \tilde{f}' \le f^{(1)}_{(n/2)+1}$.

Proof of Proposition 3. See Appendix E.

Proposition 4. *Consider a population of risk-neutral and myopic agents of size n, where each agent updates his belief at the end of each round according to a learning model that satisfies both the directional and the magnitude assumptions. Given the market constraint* $-y \le g_i^{(t)} \le y$ *, for any round* $t \ge 1$ *and agent i, there exists a finite number of rounds z in which Algorithm 1 returns a market price* $p^{(t)}$ *so that* $\lim_{t\to z} p^{(t)} = \tilde{f} = f^{(1)}_{(n+1)/2}$*, when n mod* $2 = 1$*, and* $\lim_{t\to z} p^{(t)} = \tilde{f}'$*, when n mod* $2 = 0$*, where* $f^{(1)}_{n/2} \le \tilde{f}' \le f^{(1)}_{(n/2)+1}$.

Proof of Proposition 4. See Appendix F.

In what follows, we present an example of a learning model based on anchoring theory that satisfies both the directional and the magnitude assumptions. The main idea behind anchoring theory is that agents rely too heavily on their initial beliefs when making decisions. The theory further states that agents tend to assign an equal weight to the occurrence of each outcome when they have no information regarding the occurrence of a set of exhaustive and mutually exclusive outcomes [21]. After receiving relevant information, agents then slowly update their initial beliefs in order to accommodate the new information [22].

One formal way of modeling the underlying cognitive bias in anchoring theory is by using a linear model where agents leave some support $(1-\alpha)$ on the current belief, and the rest (α) on the new information. The parameter $\alpha \in [0, 1]$ can be seen as a learning rate: the higher its value, the weaker the anchoring bias. If we consider that each agent i receives new information regarding the occurrence of the underlying outcomes whenever the market price at the end of a round t is moving away from the agent's belief, then the belief updating equation for agent i is:

$$f_i^{(t+1)} = \left(1 - \alpha\mathbb{1}\left\{\left|f_i^{(t)} - p_0^{(t)}\right| < \left|f_i^{(t)} - p^{(t)}\right|\right\}\right) f_i^{(t)} + \alpha p^{(t)}\mathbb{1}\left\{\left|f_i^{(t)} - p_0^{(t)}\right| < \left|f_i^{(t)} - p^{(t)}\right|\right\}$$

where $\mathbb{1}\left\{\left|f_i^{(t)} - p_0^{(t)}\right| < \left|f_i^{(t)} - p^{(t)}\right|\right\}$ returns 1 if the condition in the curly brackets is true, and 0 otherwise, thus satisfying the directional assumption. Since belief updating is a convex combination of the current belief and the market price at the end of a round, the prior and posterior beliefs have the same ordinal relationship to the market price, thus satisfying the magnitude assumption.

Based on previous examples and discussion, one can immediately note that when agents do not update their beliefs, the agents with beliefs closer to the median belief will find more profitable opportunities to trade than extreme-minded agents. One can then argue that the latter agents might lose interest in trading or even find the imposed participation structure unfair. We note, however, that under the above learning model, agents' beliefs become closer and closer to the median belief/interval at the end of each round since this is the point the market price is converging to. This implies that agents have more profitable opportunities to trade when they learn over time.

On a final note, it is worth mentioning that the directional assumption implicitly implies that the agent with the median belief and the agents with beliefs in the median interval never update their beliefs. This happens because the end-of-round market prices always move towards those beliefs. If one assumes that the aforementioned agents are never willing to update their beliefs, then the magnitude assumption is the only learning assumption required to prove Propositions 3 and 4.

7. Conclusion

We formally showed how prediction markets based on the logarithmic market scoring rule can have representative final prices by imposing a market structure where the market proceeds in rounds and, in each round, agents can only hold or sell up to a fixed number of Arrow-Debreu contracts. Focusing on markets over binary outcomes, we proved that the market price converges after a finite number of rounds to the median of agents' beliefs for an odd number of agents, and to a point in the median interval for an even number of agents. We argue that the equilibrium prices effectively represent all agents' beliefs because they are measures of central tendency. When it comes to the accuracy of aggregate beliefs in predicting unobserved events, it is often reported that median beliefs are more accurate than, for example, mean values [23].

Allowing agents to trade up to a fixed number of contracts per round is in spirit equivalent to providing them with similar initial budgets in the beginning of each round. Our work thus highlighted the importance of agents having similar initial budgets. We discussed how different initial budgets might lead to manipulation of the market price by wealthy agents with ulterior motives or to a market price that is unrepresentative of agents' beliefs.

We also proposed a binary-search algorithm that provides an answer to an important open question in prediction markets, namely for how long one should run a prediction market in order to obtain representative market prices. The proposed algorithm guarantees that the market price linearly converges to a point inside an interval containing an equilibrium price after a number of rounds. As the number of rounds increases, the size of that interval decreases. Thus, the final market price is guaranteed to be within a certain distance from an equilibrium price given any number of rounds. To the best of our knowledge, this is the first work that provides strong practical guidelines on how to set a prediction market and for how long one should run a prediction market in order to obtain meaningful, representative market prices.

We discussed how trading rounds provide a natural bound on the market marker's loss. In particular, we showed that our proposed bound is expected to be tighter than the original LMSR's loss bound under the assumption that the market's liquidity parameter is set according to a scheme proposed by Berg and Proebsting [17]. Finally, we showed that if agents update their beliefs when the market price at the end of each round is moving away from their current beliefs, then no learning method affects our price convergence results as long as agents' current and posterior beliefs have the same ordinal relationship to the market price.

We believe the above results are of great value to both theoreticians and practitioners. Theoreticians, on the one hand, can better understand the impact of budgets and market structure on price convergence. Practitioners, on the other hand, have now strong guidelines on how to set up a prediction market so as to obtain representative prices. An interesting open question is whether or not similar price-convergence results hold true when the underlying context has more than two possible outcomes. In those settings, the median belief is not always well-defined in a sense that it is not necessarily a probability vector. Hence, the market prices can not always converge to the median belief. The generalization of our results to such a setting is left as future work. Another interesting question regards how to ensure market price convergence for different measures of central tendency, such as trimmed mean beliefs. Having results similar to the ones in this paper for different participation structures and measures of central tendency would certainly be of great value to decision makers, who in turn would be able to choose the most appropriate participation structure based upon the desired measure of central tendency.

Acknowledgments

The author acknowledges two anonymous reviewers and the editor for the constructive comments.

Appendix A. Proof of Proposition 1

Recall that $y - g_i^{(t)}$ is the maximum number of contracts that agent i can buy given that he already holds $g_i^{(t)}$ contracts in round t. We need to find the number of contracts x that maximizes agent i's expected gain in Eq. (6), *i.e.*, $x = \arg\max_{x'} (x'f_i - C(b,p,x'))$, without violating the imposed market constraint, *i.e.*, for $0 \leq x \leq y - g_i^{(t)}$. Next, we analyze two exhaustive cases to determine the optimal value of x: 1) when the maximum number of contracts that agent i can buy is less than or equal to the number of contracts required to bring the market price to agent i's belief; and 2) when the maximum number of contracts that agent i can buy is greater than the number of contracts required to bring the market price to agent i's belief.

Case 1. $\mathbf{0 \leq S(b,p,f_i) \leq y - g_i^{(t)}}$. We first analyze the optimal value for x inside the subinterval $\left[S(b,p,f_i), y - g_i^{(t)}\right]$. Any number of contracts $x' \geq S(b,p,f_i)$ can be written as $x' = S(b,p,f_i) + x''$, where $x'' \geq 0$. Using the path-independence property in Eq. (5), the resulting market price after a purchase of x' contracts is equivalent to moving the market price to f_i by buying $S(b,p,f_i)$ contracts followed by a purchase of $x'' \geq 0$ contracts. As discussed in Section 3.2, it is never profitable to trade when $p = f_i$. Hence, the most profitable choice is to buy $S(b,p,f_i)$ contracts first, and then buy $x'' = 0$ contracts later. We now analyze the optimal value for x inside the subinterval $[0, S(b,p,f_i)]$. Any value x', for $0 \leq x' \leq S(b,p,f_i)$, can be written as $x' = S(b,p,f_i) - x''$, where $x'' \geq 0$. Using the path-independence property in Eq. (5), the resulting market price after a purchase of x' contacts is equivalent to moving the market price to f_i by buying $S(b,p,f_i)$ contracts followed by a sale of x'' contracts. Once again, it is never profitable to trade when $p = f_i$. Hence, the most profitable choice is to buy $S(b,p,f_i)$ contracts first, and then sell $x'' = 0$ later.

Case 2. $\mathbf{0 \leq y - g_i^{(t)} < S(b,p,f_i)}$. Due to the imposed market constraint, agent i cannot buy $S(b,p,f_i)$ contracts. Hence, we restrict ourselves to analyze only the subinterval $\left[0, y - g_i^{(t)}\right]$. One can then use an argument similar to the previous case to show that the most profitable choice is to buy $y - g_i^{(t)}$ contracts first, and then sell $x'' = 0$ later.

Combining the above cases, the most profitable choice is to buy $x = \min\left(S(b,p,f_i), y - g_i^{(t)}\right)$. □

Appendix B. Proof sketch of Proposition2

Recall that $-y - g_i^{(t)}$ is the maximum number of contracts that agent i can sell given that he already holds $g_i^{(t)}$ contracts in round t. We need to find the number of contracts x that maximizes agent i's expected gain in Eq. (8), *i.e.*, $x = \arg\max_{x'} (-C(b,p,-x') - x'f_i)$, without violating the imposed market constraint, *i.e.*, for $-y - g_i^{(t)} \leq x \leq 0$. Recall that negative values of x encode sales transactions. The proof of this proposition then becomes similar to the proof of Proposition 1, where one simply has to analyze two exhaustive cases to determine the optimal value for x, namely: $-y - g_i^{(t)} \leq S(b,p,f_i) \leq 0$ and $S(b,p,f_i) < -y - g_i^{(t)} \leq 0$. □

Appendix C. Proof of Theorem1

We analyze the sequence of market prices at the end of successive rounds, *i.e.*, $p^{(1)}, p^{(2)}, \ldots, p^{(z)}$. Within each round, we analyze the price movement using the path-independence property in Eq. (5).

Case 1. $\mathbf{n \bmod 2 = 1}$. We first note that when $p^{(t)} = \tilde{f}$ the final market price will no longer change in successive rounds. The reason for this is that for each agent i with belief $f_i < \tilde{f}$ willing to sell x contracts according to Proposition 2, there is at least one agent j with belief $f_j > \tilde{f}$ willing to buy x contracts according to Proposition 1, and *vice versa*. Due to the path independence property in Eq. (5), the resulting price after these trades is $Q(b,p,x) = Q(b, Q(b,p,-x), x) = p$. That is, the trades of the $(n-1)/2$ agents who have beliefs less than or equal to $\tilde{f}$ offset the trades of the $(n-1)/2$ agents who have beliefs greater than or equal to $\tilde{f}$, and the only agent able to change the market price is the agent with belief $\tilde{f}$, which in turn has no incentives to move the market price away from his belief because this is not profitable. Hence, we can restrict ourselves to show that the market price does indeed reach $\tilde{f}$ after a finite number of rounds. Without loss of generality due to symmetric considerations, assume that the initial market price at certain round $t \geq 1$ is less than the median belief, *i.e.*, $p_0^{(t)} < \tilde{f}$. Similar to the above argument, the trades of each agent i with belief $f_i \leq p_0^{(t)}$ are offset by the trades of at least one agent j with belief $f_j \geq p_0^{(t)}$, but the opposite is no longer true. Hence, using the path independence property in Eq. (5) and Proposition 1, the resulting market price after all transactions during round t is equivalent to the result of a purchase of, say, $x > 0$ contracts by the agents who do not have their trades offset by other agents' trades. Moreover, $x \leq S(b, p_0^{(t)}, \tilde{f})$, *i.e.*, the market price at the end of round t, $p^{(t)}$, is always less than or equal to the median belief $\tilde{f}$, otherwise at least one agent i with belief $f_i \leq \tilde{f}$, who previously purchased $g_i^{(t)} \leq y$ contracts, would be better off now by selling some of these contracts in order to move the market price closer to his belief. Consequently, for each round t and for every agent i, by setting $g_i^{(t)} = 0$ in the beginning of the round and forcing the constraint $|g_i^{(t)}| \leq y$ throughout each round, the market price continually increases at the end of each round until it equals the median belief, *i.e.*, $p^{(1)} < p^{(2)} < \cdots < p^{(z)} = \tilde{f}$. Given that the liquidity parameter $b \neq \infty$, the number of purchased contracts needed to move the market price from any $p_0^{(1)} \in (0,1)$ to any $\tilde{f} \in (0,1)$ is always finite, as can be seen in Eq. (4).

Case 2. n mod 2 = 0. We note note that when $f_{n/2} \leq p^{(t)} \leq f_{(n/2)+1}$, the final market price will no longer change in successive rounds since for each agent i with belief $f_i \leq f_{n/2}$ willing to sell x contracts according to Proposition 2, there is at least one agent j with belief $f_j \geq f_{(n/2)+1}$ willing to buy x contracts according to Proposition 1, and *vice versa*. Due to the path independence property in Eq. (5), the resulting price after these trades is $Q(b,p,x) = Q(b,Q(b,p,-x),x) = p$. Hence, we can restrict ourselves to show that the market price does indeed reach the median interval after a finite number of rounds. Using an argument similar to the above case, one conclude that if the initial market price at certain round $t \geq 1$ is less than $f_{n/2}$ (respectively, greater than $f_{(n/2)+1}$), then the market price continually increases at the end of each round until it equals $f_{n/2}$ (respectively, $f_{(n/2)+1}$), and the number of traded contracts required for the convergence is finite. □

Appendix D. Proof of Theorem2

Similar to the proof of Theorem 1, we analyze the sequence of market prices at the end of successive rounds: $p^{(1)}, p^{(2)}, \ldots, p^{(z)}$. Within each round, we analyze the price movement using the path-independence property.

Case 1. n mod 2 = 1. Recall from Corollary 1 that the final market price at the end of a round t is no farther from the median belief than the market price in the beginning of t, *i.e.*, $|\tilde{f} - p^{(t)}| \leq |\tilde{f} - p_0^{(t)}|$. Hence, due to the path independence property in Eq. (5), the resulting market price after all the trades in any given round t is equivalent to the resulting market price after the agent with median belief trades a certain number of contracts in order to move the market price towards his belief. We can then shift our analysis to a single-agent trade. Consider the values $0 \leq lb' < \tilde{f} < ub' \leq 1$ such that $x' = S(b, lb', \tilde{f})$ and $x'' = S(b, ub', \tilde{f})$, *i.e.*, lb' (respectively, ub') is the lowest (respectively, greatest) possible market price such that the agent with median belief is able to drive the market price to his belief in one round by buying x' (respectively, selling x'') contracts. That is, if the market price at the end of round t is either lb' or ub', then the market price at the end of round $t+1$ will be equal to the median belief. We can now restrict ourselves to show that at least one of the following inequalities is true after a finite number of rounds: $lb \geq lb'$ or $ub \leq ub'$, where lb and ub are the bounds of the interval where the median belief lies according to Algorithm 1. We can rephrase the above to say that the inequality $ub-lb < ub'-lb'$ has to be true after a finite number of rounds. Using the function in Eq. (10), we then have that we need at most $\log_{0.5}(ub' - lb')$ rounds for the above inequality to be true. By construction, $0 < ub' - lb' < 1$, which means that the value $\log_{0.5}(ub' - lb')$ is finite.

Case 2. n mod 2 = 0. We first note that when $f_{n/2} \leq p^{(t)} \leq f_{(n/2)+1}$, for any round $t \geq 1$, the final market price will no longer change in successive rounds (see Theorem 1). If $f_{n/2} = f_{(n/2)+1}$, we can then treat the agents with beliefs $f_{n/2}$ and $f_{(n/2)+1}$ as a single agent with median belief $\tilde{f}$, and thus Case 1 applies. Otherwise, we can restrict ourselves to show that the inequality $ub - lb \leq f_{(n/2)+1} - f_{n/2}$ is true after a finite number of rounds, *i.e.*, the range is less than or equal to the median interval, which means that an equilibrium point is reached. Using the function in Eq. (10), that inequality holds true after at most $\log_{0.5}(f_{(n/2)+1} - f_{n/2})$ rounds. Since $0 < f_{(n/2)+1} - f_{n/2} \leq 1$, then $\log_{0.5}(f_{(n/2)+1} - f_{n/2}) < \infty$. □

Appendix E. Proof of Proposition3

Similar to the proofs of Theorem 1 and Theorem 2 in Appendices C and D, we next analyze the sequence of market prices at the end of successive rounds, *i.e.*, $p^{(1)}, p^{(2)}, \ldots, p^{(z)}$. Within each round, we analyze the price movement using the path-independence property in Eq. (5). We consider the cases where the population contains an odd number and an even number of agents.

Case 1. n mod 2 = 1. Without loss of generality due to symmetric considerations, assume that the initial market price at certain round $t \geq 1$ is less than the median belief, *i.e.*, $p_0^{(t)} < \tilde{f}$. Due to Corollary 1, the market price at the end of round t is no farther from the median belief than the market price in the beginning of round t. Hence, Corollary 1 together with the directional assumption imply that the agent with median belief never learns from the market price because the market price is always moving towards his belief. Moreover, due to the magnitude assumption, the agents who update their beliefs, *i.e.*, those who have beliefs less than or equal to the market price at the end of round t, do not learn beyond the median belief. This happens because the market price at the end of round t is less than or equal to the median belief. Then, similar to the proof of Case 1 of Theorem 1, the market price at the end of each round monotonically increases until it equals the median belief. Given that the liquidity parameter $b \neq \infty$, the number of purchased contracts needed to move the market price from any $p_0^{(t)} \in (0,1)$ to any $\tilde{f} \in (0,1)$ is always finite (see Eq. (4)).

Case 2. n mod 2 = 0. This case is similar to the above case when one observes that the directional assumption implies that the agents with initial beliefs $f_{n/2}^{(1)}$ and $f_{(n/2)+1}^{(1)}$ never learn from the market price because the market price is always moving towards their beliefs. Moreover, due to the magnitude assumption, the agents who update their beliefs do not learn beyond $f_{n/2}^{(1)}$ and $f_{(n/2)+1}^{(1)}$. □

Appendix F. Proof sketch of Proposition4

The proof of this proposition is similar to the proofs of Proposition 3 and Theorem 2 after one observes the following points: (1) due to Corollary 1, the market price at the end of any round is no farther from the median of agents' beliefs than the market price in the beginning of that round. Consequently, due to the directional assumption, the agent with median belief, for an odd number of agents, and the agents with beliefs in the median interval, for an even number of agents, never update their beliefs; and (2) due to the magnitude assumption, the agents who update their beliefs never change the median belief. □

References

[1] I. Blohm, C. Riedl, J. Füller, J.M. Leimeister, Rate or trade? Identifying winning ideas in open idea sourcing, Inf. Syst. Res. 27 (1) (2016) 27–48.

[2] J. Berg, R. Forsythe, F. Nelson, T. Rietz, Results from a dozen years of election futures markets research, in: C.A. Plott, V.L. Smith (Eds.), Handbook of Experimental Economic Results, vol. 1, 2008. pp. 742–751.

[3] K.-Y. Chen, C.R. Plott, Information aggregation mechanisms: concept, design and implementation for a sales forecasting problem, Technical report, California Institute of Technology, Division of the Humanities and Social Sciences. 2002,

[4] D.M. Pennock, S. Lawrence, C.L. Giles, F.A. Nielsen, The real power of artificial markets, Science 291 (5506) (2001) 987–988.

[5] P. Milgrom, N. Stokey, Information, trade and common knowledge, J. Econ. Theory 26 (1) (1982) 17–27.

[6] R. Hanson, Combinatorial information market design, Inf. Syst. Front. 5 (1) (2003) 107–119.

[7] Y. Chen, D. Pennock, Designing markets for prediction, AI Magazine 31 (4) (2010) 42–52.

[8] R. Hanson, Logarithmic market scoring rules for modular combinatorial information aggregation, J. Prediction Markets 1 (1) (2007) 1–15.

[9] C.F. Manski, Interpreting the predictions of prediction markets, Econ. Lett. 91 (3) (2006) 425–429.

[10] A.J. Storkey, Machine learning markets, Proceedings of the 14th International Conference on Artificial Intelligence and Statistics, 2011. pp. 716–724.
[11] C. Deck, S. Lin, D. Porter, Affecting policy by manipulating prediction markets: experimental evidence, J. Econ. Behav. Organ. 85 (2013) 48–62.
[12] M. Ostrovsky, Information aggregation in dynamic markets with strategic traders, Econometrica 80 (6) (2012) 2595–2647.
[13] R. Sethi, J.W. Vaughan, Belief aggregation with automated market makers, Comput. Econ. 48 (1) (2016) 155–178.
[14] L. Jian, R. Sami, Aggregation and manipulation in prediction markets: effects of trading mechanism and information distribution, Manag. Sci. 58 (1) (2012) 123–140.
[15] A. Graefe, J.S. Armstrong, Comparing face-to-face meetings, nominal groups, delphi and prediction markets on an estimation task, Int. J. Forecast. 27 (1) (2011) 183–195.
[16] L.J. Savage, Elicitation of personal probabilities and expectations, J. Am. Stat. Assoc. 66 (336) (1971) 783–801.
[17] H. Berg, T.A. Proebsting, Hanson's automated market maker, J. Prediction Markets 3 (1) (2009) 45–59.
[18] Y. Chen, S. Dimitrov, R. Sami, D.M. Reeves, D.M. Pennock, R.D. Hanson, L. Fortnow, R. Gonen, Gaming prediction markets: equilibrium strategies with a market maker, Algorithmica 58 (4) (2010) 930–969.
[19] P.P. Wakker, Prospect Theory: For Risk and Ambiguity, Cambridge University Press. 2010.
[20] E. Servan-Schreiber, J. Wolfers, D.M. Pennock, B. Galebach, Prediction markets: does money matter? Electron. Markets 14 (3) (2004) 243–251.
[21] C. Fox, Y. Rottenstreich, Partition priming in judgment under uncertainty, Psychol. Sci. 14 (3) (2003) 195–200.
[22] K.E. See, C.R. Fox, Y.S. Rottenstreich, Between ignorance and truth: partition dependence and learning in judgment under uncertainty, J. Exp. Psychol. Learn. Mem. Cogn. 32 (6) (2006) 1385–1402.
[23] J.-C. Larreche, R. Moinpour, Managerial judgment in marketing: the concept of expertise, J. Market. Res. (1983) 110–121.

Dr. Arthur Carvalho is an Assistant Professor of Information Systems and Analytics at Farmer School of Business, Miami University. His research interests lie at the intersection of Information Systems, Analytics, and Decision Analysis. In particular, Dr. Carvalho works on the development of mathematical techniques and decision support systems that assist decision makers in making better decisions, with a particular focus on applications in the crowdsourcing and energy domains. Dr. Carvalho's research has been published in academic journals such as *Annals of Mathematics and Artificial Intelligence, APA Decision, INFORMS Decision Analysis, Judgement and Decision Making, The Journal of Prediction Markets*, among others. Before joining Miami University, Dr. Carvalho worked as an Assistant Professor at Rotterdam School of Management, Erasmus University. He received his MMath and PhD in Computer Science from the University of Waterloo in 2010 and 2014.