Belief Aggregation and Trader Compensation in Infinite Outcome Prediction Markets

Blake Martin, Mithun Chakraborty, Sindhu Kutty
Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI

ELECTRICAL ENGINEERING & COMPUTER SCIENCE
UNIVERSITY OF MICHIGAN

Introduction

- Prediction markets are incentive-compatible mechanisms designed to **elicit the personal beliefs** of traders about a future uncertain event and **aggregate those beliefs** into the market price
- Prediction markets have been empirically observed to outperform polls as they have built-in **financial incentives** and **timely responses**
- We study the impact of traders' **informativeness**, **budget**, and the **sequence** in which they trade on **aggregation** properties and trader **compensation** under a new prediction market design

Background

- The **Bernoulli probability distribution** models an uncertain binary event with success parameter p

$$f(x;p) = p^x(1-p)^{1-x}, \quad \text{for } x \in 0,1$$

- The **Beta** distribution is **conjugate prior** for the Bernoulli distribution success parameter and has pdf given by

$$f(p;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

- Given a Beta prior over p and a sample of Bernoulli observations, the **Bayesian posterior update** is given by

$$\hat{\alpha} := \alpha + \sum_{i=1}^{N} x_i$$
$$\hat{\beta} := \beta + \sum_{i=1}^{N} 1 - x_i$$

- The **gamma** and **digamma** functions are used in computing statistics associated with the Beta distribution and are given below

$$\Gamma(n) = (n-1)! \qquad \psi(n) = \frac{\mathrm{d}\Gamma(n)/\mathrm{d}n}{\Gamma(n)} \qquad \text{for } n \in Z^+$$

Trader Model

- We assume traders are myopic, risk-neutral, and rational
- Trader **informativeness** is modeled by N, the size of their private sample of Bernoulli observations at each trade

$$x_1, \ldots, x_N \overset{iid}{\sim} \text{Bern}(p)$$

- Traders may also have limiting **budgets** B, modeled by an additional constraint of using the largest observation sub-sequence of length k where worst-case losses do not exceed B

$$\max_p \left\{ C\left(\eta + \left[\sum_{i=1}^{k} x_i, k - \sum_{i=1}^{k} x_i\right]^T\right) - C(\eta) - \left[\sum_{i=1}^{k} x_i, k - \sum_{i=1}^{k} x_i\right] \cdot [\ln p, \ln(1-p)] \right\} \leq B$$

- Trades proceed as follows:
  - Trader **samples from Bernoulli** distribution according to their informativeness
  - Trader **updates Beta posterior beliefs** according to $\hat{\alpha} := \alpha + \sum_{i=1}^{N} x_i$, $\hat{\beta} := \beta + \sum_{i=1}^{N} 1 - x_i$
  - Trader purchases shares such that **market posterior matches private beliefs**

Aggregation Mechanism: Beta-Bernoulli Market

- **Securities and payoffs:** The market has two securities with payoffs given by $\ln p_{true}$ and $\ln(1-p_{true})$ when $p_{true} \in [0,1]$, the true Beta forecast variable corresponding to the Bernoulli success parameter, is revealed
- **Outstanding shares and aggregated belief:** The market maintains and displays outstanding shares $\eta = [\eta_1, \eta_2]$ which also corresponds to market belief given by $\mathrm{BETA}(p;\alpha,\beta)$ where $\eta_1 = \alpha - 1, \eta_2 = \beta - 1$. This gives us expectations of the Bernoulli parameter and payoffs

$$\mathbb{E}_\eta[p] = \frac{\eta_1+1}{\eta_1+\eta_2+2} \qquad \mathbb{E}_\eta[\ln p] = \psi(\alpha) - \psi(\alpha+\beta)$$
$$\mathbb{E}_\eta[\ln(1-p)] = \psi(\beta) - \psi(\alpha+\beta)$$

- **Cost Function:** This gives us the cost of a trade $C(\eta_{new}) - C(\eta_{old})$, the prices of each security $\nabla_\eta C(\eta)$, and is given by

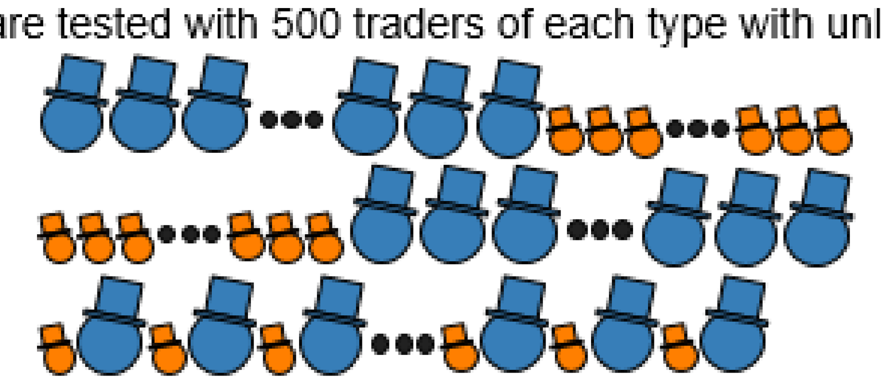
$$C(\eta) = \ln\Gamma(\eta_1+1) + \ln\Gamma(\eta_2+1) - \ln\Gamma(\eta_1+\eta_2+2)$$

Informativeness and Sequence Results

- **Experimental Setup** High Informativeness $\leftrightarrow N = 5$ Low Informativeness $\leftrightarrow N = 1$

Three sequences are tested with 500 traders of each type with unlimited budgets

1) High Info (HI) First:
2) Low Info (LI) First:
3) Interleaved:



- We report market convergence (Fig. 1), overall belief aggregation (Fig. 2) and average compensation for each trade type (Table 1)

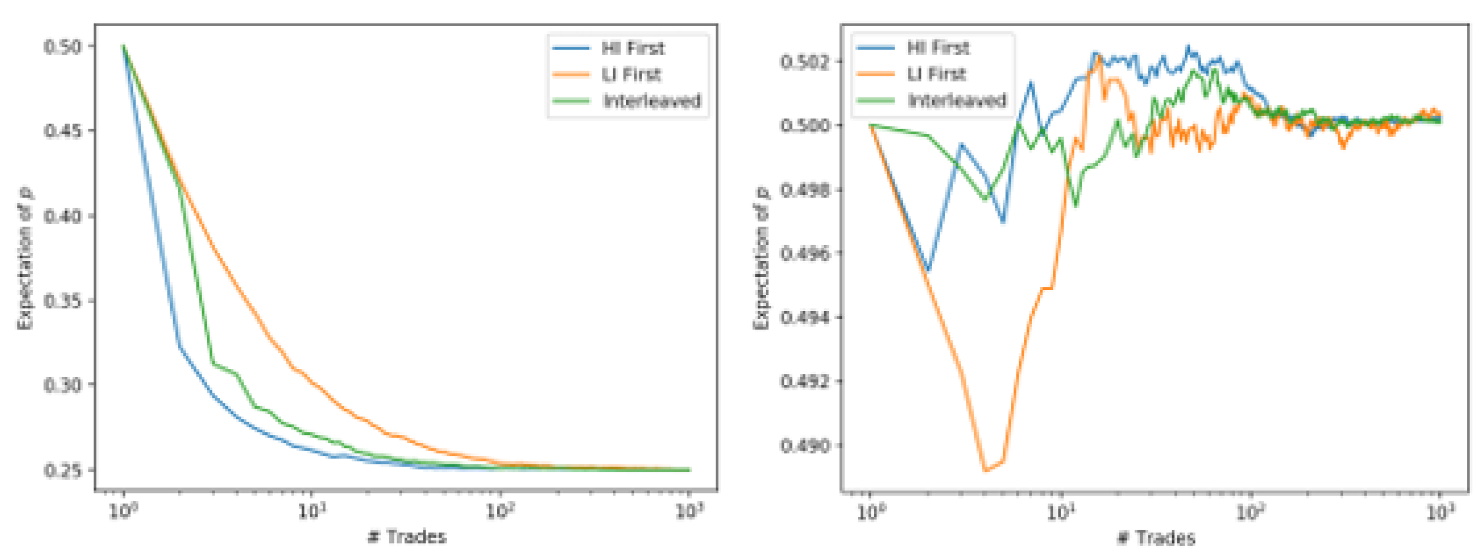


Figure 1: The market's expectation of the Bernoulli parameter as a function of the number of completed trades for different trader orderings, averaged over 1000 simulations with $p_{true} = 0.25$ (left) and $p_{true} = 0.5$ (right). Horizontal axis is logarithmic.

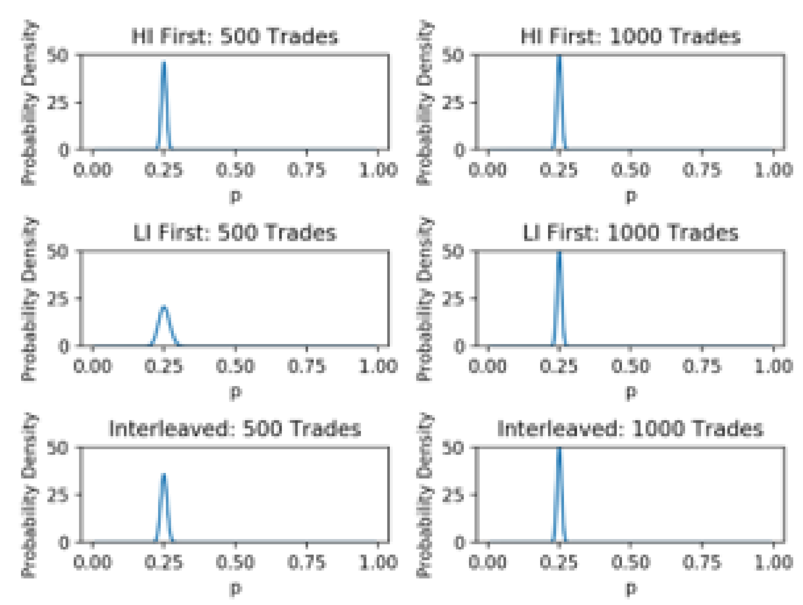


Figure 2: Posterior distribution (market belief) after 500 and 1000 trades for different trader orderings averaged across 1000 simulations with $p_{true} = 0.25$.

p_{true}	Sequence	Avg HI ($\times 10^3$)	Avg LI ($\times 10^3$)
0.75	**HI First**	6.66 ± 0.09	0.18 ± 0.05
	LI First	**1.75 ± 0.12**	**5.07 ± 0.09**
	Interleaved	5.51 ± 0.15	1.29 ± 0.13
0.50	**HI First**	6.37 ± 0.08	0.17 ± 0.05
	LI First	**1.75 ± 0.11**	**4.78 ± 0.09**
	Interleaved	5.56 ± 0.14	0.99 ± 0.12
0.25	**HI First**	6.70 ± 0.08	0.21 ± 0.05
	LI First	**1.80 ± 0.11**	**5.04 ± 0.09**
	Interleaved	5.52 ± 0.14	1.26 ± 0.12

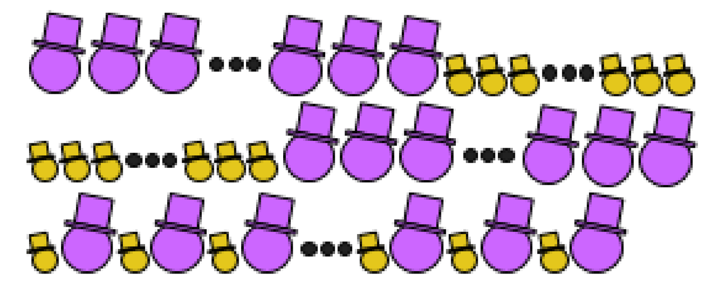
Table 1: Compensation of HI and LI traders for different trader orderings and p_{true} values. This table reports the compensation averaged over 1000 simulations with 95% confidence bounds.

Budget and Sequence Results

- **Experimental Setup** Unlimited Budget $\leftrightarrow B_0 = \infty$ Limited Budget $\leftrightarrow B_0 \in (0,\infty)$

Three sequences are tested with 500 traders of each type with constant informativeness ($N = 5$ per trade) and initial budgets B_0. There are **25 successive market instances** where agents trade, receive compensation, and update their budgets before proceeding to the next round.

1) Unlimited Budget (UB) First:
2) Limited Budget (LB) First:
3) Interleaved:



- We report the change in average LB trader budget over market rounds

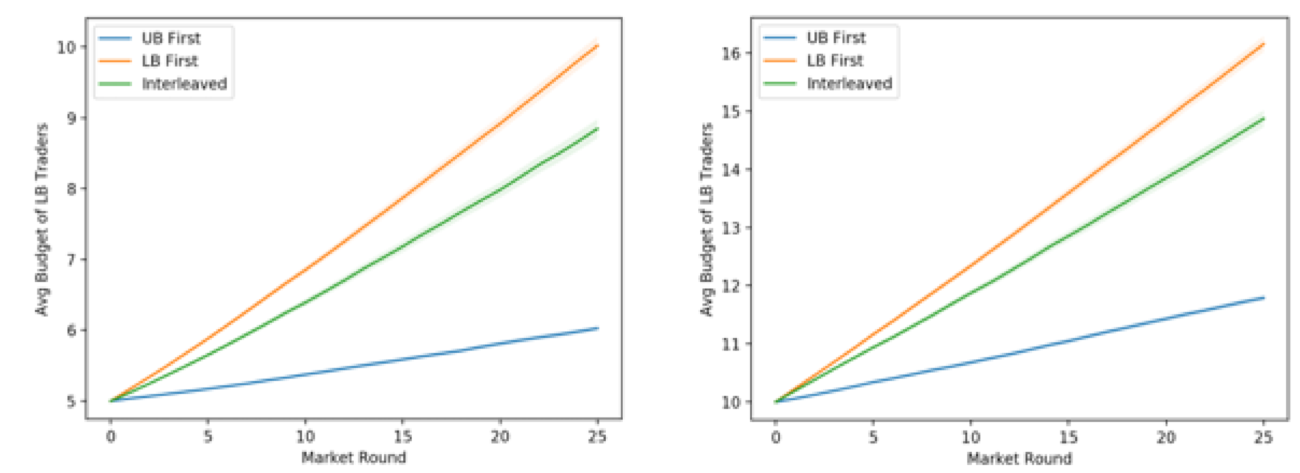


Figure 3: Change in budget of LB traders with an initial budget of $B_0 = 5$ (left) and $B_0 = 10$ (right) over rounds 1 – 25, averaged over 1000 simulations with $p_{true} \sim U[0.05, 0.95]$ for different trader orderings.

Conclusions

1. The expected **compensation** of a trader depends not only on her **informativeness** but also strongly on the **sequence** (Table 1)
2. Notably, **LI traders receive higher compensation** on average than HI traders when they **trade first**. This phenomenon underscores the advantage gained by injecting information into the market early as **higher marginal value of information** leads to **higher compensation**
3. Traders with the same overall informativeness induce the same **aggregation characteristics** at convergence **regardless of sequence**, but number of trades required to reach **steady-state** is largely impacted by **trader ordering** (Figs. 1-2)
4. All LB agents can trade on **more information** in round 25 than in the initial rounds due to **increased budget regardless of sequence**, though rate of change in budgets is affected by trader ordering (Fig. 3)
5. Future work includes further experiments with other **trader models** as well as theoretical analysis to quantify the impact of the **sequencing-informativeness-budget interplay**

Significant References

- Jacob Abernethy, Sindhu Kutty, Sébastien Lahaie, and Rahul Sami. "Information Aggregation in Exponential Family Markets." EC 2014
- Martin Wainwright and Michael Jordan. "Graphical Models, Exponential Families, and Variational Inference." Foundations and Trends in Machine Learning 2008
- Robin Hanson. "Logarithmic Market Scoring Rules for Modular Combinatorial Information Aggregation." Journal of Prediction Markets 2007

COLLEGE OF ENGINEERING
ELECTRICAL ENGINEERING & COMPUTER SCIENCE
UNIVERSITY OF MICHIGAN