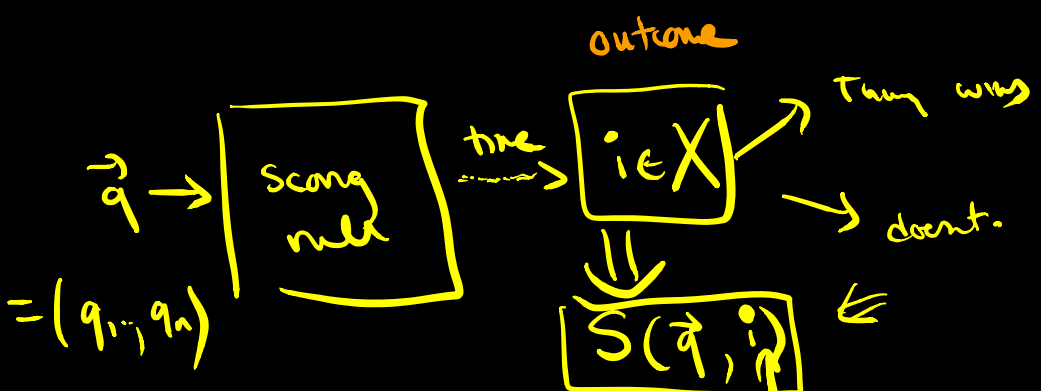
Scoring rules

- A different kind of mechanism design problem: how to elicit a good prediction of an uncertain event?
  - Weather forecaster: will it rain tomorrow?
  - Political pundit: will a Democrat or Republican win next election
  - Microsoft employee: will the next version of MS Office ship on time?
- How should we evaluate the quality of a prediction/pay based on the quality of predictions/ incentivize the work needed to output the best possible prediction?

Scoring rules

- X finite set of possible outcomes of uncertain event. $X = \{sun, rain, snow\}$
- A scoring rule is a real-valued function $S(\vec{q}, i)$
  - $\vec{q}$ is a probability distribution over X (a prediction)
  - i is some outcome in X (the realized outcome)



Model for incentives

- Forecaster has a belief $\vec{p}$, prob distribution over X.
- Forecaster will choose prediction $\vec{q}$ to maximize expected score

forecaster's goal: report $\vec{q}$ to maximize $E_{i \sim p}[S(\vec{q}, i)]$

report $\vec{q} \Rightarrow \sum_{i=1}^{n} p_i S(\vec{q}, i)$

Strictly proper scoring rules

- X finite set of possible outcomes of uncertain event.
- A scoring rule is a real-valued function S(q,i)
  - q is a probability distribution over X (a prediction)
  - i is some outcome in X (the realized outcome)

- A scoring rule is strictly proper if, no matter what the true belief $\vec{p}$ of the forecaster is, her unique best response is to report truthfully, i.e. to set $\vec{q} = \vec{p}$.

Strictly proper scoring rules

- X finite set of possible outcomes of uncertain event.
- A scoring rule is a real-valued function S(q,i)
  - q is a probability distribution over X (a prediction)
  - i is some outcome in X (the realized outcome)

$\sum_{i=1}^{n} p_i \, S(\vec{q}, i)$

- A scoring rule is strictly proper if, no matter what the true belief p of the forecaster is, her unique best response is to report truthfully, i.e. to set q = p.

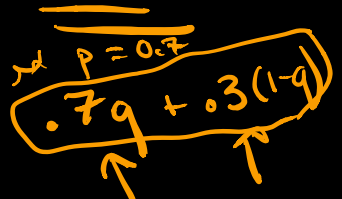
Example: $S(\vec{q}, i) = q_i$

baby $(p, 1-p)$

report $(q, 1-q)$

Exp payoff $= p\,q + (1-p)(1-q)$

given p what q maximizes this?



Quadratic scoring rule

$$S(\vec{q}, i) = q_i - \frac{1}{2}\sum_{j \in X} q_j^2$$

$q_i = 1$ for some i $\quad$ $q_j = 0$ $\forall j \neq i$ $\quad$ i happens. $1 - \frac{1}{2} = \frac{1}{2}$

if i doesn't happen $\quad -\frac{1}{2}$

$q_i = \frac{1}{n}$ $\quad$ no matter what payoff $\geq \frac{1}{2n}$

$$S(\vec{q}, i) = q_i - \frac{1}{2} \sum_{j \in X} q_j^2$$

GSR is strictly proper

$$E(\text{score}) = \sum_i p_i q_i - \frac{1}{2} \sum_{i \in X} p_i \sum_{j \in X} q_j^2$$

is max at $\vec{p} = \vec{q}$

$$\frac{\partial}{\partial q_k} E(\text{score}) = p_k - \sum_{i \in X} p_i q_k$$

$aS + b$

$+\frac{1}{2}$

$$q_k = \frac{p_k}{\sum_{i \in X} p_i} = p_k$$

$\sum_{i \in X} p_i = 1$

Logarithmic scoring rule

$S(q, i) = \ln q_i$

add $\ln |X|$ $\quad |X| = n$

forecasters can guarantee nonneg expability.

$\vec{q} = (\frac{1}{n}, \ldots \frac{1}{n})$

$E(\text{score}) = \sum_{i=1}^{n} p_i \ln(\frac{1}{n}) + \ln(n) = 0$

$-\ln(n)$

$q_i = 0$ | $q_i \geq \varepsilon$

Logarithmic scoring rule is strictly proper.

- incentivizes honest feedback
- prediction markets

Incentivizing honest feedback

- Example: peer grading, where students grade the assignments of other students.

- How to incentivize accurate grading, without direct verification?

Model

- n players (graders of an assignment, say in MOOC)
- Player i has a "signal" s_i → grader's true opinion of quality of homework.
- Each player submits a report r_i to a mechanism.
- Mechanism pays player $\pi_i(r_1, \ldots, r_n)$

Assume signals $(s_1, \ldots, s_n)$ drawn from correlated distn D.

Example:

n=2

		$s_2=0$ (bad)	$s_2=1$ (good)
bad	$s_1=0$	0.3	0.1
good	$s_1=1$	0.1	0.5
		0.4	0.6

$\Pr(s_2=0 \mid s_1=0) = \frac{3}{4}$

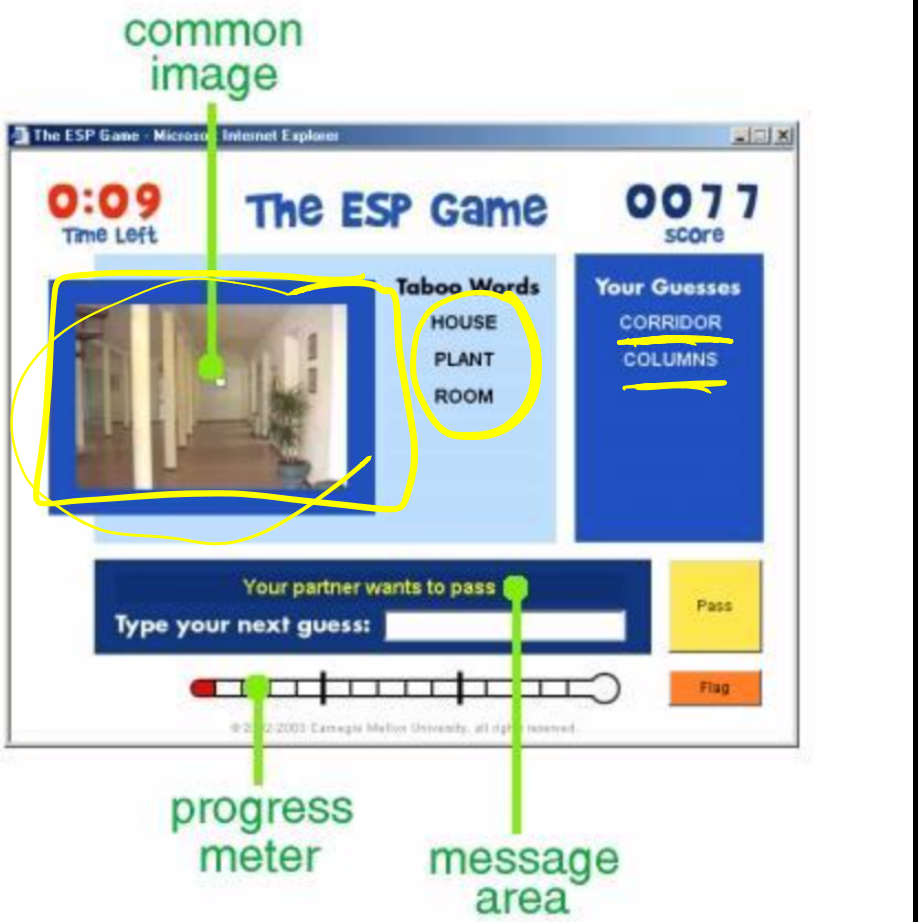
$\Pr(s_2=1 \mid s_1=0) = \frac{1}{4}$

How to choose payout rule $\pi_1(\vec{r}) \ldots \pi_n(\vec{r})$ to incentivize truthful reporting?

Output Agreement

reward agreement.

- For each player i
  - Pick a random player $j \neq i$
  - Set payoff π_i equal to 1 if they agree, 0 otherwise.


common image
The ESP Game - Microsoft Internet Explorer
0:09
Time Left
The ESP Game
0077
score
Taboo Words
HOUSE
PLANT
ROOM
Your Guesses
CORRIDOR
COLUMNS
Your partner wants to pass
Type your next guess:
Pass
Flag
progress meter
message area

Output Agreement

- For each player i
  - Pick a random player $j \neq i$
  - Set payoff π_i equal to 1 if they agree, 0 otherwise.

Is it a Nash eq to report truthfully

		bad $s_2=0$	good $s_2=1$
bad	$s_1=0$	0.3	0.1
good	$s_1=1$	0.1	0.5
		0.4	0.6

$s_2 = 0$

$\Pr(s_1=0 \mid s_2=0) = \frac{3}{4}$

	0	1
0	0.1	0.2
1	0.2	0.5

$\Pr(s_2 = x \mid s_1 = x) > \Pr(s_2 = y \mid s_1 = x) \quad \forall_y$

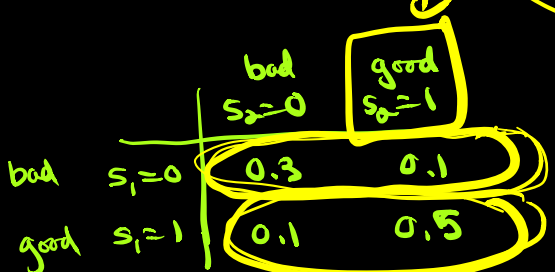
Mechanism has bad NE: everyone report good

Peer prediction mechanism

- Suppose the distribution D over signals is known to mechanism.
- For each player i
  - Pick a random player $j \neq i$
  - Let $D_j(r_i)$ be the distribution of s_j conditioned on $s_i = r_i$
  - Set i's payoff $\pi_i := S(D_j(r_i), r_j)$

treat player's report as a prediction of the distribution of other player's signal

$r_1 = 0 \Rightarrow$ $D_2(0 \mid r_1 = 0) = \frac{3}{4}$, $D_2(1 \mid r_1 = 0) = \frac{1}{4}$



Problems

- Requires advance knowledge of distribution.
- Other non-truthful and "bad" equilibria.

- In experiments:
  - Participants coordinate on high-payoff but uninformative equilibria
  - Empirically, people give better/truthful reports when paid a fixed reward (indep of their report).

Prediction Markets

- Suppose you're interested in an uncertain event e.g.,
  - Will Trump be reelected?
  - Will there be a Covid-19 vaccine by the end of 2020?
  - Who will win the next superbowl?

Pred markets: stock market for uncertain events like political outcomes

IEM PredictIt,

Prediction markets

- Idea: say want to predict which of two candidates A or B will win election.
- Create two securities a and b:
  - Each share of security a will pay out $1 if A wins.
  - Each share of security b will pay out $1 if B wins.
- Allow people to buy and sell these securities.
- Suppose current price of a is 75 cents (and b is 25 cents) and you believe A will win with probability p.
- What do you do?

PredictIt Markets Support Insights Leaderboards Login Sign Up

Dem. Nomination | Prez. Election | Donald Trump | Congress | U.S. Government | World

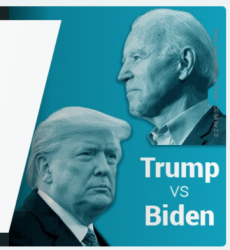
2020 presidential election winner?
Donald Trump
49¢ NC
Joe Biden
43¢ NC
77.8M Shares Traded

Electoral College margin of victory?
Dems by 60 - 99
11¢ NC
GOP by 60 - 99
11¢ 1¢▲
1.4M Shares Traded

Popular Vote margin of victory?
Dems by 1.5% - 3.0%
14¢ NC
Dems by 3.0% - 4.5%
14¢ NC
406K Shares Traded

Will Pence be 2020 GOP VP nominee?
Yes
81¢ 2¢▼
No
19¢ 2¢▲
455K Shares Traded

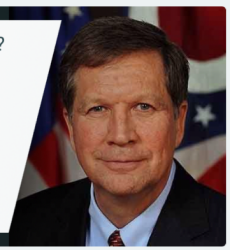
Trump vs. Biden RCP leader on 5/12?
Biden +4.4% to 4.7%
63¢ NC
Biden +4.8% to 5.1%
17¢ 2¢▲
108K Shares Traded



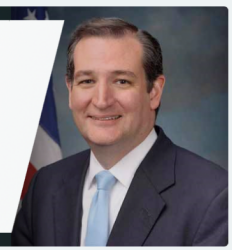
Will the 2020 GOP nominee be a woman?
Yes
5¢ 1¢▲
No
95¢ 1¢▼
74,571 Shares Traded

Will Zuckerberg run in 2020?
Yes
2¢ NC
No
98¢ NC
104K Shares Traded

Will Kasich run in 2020?
Yes
4¢ NC
No
96¢ NC
155K Shares Traded

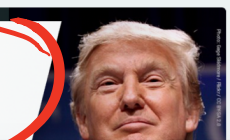


Will Cruz run in 2020?
Yes
4¢ NC
No
96¢ NC
95,612 Shares Traded

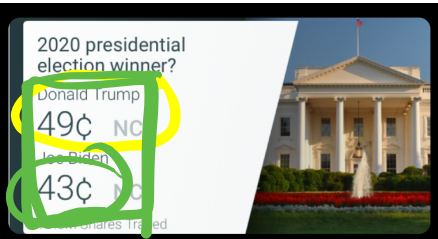


Will the Rock run in 2020?
Yes
1¢ NC

Will Trump be the 2020 GOP nominee?
Yes
90¢ NC



Will Scarborough run in 2020?
Yes
2¢ NC



you believe that chance that Trump will win is 52%

$$\text{Exp}(\text{payoff}) = \$1\ 0.52 - 0.49$$
$$= 0.03$$

Prediction markets

- Idea: say want to predict which of two candidates A or B will win election.
- Create two securities a and b:
  - Each share of security a will pay out $1 if A wins.
  - Each share of security b will pay out $1 if B wins.
- Allow people to buy and sell these securities.

- Interpret market price as the market's "belief" that the candidate will win the election.

- Market aggregating beliefs of all participants => "consensus opinion".

Legality Issues

- IEM, PredictIt circumvent regulation through a no-action letter by CFTC which condones IEM
  - Non-profit and used for research purposes
  - Stakes are small
- Several prediction markets with fictitious currency.
- No real path to establishing legal real-money prediction markets.

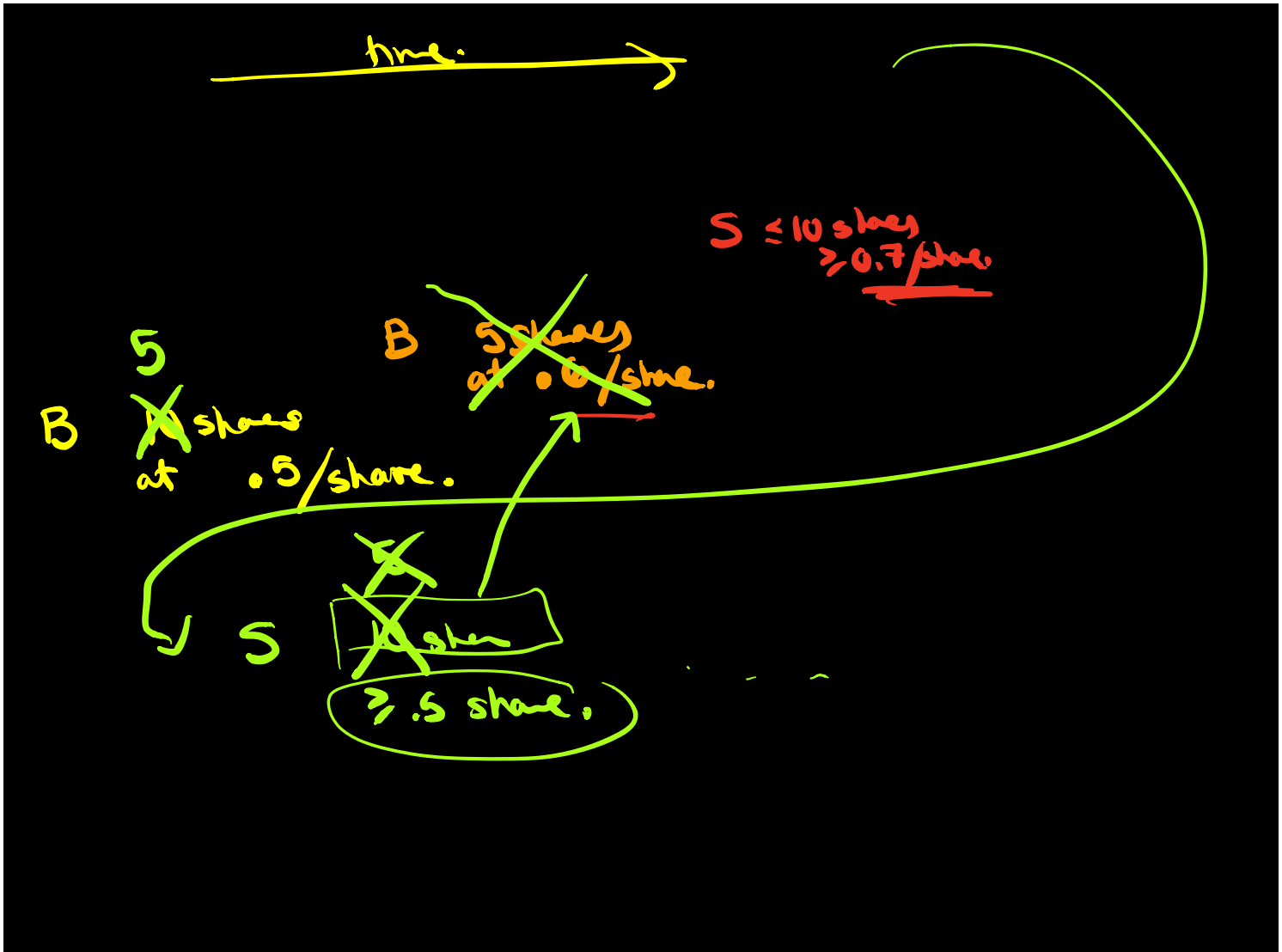
Accuracy

- Prediction markets vs polls
- Historically, prediction markets have done pretty well
  - People are better at predicting what other people will do than themselves.

bad in 2016

Basic prediction market (e.g. IEM)

- Use continuous double auctions
  - Trader can submit a buy or sell order any time.
  - An order:
    - Price
    - Max number of shares to be bought/sold.
    - Expiration date.
  - Trades are executed greedily (with nuances).


time
S ≤ 10 shares
≥ 0.7/share
B 5 shares at .5/share.
B 5 shares at .6/share.
S shares
≥ .5 share.

The New York Times

PLANS AND CRITICISMS; Pentagon Prepares A Futures Market On Terror Attacks

By Carl Hulse

July 29, 2003

The Pentagon office that proposed spying electronically on Americans to monitor potential terrorists has a new experiment. It is an online futures trading market, disclosed today by critics, in which anonymous speculators would bet on forecasting terrorist attacks, assassinations and coups.

Traders bullish on a biological attack on Israel or bearish on the chances of a North Korean missile strike would have the opportunity to bet on the likelihood of such events on a new Internet site established by the Defense Advanced Research Projects Agency.

The Pentagon called its latest idea a new way of predicting events and part of its search for the "broadest possible set of new ways to prevent terrorist attacks." Two Democratic senators who reported the plan called it morally repugnant and grotesque. The senators

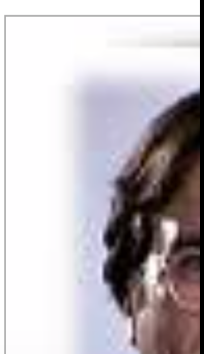
Pentagon kills 'terror futures market'

Senate urged Defense Dept. to scrap system to predict events

By John W. Schoen
msnbc.com

July 29 — A controversial plan to set up a "futures market" to use market forces to help predict political upheaval in the Middle East has been scrapped. The Pentagon Tuesday agreed to abandon the plan, the Senate Armed Services Committee chairman said, after Senate Democrats Monday blasted the plan as nothing more than state-sponsored "gambling on terrorism."

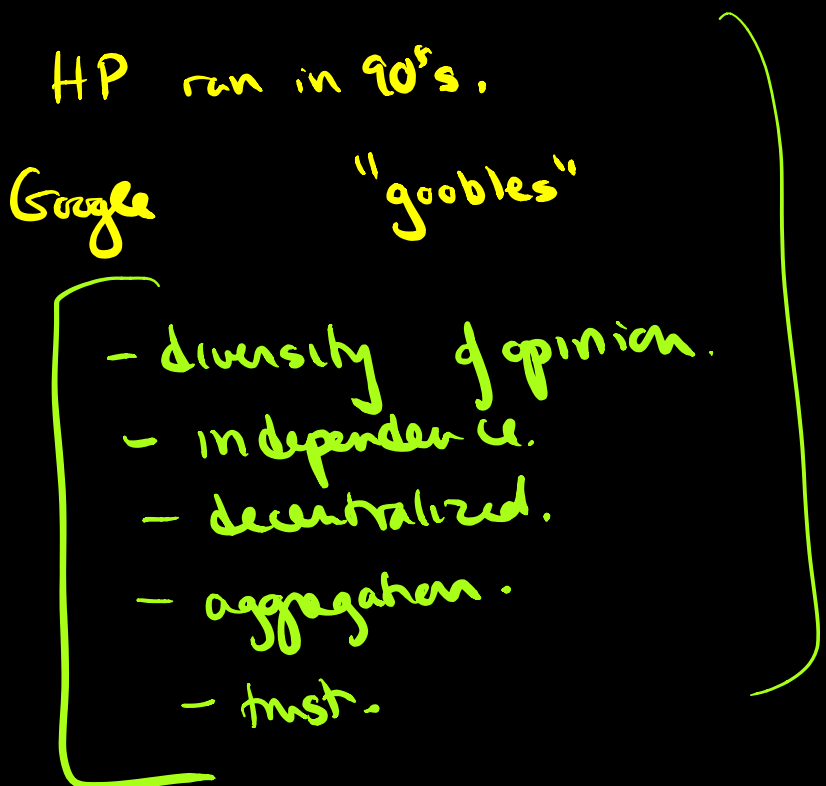
SEN. JOHN WARNER, R-Va., said Monday he spoke by phone with the program's director, "and we mutually agreed that this thing should be stopped."



Warner announced the decision not long after Senate Democratic Leader Thomas Daschle took to the floor to denounce the program as "an incentive actually to commit acts of terrorism."

"This is just wrong," declared Daschle, D-S.D.

The Wisdom of Crowds [Surowiecki] (2004)



Another Approach – Market Scoring Rules

- CDAs work well for "thick" markets – lots of traders, but not in
  - "thin" markets – few traders
  - "illiquid" markets -- large "bid-ask spread"
- Different approach: automated market-maker
  - At any time there is a price, and the market is always happy to buy or sell shares at this price.
  - Price evolves as shares are bought and sold.

Automated Market Makers

- Implemented using strictly proper scoring rule that is "shared" by all the players.
- Let S be a strictly proper scoring rule.

Initialize $p^0 = (\frac{1}{n}, \frac{1}{n}, \ldots \frac{1}{n})$ dist over X.

at any time t any player can update $p^{t-1} \rightarrow p^t$

When outcome $i \in X$ is realized, payout to player who $p^{t-1} \rightarrow p^t$ update is

$$S(p^t, i) - S(p^{t-1}, i)$$

paid out according to extent to which report improved prediction.

Properties:

① Market maker has bounded financial loss [specifically for log scoring rule] if it runs for T steps.

total payout $\underbrace{S(p^T, i)}_{-ve.} - S(p^0, i)$

$$-S(p^0, i) = -\ln\left(\frac{1}{n}\right) = \ln(n)$$

② If [players are myopic and each player trades once in a fixed order] then unique best response for each player is to update to their true belief.

$\vec{p}$ my true belief.

I will report p^* to max

$$E_{i \sim p}\left[S(p^*, i) - S(p^{t-1}, i)\right]$$

in best interest to report $p^* = p$.

Suppose betting on is whether 2 fair indep coins will both come up Hs.

$(p, 1-p)$

$\left(\frac{1}{4}, \frac{3}{4}\right)$

Suppose Alice knows outcome of coin 1 & she knows that its tails. should report $(0,1)$

$(.5, .5)$

& Bob knows outcome of 2nd coin toss & its Hs

$(1,0) \longrightarrow (0,1)$

What does this do?

- Player is rewarded according to extent her report improves the prediction.
- Final prediction is last distribution.
- Predictions tend to settle down.