Preventing Arbitrage from Collusion when Eliciting Probabilities

Rupert Freeman
Microsoft Research

David M. Pennock
Rutgers

Dominik Peters
Carnegie Mellon

Bo Waggoner
CU Boulder

Poster #71

Eliciting Probabilities

- We want to know the probability of an event, e.g., "AAAI-21 will get > 10,000 submissions"
- Experts have a belief about that probability
- We have some money lying around
- Idea: give money to experts in a way that incentivizes truth-telling (and high-quality estimates), by conditioning payment on report and outcome
- If someone reports p = 0.9, give them a lot of money if event occurs, and little money if it doesn't

Proper scoring rules

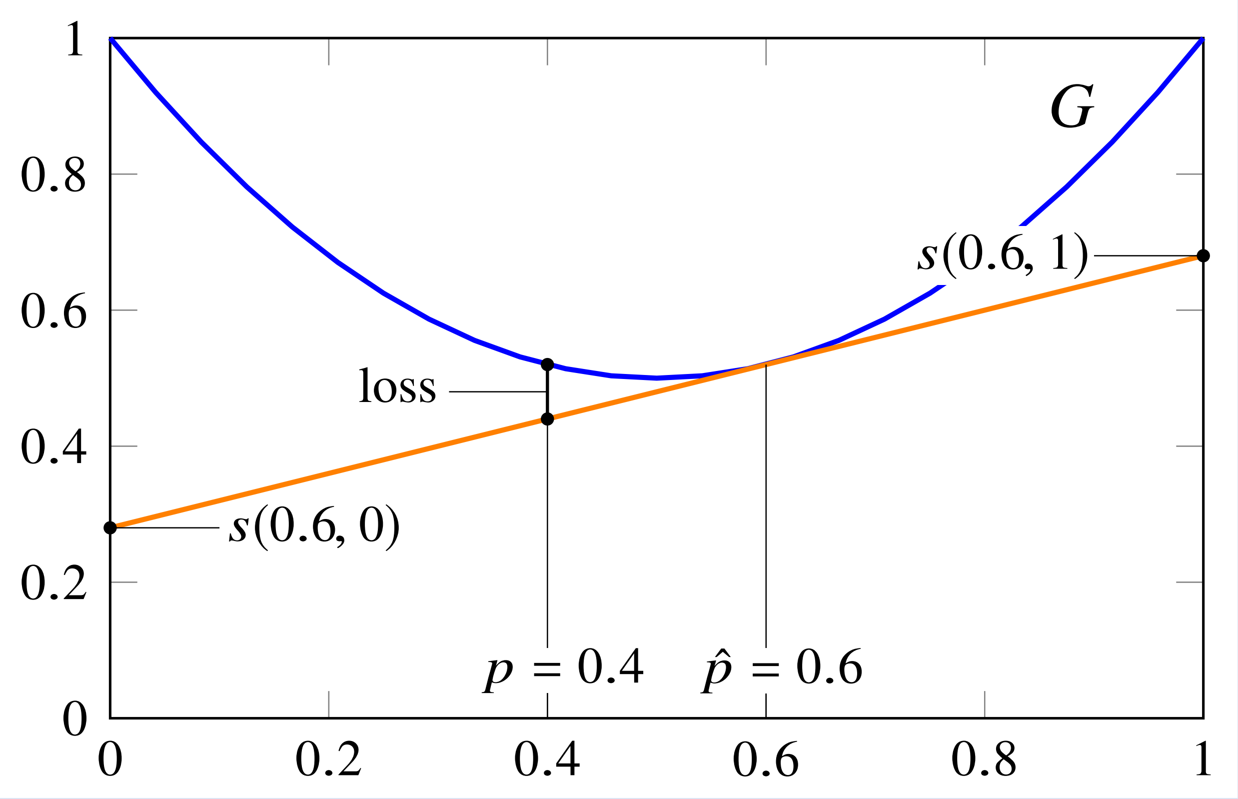
- Brier [1950] proposed such a payment scheme

- $s(\hat{p}, 0) = 1 - \hat{p}^2$
- $s(\hat{p}, 1) = 1 - (1 - \hat{p})^2$

	X = 0	X=1
$\hat{p} = 0.4$	$0.84	$0.64
$\hat{p} = 0.6$	$0.64	$0.84
$\hat{p} = 0.8$	$0.36	$0.96
$\hat{p} = 1.0$	$0.00	$1.00

- Easy calculus: if agent wants to maximize expected payout, it is uniquely optimal to report $\hat{p} = p$.
- Formally, $(1 - p)s(\hat{p}, 0) + ps(\hat{p}, 1)$ is uniquely maximized for $\hat{p} = p$.
- So: any misreport gives *strictly* less expected payout. This property is known as being *strictly proper.*

Proper scoring rules and strict convexity



Theorem (Savage 1971): Every strictly proper scoring rule
is defined by (sub)tangents of some strictly convex function *G*

Note: G(p) is the expected payout when truthfully reporting p.

Collusion and arbitrage opportunities

- Want to get estimates from multiple experts
- Easy! Just offer each of them a Brier payment
- Each expert has strict incentives to report truthfully
- French (1985) noticed a problem: if agents can collude, they can extract higher payments

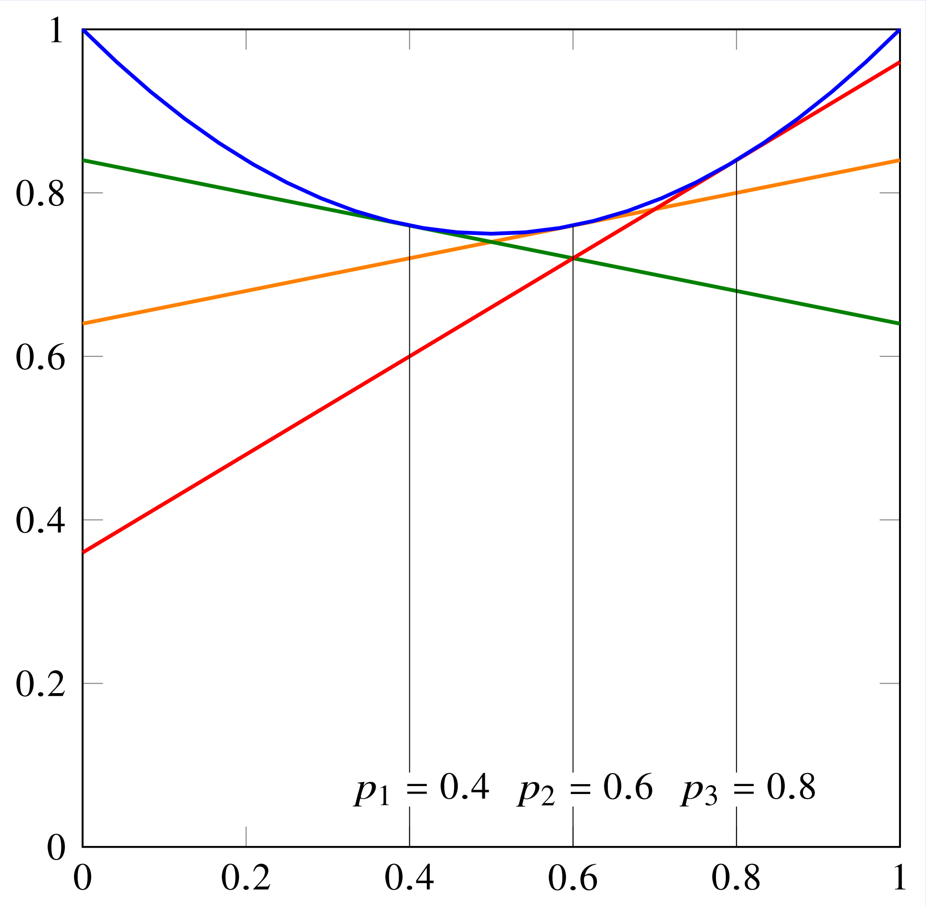
Collusion and arbitrage opportunities

- Assume:
  - Agents know each other
  - They can communicate beliefs before reporting
  - They can transfer money among themselves
- Then it is better for them to report their average belief
- Hopefully uncommon due to coordination difficulties
  - but forecasters sometimes work in groups (GJP), and there is a profit motive for intermediaries
- Bad:
  - If principal wants to aggregate reports, aggregate gets distorted
  - If agents all pretend to have the same belief, principal may be overconfident in aggregate
  - Difficult to identify the best forecasters

Collusion and arbitrage opportunities

	X = 0	X=1
$p_1 = 0.4$	\$0.84	\$0.64
$p_2 = 0.6$	\$0.64	\$0.84
$p_3 = 0.8$	\$0.36	\$0.96
$\sum$	**\$1.84**	**\$2.44**

	X = 0	X=1
$p_1 = 0.6$	\$0.64	\$0.84
$p_2 = 0.6$	\$0.64	\$0.84
$p_3 = 0.6$	\$0.64	\$0.84
$\sum$	**\$1.92**	**\$2.52**



Formal model

- A multi-agent payment scheme is a function $\Pi: [0,1]^n \times \{0,1\} \to \mathbb{R}^n$, so if $\boldsymbol{p} = (p_1, \ldots, p_n)$ is a vector of beliefs, then $\Pi_i(\boldsymbol{p}, x)$ is the payout to agent i in outcome x.
- Π is strictly proper if for each fixed reports $\boldsymbol{p}_{-i}$ of other agents, the induced scoring rule for i is strictly proper.
- Π admits arbitrage if there exists a coalition $C \subseteq N$, and vectors $\boldsymbol{q}$ and $\boldsymbol{r}$ with $q_i = r_i$ for all $i \notin C$ s.t.
  - $\sum_{i \in C} \Pi_i(\boldsymbol{q}, 0) \geq \sum_{i \in C} \Pi_i(r, 0)$ and
  - $\sum_{i \in C} \Pi_i(\boldsymbol{q}, 1) \geq \sum_{i \in C} \Pi_i(r, 1)$ and
  - one of these is strict.

Known results about arbitrage

- French (1985)
  - Every concave scoring rule admits arbitrage
- Chun and Shachter (UAI 2011)
  - Every scoring rule admits arbitrage
  - Market scoring rules (Hanson 2003) admit arbitrage
  - Competitive scoring rules (Kilgour and Gerchak 2004; Lambert et al. 2008) admit arbitrage
  - All these rules admit arbitrage at *every input profile* except when there is total agreement $p_1 = \cdots = p_n$.
  - "It is still an open question whether there is any strictly proper mechanism that does not admit arbitrage, but it seems unlikely."

Our Mechanisms

- We propose two payment schemes.
- Mechanism 1:
  - Strictly proper
  - Arbitrage-free for bounded reports $\epsilon \leq p_i \leq 1 - \epsilon$
    - bounding reports is a common restriction, e.g. in systems based on the logarithmic scoring rule, or on PredictIt
- Mechanism 2:
  - Weakly proper, and truth-telling is the only undominated strategy
  - Arbitrage-free

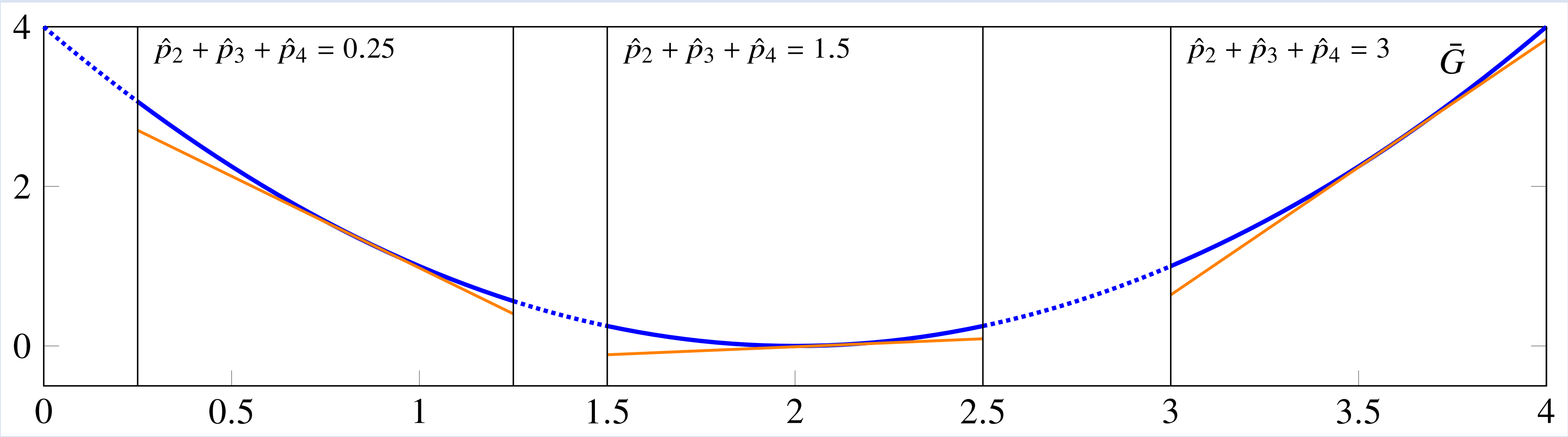
Mechanism 1

- Defined by tangents of $G(\widehat{p_i}) = \left(\sum_{j \in N} \widehat{p_j} - \frac{n}{2}\right)^k$, where k is an even integer
- For smaller ϵ, choose larger k
- Explicit formula:

$$\Pi_i(\boldsymbol{p}, x) = \left(\sum_{j \in N} \widehat{p_j} - \frac{n}{2}\right)^k + k(x - p_i)\left(\sum_{j \in N} \widehat{p_j} - \frac{n}{2}\right)^{k-1}$$

- If k is large and $\sum_{j \in N} \widehat{p_j} \approx \frac{n}{2}$, then payments are not very responsive to changes in individual reports.

Mechanism 1



Payouts to agent 1 (of a total of 4 agents). Agent 1 truthfully reports $p_1 = 0.6$.
Horizontal axis denotes the the sum $\sum_{i \in N} p_i$ of reports.

Proof idea for arbitrage-freeness:
total payment to a group *C* is a function of only the *sum* of their reports,
and this function is increasing for *x=1* and decreasing for *x=0* (for bounded reports).

Mechanism 2(a)

- Aim: get full arbitrage-freeness (w/o bounded reports)
- Weaken strictly proper to weakly proper
- Then it is possible to pay each agent independently while avoiding arbitrage.
- A scoring rule is *t*-choice if it is defined by a function *G* that is piecewise linear with *t* pieces.
- **Theorem**. Paying agents independently according to a weakly proper scoring rule *s* is arbitrage-free if and only if *s* is 1-choice or 2-choice.
- Example: If *x=1*, pay $1 to agents with report ≥ 0.5, and $0 to others. If *x=0*, pay agents with report ≤ 0.5.

Mechanism 2(b)

- Truth-telling is not the only undominated strategy in Mechanism 2(a).
- Alternative: pay each agent the Brier score of the median report $med(p_1, \ldots, p_n)$.
- Theorem. This payment scheme is arbitrage-free, weakly proper, and truth-telling is the only undominated strategy.
- But: this rule pays all agents the same. So if $p_1 = 0$ and $p_n = 1$, they get the same payment…

Mechanism 2(c)

- Idea: Use linear combination of 2(a) and 2(b) to get
  - the distinguishing payments of 2(a)
  - the undominated properness of 2(b)
  - the arbitrage-freeness of 2(a) and of 2(b)
- Distinguishing payments and undominated properness is preserved under linear combinations.
- But arbitrage-freeness is not: 50% of 2(a) + 50% of 2(b) admits arbitrage.
- **Theorem.** $1 - \epsilon$ of 2(a) + ϵ of 2(b) is arbitrage-free, where $\epsilon = 1/(n+1)$.

Beyond binary events

- Discussion has focused on yes/no events, $x \in \{0,1\}$
- All the notions make sense for events with several outcomes, e.g. number of submissions to AAAI-21 could be {<7k, 7k-8k, 8k-9k, >9k}.
- Agents then report a probability distribution over these outcomes.
- Our mechanisms can be extended to work for non-binary events using an inductive construction.

Conclusion

- Collusion and arbitrage are problems when using scoring rules in a multi-agent setting.
- Long-standing open question: can we avoid collusion while keeping individual incentives?
- We give partially positive answers.
- Open: is there a strictly proper scheme that is fully arbitrage-free?
- Open: Is there a mechanism similar to our Mechanism 1 that is more responsive to individual reports?
- Open: Might there be an impossibility when adding budget balance?

Preventing Arbitrage from Collusion when Eliciting Probabilities

Rupert Freeman
Microsoft Research

David M. Pennock
Rutgers

Dominik Peters
Carnegie Mellon

Bo Waggoner
CU Boulder

Poster #71