Information Aggregation in Prediction Markets

Rahul Sami
University of Michigan School of Information
Current affiliation: Google, Bangalore

NMI Workshop on Game Theory and Mechanism Design

Talk Goals

- Introduction to scoring rule and market scoring rule mechanisms
- Theoretical analysis of strategies
  - with connections to information theory
- Small peek into experimental methods

Information/Prediction Markets

Pays $1 if India beats AUS
Pays $0 otherwise

Markets designed to aggregate traders' information.

- Issue securities with value contingent on future event.
- Trading price is taken as a prediction of future value.
- Once event occurs, security is cashed out for money

Example:Iowa Electronic Markets

iemweb.biz.uiowa.edu/quotes/Pres12_quotes.html

isited ▾ Getting Started Latest Headlines ▾ Ann Arbor – Mic...

Markets ▫ 2012 Presidential ▫ 2012 Presidential Data ▫ 2012 Presidential Quotes

Market Quotes: Pres12_VS
2012 Presidential Election Vote Share Market.

Quotes current as of **14:15:04 CST, Monday, October 31, 2011**.

Symbol	Bid	Ask	Last	Low	High	Average
UDEM12_VS	0.478	0.497	0.499	---	---	---
UREP12_VS	0.511	0.518	0.510	---	---	---

| Prospectus | Price History | Graph |

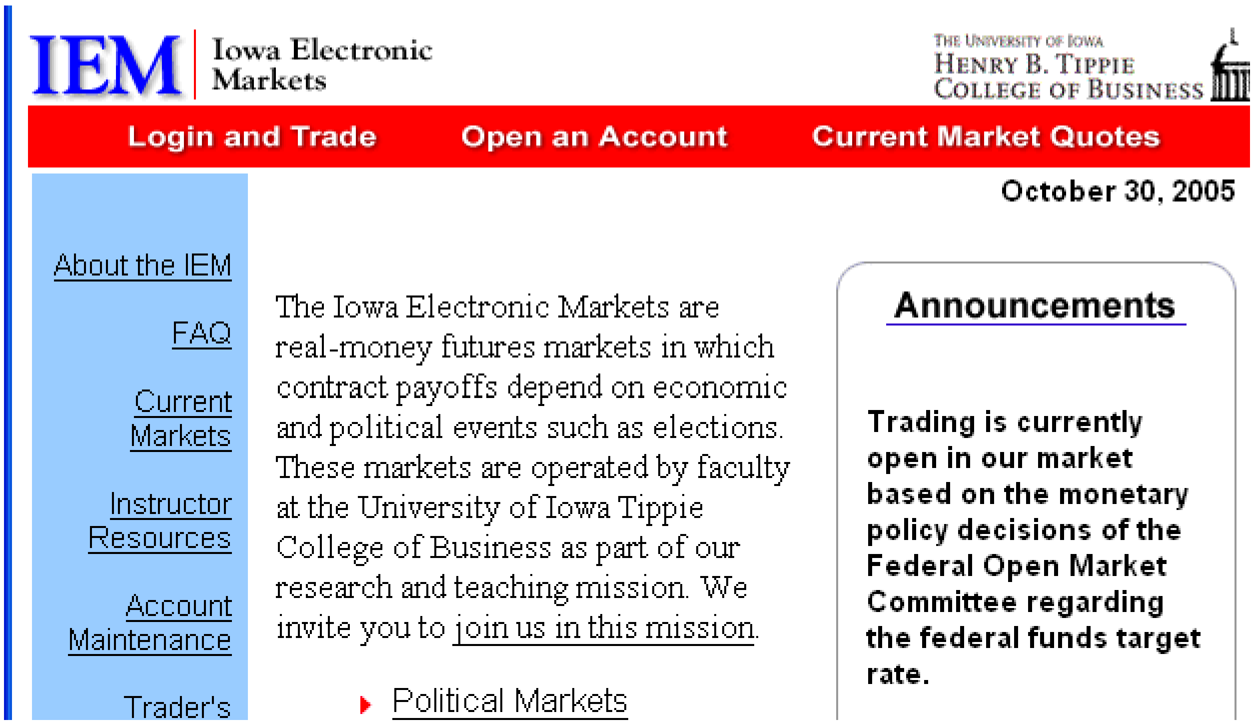
Market Quotes: Pres12_WTA
2012 Presidential Election Winner-Take-All Market.

Quotes current as of **14:15:04 CST, Monday, October 31, 2011**.

Symbol	Bid	Ask	Last	Low	High	Average
DEM12_WTA	0.506	0.513	0.510	0.510	0.510	0.510
REP12_WTA	0.491	0.493	0.492	0.490	0.492	0.490

| Prospectus | Price History | Graph |

Markets aggregate information

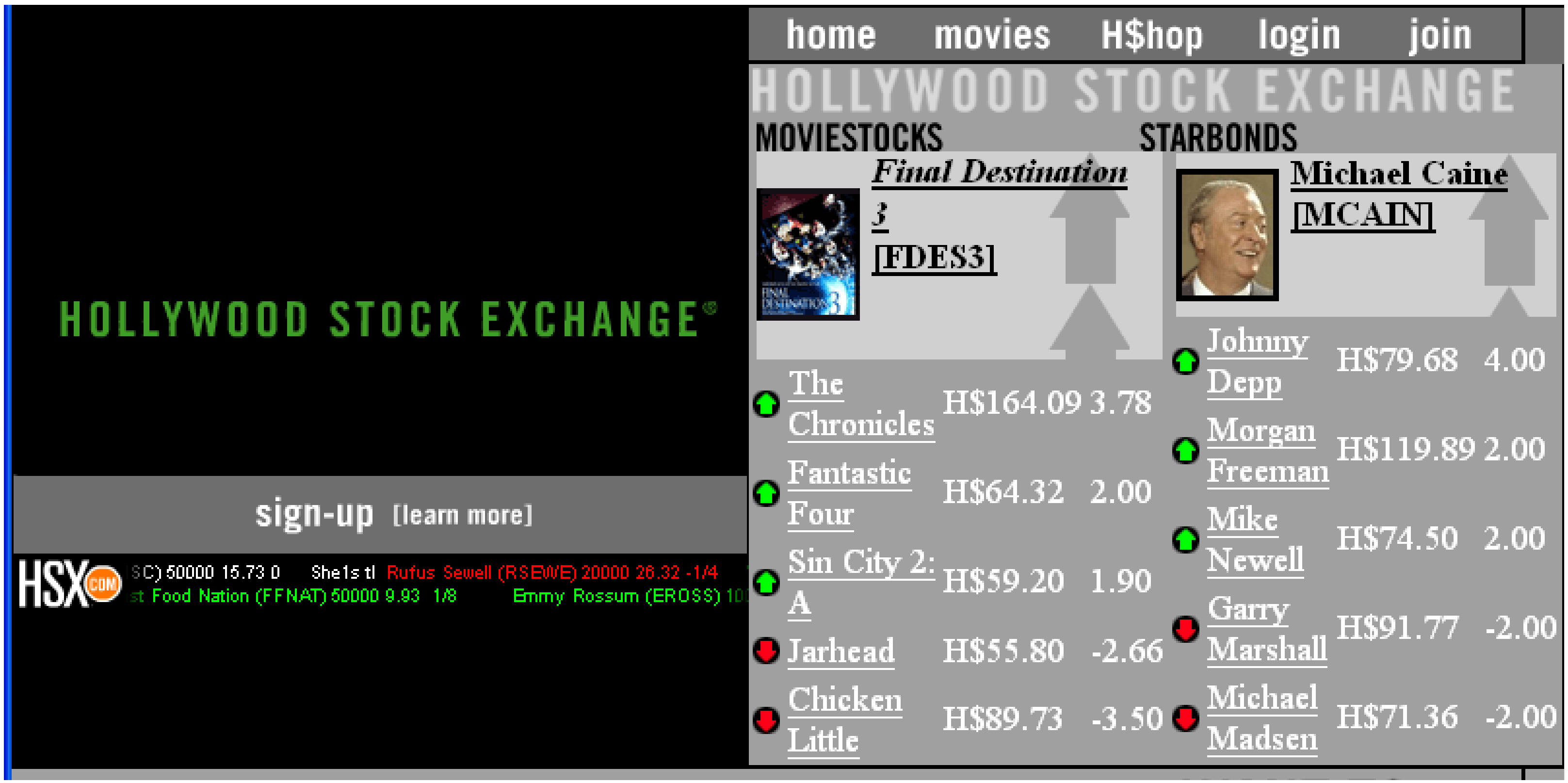


The Iowa Electronic Market predicts election outcomes better than opinion polls [Forsythe *et al.* '99].

Markets aggregate information

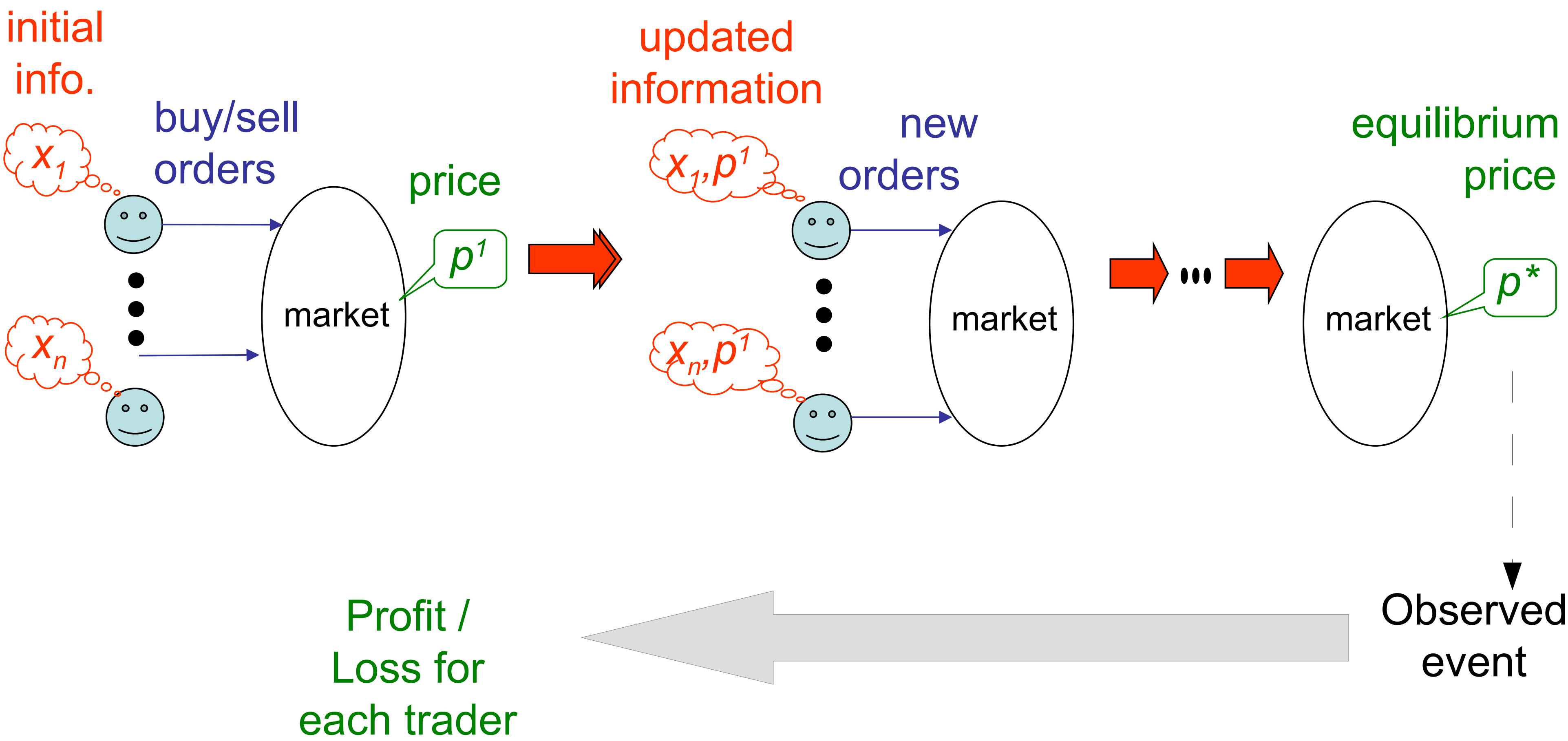
Sports betting markets provide unbiased forecasts of game outcomes [Gandar *et al.* '98; Debnath *et al.* '03]

Markets Designed for Aggregation



Markets sometimes deployed primarily for information aggregation (*e.g.*, IEM, Hollywood Stock Exchange)

Market as Incentive Mechanism



equilibrium price p^* $\Leftrightarrow$ $f(x_1, x_2, \ldots, x_n)$

Goal: Profit incentive should induce optimal aggregation

Single Forecaster Incentives

Decision Maker

q

Expert

What is the probability of rain tomorrow?

RAIN or DRY

Reward

Scoring rules

A *scoring rule* is a rule that is used to compute the reward for a forecaster.

- Depends on the forecast probability of rain q
- Must also depend on actual outcome
- For this example, scoring rule consists of two functions:

$$S_{RAIN}(q)\ ,\quad S_{DRY}(q)$$

Example: Linear scoring rule

- If you say “It will rain with probability p” and it rains =>your reward is \$p
- If you say “it will rain with probability p” and it is dry=> your reward is \$(1-p)

If you think the probability is 80% of rain (and say so), what is your expected reward?

Proper scoring rule

Proper scoring rules satisfy the following property:

If a forecaster believes the probability of an event is p, her expected reward is maximized by reporting q=p.

There are several well-known proper scoring rules:

- *Quadratic Scoring Rule [Brier 1952]*
- *Logarithmic Scoring Rule [Good 1950]*
- *Spherical Scoring Rule*

✗ Linear scoring rule is *not* proper

Logarithmic Scoring Rule

$$S_{RAIN}(q) = \log q$$

$$S_{DRY}(q) = \log(1-q)$$

Log scoring rule is a proper scoring rule:

Logarithmic Scoring Rule

$S_{RAIN}(q) = log\ q$

$S_{DRY}(q) = log(1\text{-}q)$

Log scoring rule is a proper scoring rule:

$$
\begin{aligned}
\text{Ep [Score(q)]} &= \text{p log q + (1-p) log (1-q)} \\
&= \text{p log p + (1-p) log p +} \\
&\quad\ \text{p log (q/p) + (1-p) log [(1-q)/(1-p)]}
\end{aligned}
$$

Logarithmic Scoring Rule

$S_{RAIN}(q) = log\ q$

$S_{DRY}(q) = log(1\text{-}q)$

Log scoring rule is a proper scoring rule:

$$
\begin{aligned}
Ep\ [\ \mathrm{Score}(q)\ ] &= p \log q + (1-p) \log (1-q) \\
&= p \log p + (1-p) \log p + \\
&\quad\ p \log (q/p) + (1-p) \log [(1-q)/(1-p)] \\
&= -H(p) - D(p \,||\, q)
\end{aligned}
$$

Entropy KL-divergence

Optimal q : q=p => Expected score = -*H(p)*

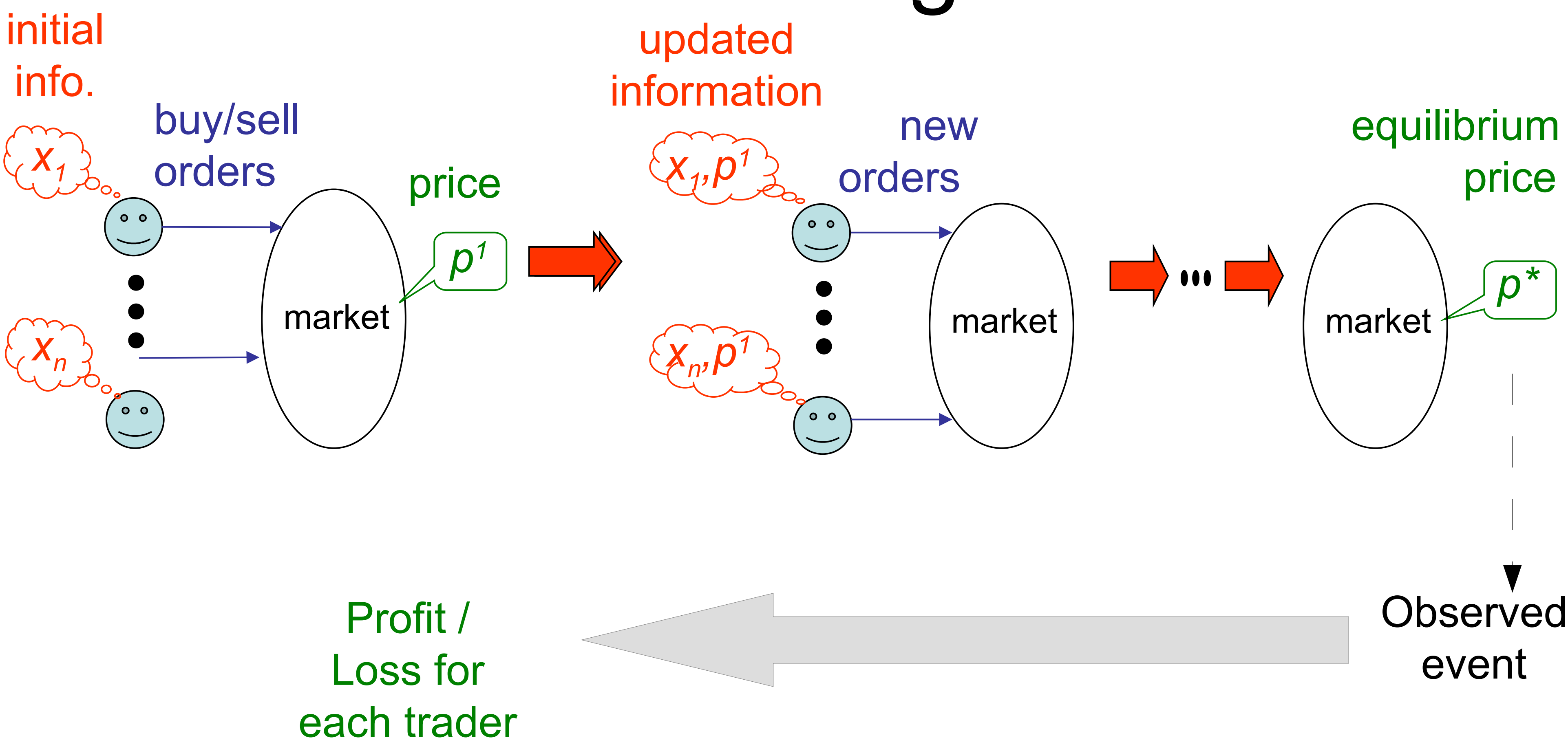
Logarithmic Scoring Rule: Variations

$$S_{RAIN}(q) = a + b \log q$$

$$S_{DRY}(q) = a + b \log(1-q)$$

Constants a,b control scale and absolute value of rewards while retaining strategic properties.

Multiple Forecasters: Market Setting

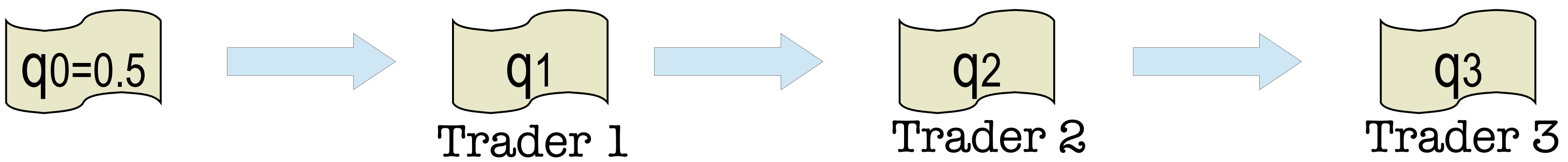


equilibrium price p^* $\Leftrightarrow$ $f(x_1, x_2, \ldots, x_n)$

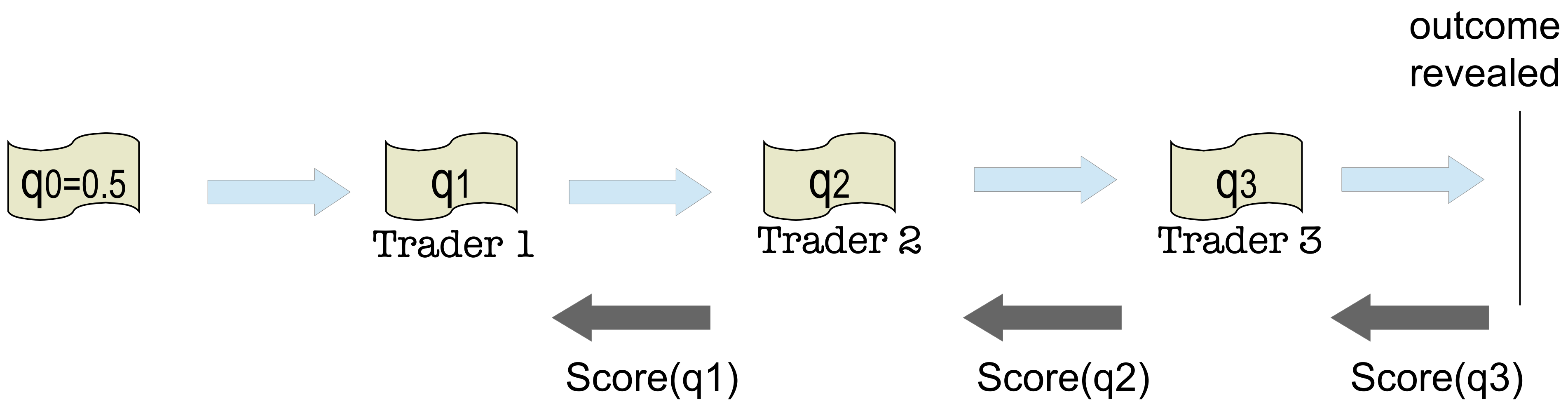
Goal: Profit incentive should induce optimal aggregation

Market Scoring Rules [Hanson 03]

Market based on trading scoring rules

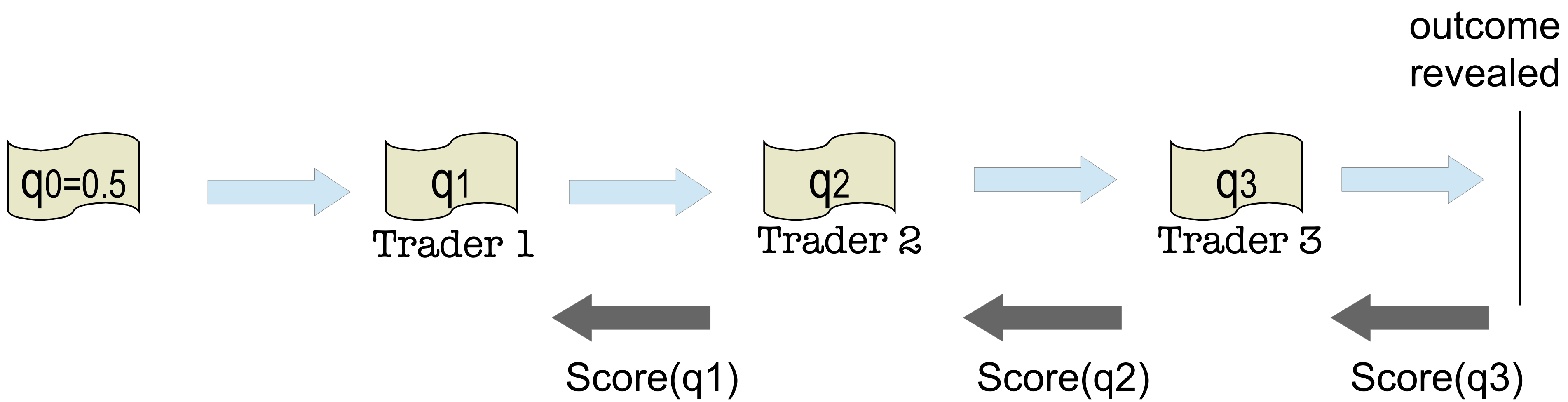


Market Scoring Rule Payoffs



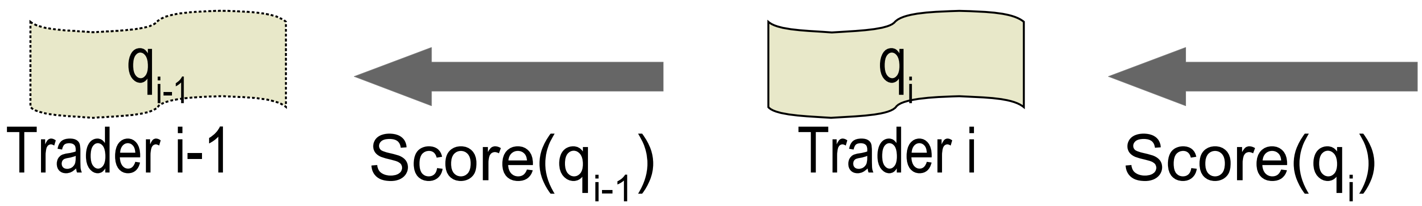
- Market maker rewards last trader
- Each trader pays previous trader’s reward

Market Scoring Rule Payoffs



- Market maker rewards last trader
- Each trader pays previous trader’s reward

Strategies with Market Scoring Rules



- Profit of trader *i:*

$$\text{Profit(i)} = \text{Score}(q_i) - \text{Score}(q_{i-1})$$
$$\text{E(Profit(i))} = \text{E(Score}(q_i) - \text{E (Score}(q_{i-1})$$

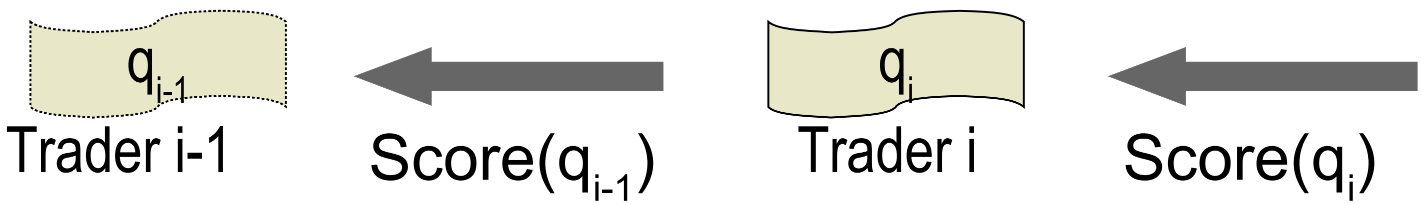
- Truthful reporting is “myopically” optimal strategy.
  - ie., if you rule out misleading other traders to make a long-term profit

Market maker's gain or loss

- Market operator can specify a maximum "endowment" for the market maker

- This controls:
  - how much MM can win or lose
  - how sensitive instantaneous prices are to one unit bought/sold
  - Equivalently: the constant multiplier 'b' in the underlying scoring rule:
    $S_{YES}(q) = b \log p \quad ; S_{NO}(q) = b \log(1\text{-}p)$

Information-theoretic characterization of profit



- Profit of trader *i:*

$$\boxed{\begin{aligned}&\text{Profit}(i) = \text{Score}(q_i) - \text{Score}(q_{i-1})\\&E(\text{Profit}(i)) = E[\text{Score}(q_i)] - E[\text{Score}(q_{i-1})]\end{aligned}}$$

- *Assuming all reports are truthful:*
  - *Expected Profit of trader i =*

$$b[\ H(\text{Event}\ |x_1, x_2, ..x_{i-1}) - H(\text{Event}\ |\ x_1, x_2, \ldots x_i)\ ]$$

Alternative view: automated market maker

The market scoring rule can also be viewed as an automated market maker

- "Instantaneous prices" are set based on current probability
- For log-MSR, if M shares on outcome X and N shares on outcome Y have been sold,

instantaneous price of X =

$$\frac{e^{M/b}}{e^{M/b} + e^{N/b}}$$

- updated for every little additional unit bought or sold
- Payoff of outcome that happens =1
- Trader buys/sells until price exactly matches her expected value for the security

Mathematically equivalent to previous description!

Market Scoring Rules : Summary of Basic Properties

[Hanson03]

- Truth-revealing is strategically optimal (myopically)

- Can be implemented as a price-setting market maker

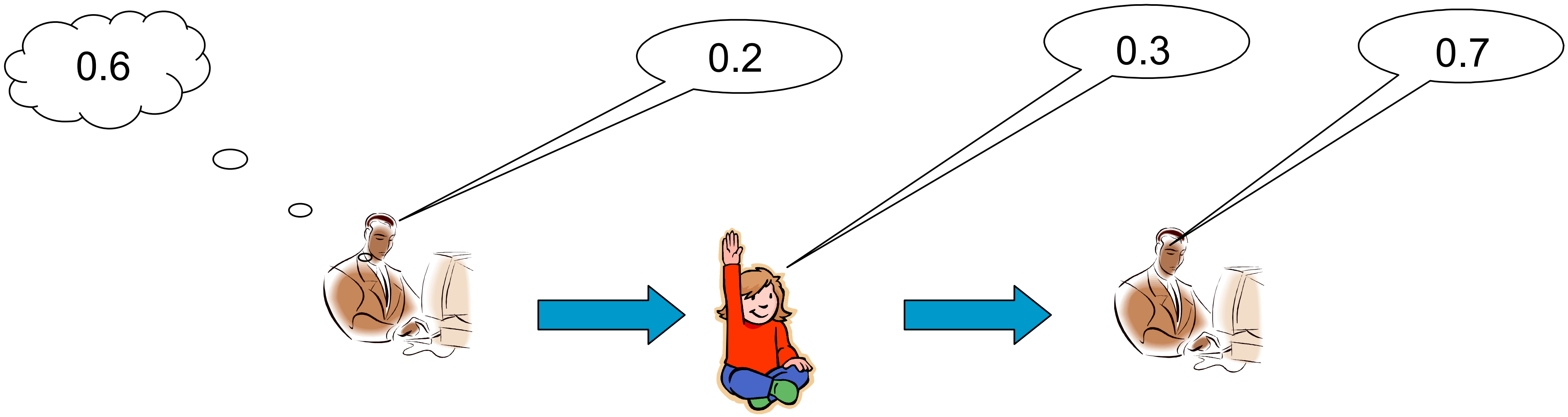
- Market-maker's loss / subsidy is bounded (regardless of number of traders)

- Expected profit connected to entropy

Long-term strategies?

Long-term Incentives to be Untruthful

Is it ever profitable to bluff and correct?



Motivating example

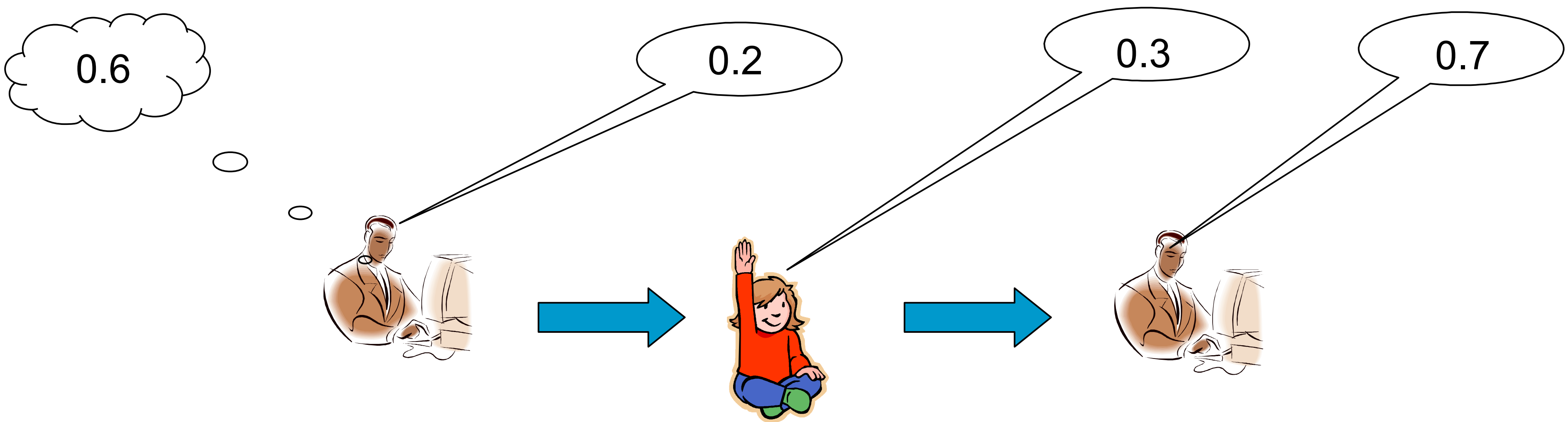
- Trader1 information:
  '1' : IND focus on batting (with prior 0.49),
  '0' IND focus on bowling (with prior 0.51)
- Trader2 information:
  '1' : AUS focus on batting (with prior 0.49),
  '0' : AUS focus on bowling (with prior 0.51)
- True outcome: XOR(x_1, x_2)

Motivating example

- True outcome: $XOR(x_1, x_2)$, prior = 0.5
- If Trader 1 sees '1':
  - Truthful:
    - Move price $0.5 \rightarrow 0.49$
    - Trader 2 moves from 0.49 to 1 or 0
  - Bluff:
    - Move price from $0.5 \rightarrow 0.51$
    - Trader 2 moves to 0 or 1
    - Trader 1 flips price to 1 or 0!

Single market: Is honest play optimal?

[Dimitrov, S. '07]



- Assumption: traders get **independent** signals
- Thm: Generically, honest play is not an equilibrium strategy

A different model: conditional independence

[Chen, Reeves, Pennock, Hanson, Fortnow, Gonen ‘07]

- Truthful reporting is an equilibrium strategy!
  - Assume that signals are **conditionally independent**, conditioned on the (unknown) true value

Resolving the different results:

[CDSRPHFG ‘10]:

- Critical factor: Are signals *substitutes* or *complements*?
  - Value of signal is reduction in entropy due to signal
  - May be different before/after knowing other signal
- Truthful reporting is an equilibrium in former case

Handling Complementarity: Discounted Market Scoring Rule

- One solution: discount profits over time [DS'07]
  - Second round payoff is $\delta b(\log q_2 - \log q_1)$,..
- Bluffing still possible..
- But, market converges to the optimal price:
  - Thm: In any weak-perfect bayesian equilibrium, the distribution of prices p_t after t trades each satisfies: $E\,[D(p^*||p_t)] < c\delta^{at}$

What happens in the real world?

Lab experiments are a good first step at testing theory predictions

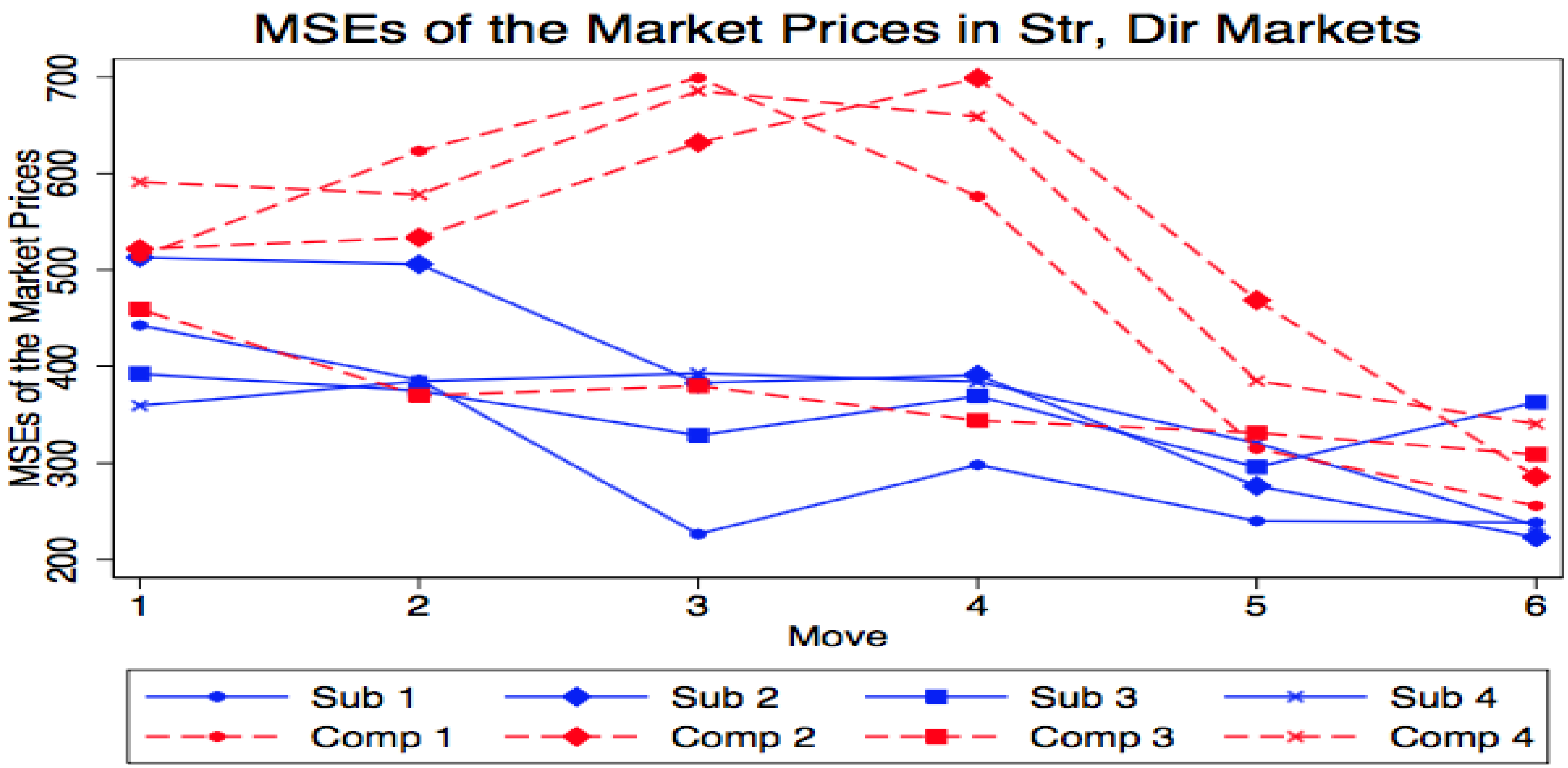
Experiments: Effect of information structure and market form

[Jian, S., 2010]

Experiment design

- 2-player markets with repeated play
- 8 treatments
  - Independent (complements) vs. CI (substitute)
  - Interface variations: prediction vs. trading
  - Structured vs. Unstructured
- 4 sessions/treatment, 8 subjects/session
- Measure: Intermediate and Final price accuracy

Results: Structured trading markets



- ✓ Result: Error in Complementary case after 2,4 rounds was significantly higher than in Substitutes case.

Results: Unstructured trading markets

- Result: Error in Complementary case after 2 trades was not significantly higher than in Substitutes case.

Takeaways:

- Theoretical model was predictive when trading format exactly matched model
-
- .. but real-world natural trading is more complex for participants (and analysts!)

- Bluffing strategies were used in both complements and substitutes treatments (more in complements)
- Aside: Structured trading helps with more effective aggregation

Conclusion

Prediction markets are an exciting class of mechanisms to study!

- Real-world applications and success stories
- Information-theoretic measures of value
- Rich strategic problems

A few directions for future work:

- Better modeling of real market microstructures
- More complete analysis of information settings and strategies
- Other market forms, scoring rules, etc.