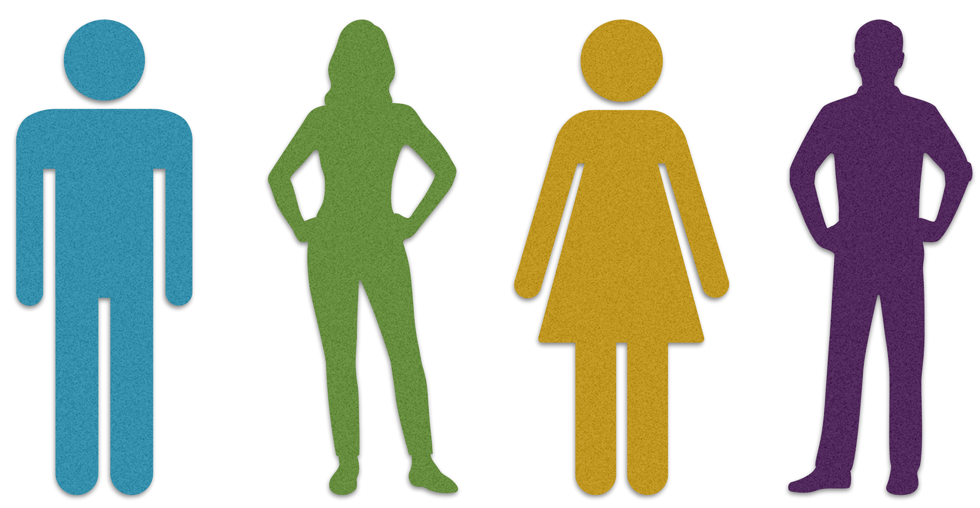
Eliciting Probabilities: Preventing Collusion

Rupert Freeman Microsoft Research NY

David M. Pennock Microsoft Research NY

Dominik Peters Carnegie Mellon University

Bo Waggoner CU Boulder



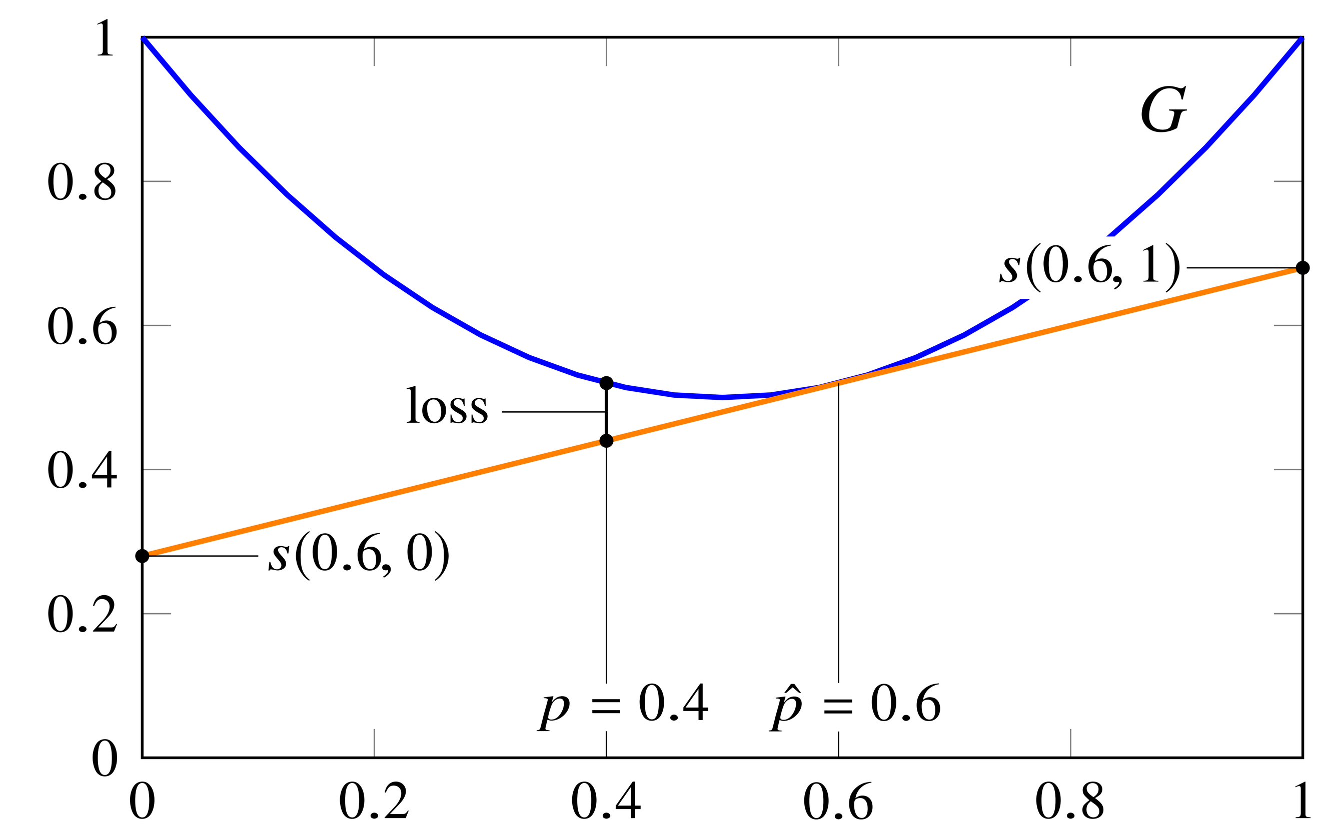
0.3	0.7	0.8	0.6
payment if it happens			
\$0.51	\$0.91	\$0.96	\$0.84
payment if it doesn't			
\$0.91	\$0.51	\$0.36	\$0.64

Agents estimate probability that an event occurs.
(e.g. that AAAI-21 gets > 10,000 submissions)
We want to know their estimates.
We offer them money to tell us.

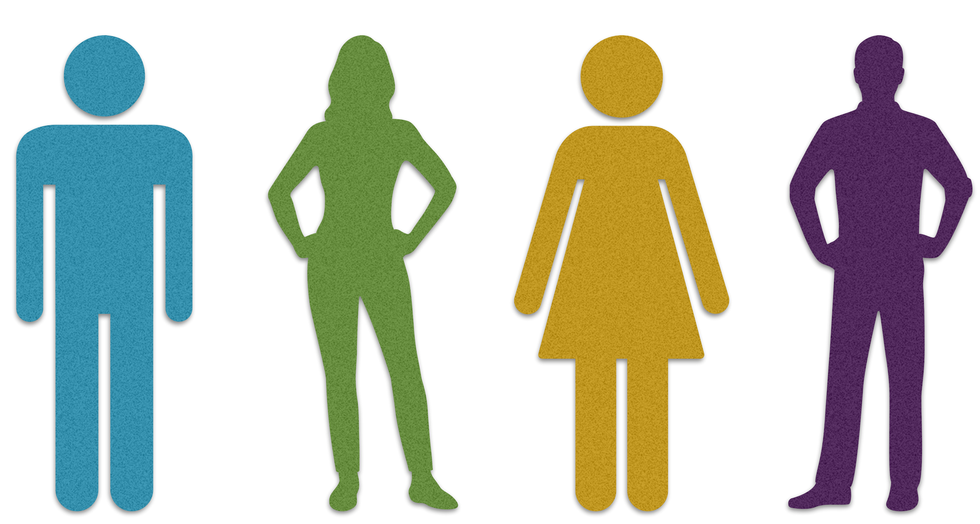
Brier's solution [1950]:

← pay $\$(1-(1-p)^2)$ if event happens

← pay $\$(1-p^2)$ if it does not



Theorem (Brier 1950): Agents uniquely maximize expected payoff with truthful reports.



0.6	0.6	0.6	0.6
payment if it happens			
\$0.84	\$0.84	\$0.84	\$0.84
payment if it doesn't			
\$0.64	\$0.64	\$0.64	\$0.64

French [1985]: Agents can collude!
If they share their beliefs and each report average,
we pay them strictly more whether or not event happens.
Agents can share the profits: a riskless manipulation.

Chun and Shachter [UAI 2011]: Similar riskless collusion possible for all scoring rules, for market scoring rules, and for competitive scoring rules.

Big question: Is there a payment scheme that avoids collusion, and incentivizes truthful reporting?

Mechanism 1

Strictly incentive-compatible

No collusion possible

if reports are bounded ($\varepsilon < p_i < 1-\varepsilon$)

Payoffs:

$$M_i^k(\hat{\mathbf{p}}, x) = \left(\sum_{j=1}^n \hat{p}_j - \frac{n}{2}\right)^k + k(x-\hat{p}_i)\left(\sum_{j=1}^n \hat{p}_j - \frac{n}{2}\right)^{k-1}$$

for k chosen large enough (depending on ε)

Mechanism 2

Weakly incentive-compatible

No collusion possible

Also "weakly distinguishing"

Payoff:

Pay agents using a 2-piecewise-linear scoring rule, plus the Brier score of the median report

Presented at AAAI-2020, New York City