No-Regret and Incentive-Compatible Online Learning

Rupert Freeman (MSR NYC), David M. Pennock (DIMACS, Rutgers), Chara Podimata (Harvard) and Jennifer Wortman Vaughan (MSR NYC)

The Main Question

In deployments of the *prediction with expert advice* problem experts' reputation or financial compensation is often tied to the influence that they have on the prediction of the learner.

→ Experts might report differently from their **true** belief if that boosts their influence.

E.g., FiveThirtyEight ranks pollsters according to their accuracy:

FiveThirtyEight

Search for a pollster

POLLSTER	METHOD	LIVE CALLER WITH CELLPHONES	NCPP/ AAPOR/ ROPER	POLLS ANALYZED	ADVANCED +/-	PREDICTIVE +/-	538 GRADE	BANNED BY 538	MEAN-REVERTED BIAS
SurveyUSA	IVR/ online/ live		●	777	-1.1	-0.9	A		D+0.1
Rasmussen Reports/Pulse Opinion Research	IVR/ online			711	+0.2	+0.6	C+		R+1.5
Zogby Interactive/ JZ Analytics	Online			464	+0.6	+1.0	C		R+0.9
Mason-Dixon Polling & Research Inc.	Live	●		420	-0.5	-0.3	B+		R+0.7
Public Policy Polling	IVR/ online			411	-0.4	0.0	B		D+0.3
YouGov	Online			375	-0.4	+0.1	B		D+0.3

Do there exist online learning algorithms that are **simultaneously** no-regret for the learner and incentivize the experts to report their true beliefs?

Model

Repeated Interaction Protocol

For round $t \in [T]$:

1. Learner maintains a distribution $\boldsymbol{\pi}_t = (\pi_{1,t}, \dots, \pi_{K,t})$ over experts
2. (a) Each expert $i \in [K]$ has a private belief $b_{i,t} \in [0,1]$ about the outcome
   (b) Each expert $i \in [K]$ reports prediction $p_{i,t} \in [0,1]$
3. With probability $\boldsymbol{\pi}_{i,t}$, learner chooses prediction $\bar{p}_t = p_{i,t}$
4. Learner and experts observe event realization $r_t \in [0,1]$
5. All predictions have an associated bounded proper loss $\ell(\bar{p}_t, r_t)$ and $\ell(p_{i,t}, r_t)$
6. Learner updates $\boldsymbol{\pi}_t \to \boldsymbol{\pi}_{t+1}$

Full Information: All losses are observed
Partial Information: Only the loss of the learner's prediction is observed

Experts:

- Wish to increase leverage over the decisions of learner.
- Mathematically: $p_{i,t} = \arg\max \mathbb{E}_{r_t \sim \mathrm{Bern}(b_{i,t})}\left[\pi_{i,t+1} \middle| \left(b_{j,t}, p_{j,t}\right)_{j \in [K]}\right]$

[Roughgarden and Schrijvers '17]: Consider a similar problem to us, but experts care about unnormalized weights

Learner's Goals

1 **Incentive Compatibility**: Experts reporting their true belief maximizes their probability of being chosen, in expectation over $r_t \sim \mathrm{Bern}(b_{i,t})$

2 **Minimize Regret**

$$R(T) = \sum_{t=1}^{T} \ell(\bar{p}_t, r_t) - \min_{i \in [K]} \sum_{t=1}^{T} \ell(p_{i,t}, r_t) \le o(T)$$

Main Result

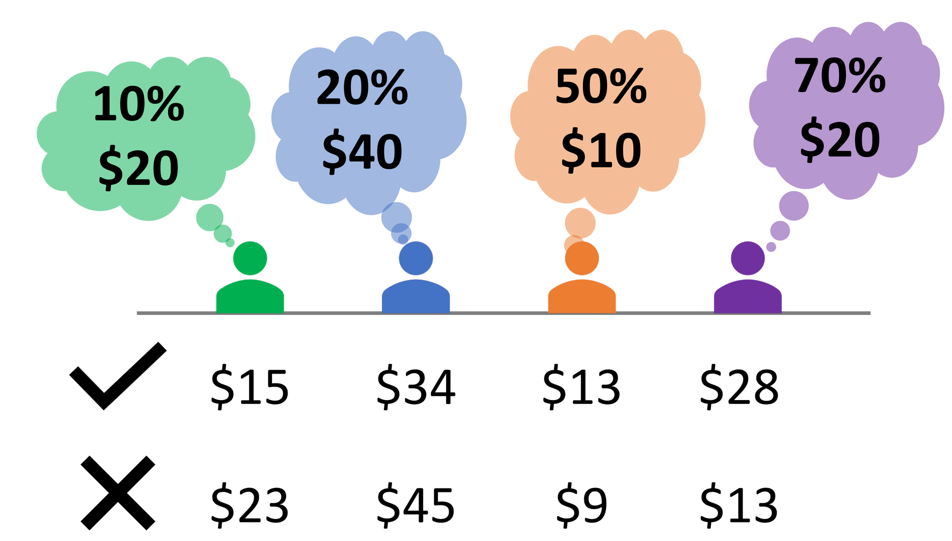
Two incentive-compatible online learning algorithms:

1) Weighted Score Update (WSU):
   $R(T) = O\left(\sqrt{T \ln K}\right)$ for the full information setting
2) Weighted Score Update with Uniform Exploration (WSU-UX):
   $R(T) = T^{\frac{2}{3}}(K \ln K)^{\frac{1}{3}}$ for the partial information setting

A Seemingly Unrelated Problem

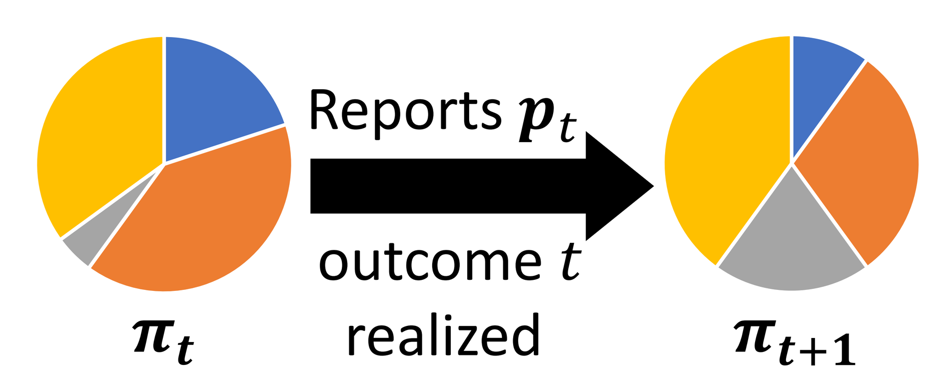
Construction of incentive-compatible wagering mechanisms

Probability of winning WNBA championship?



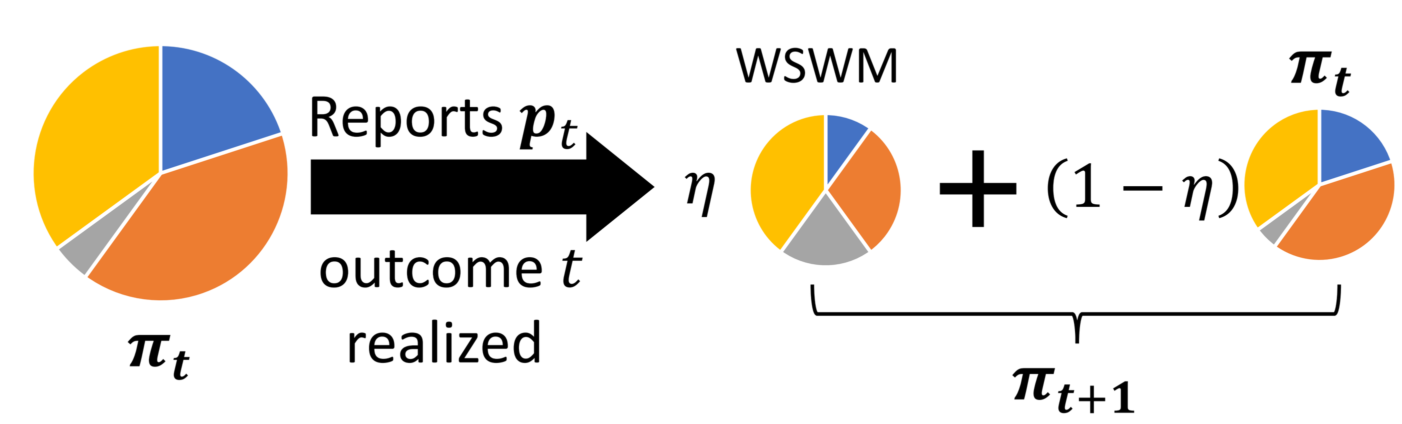
[Lambert et al. '08] defined a class of incentive-compatible wagering mechanisms: **W**eighted **S**core **W**agering **M**echanisms (WSWM)

Online Learning algorithms **are** Wagering Mechanisms!



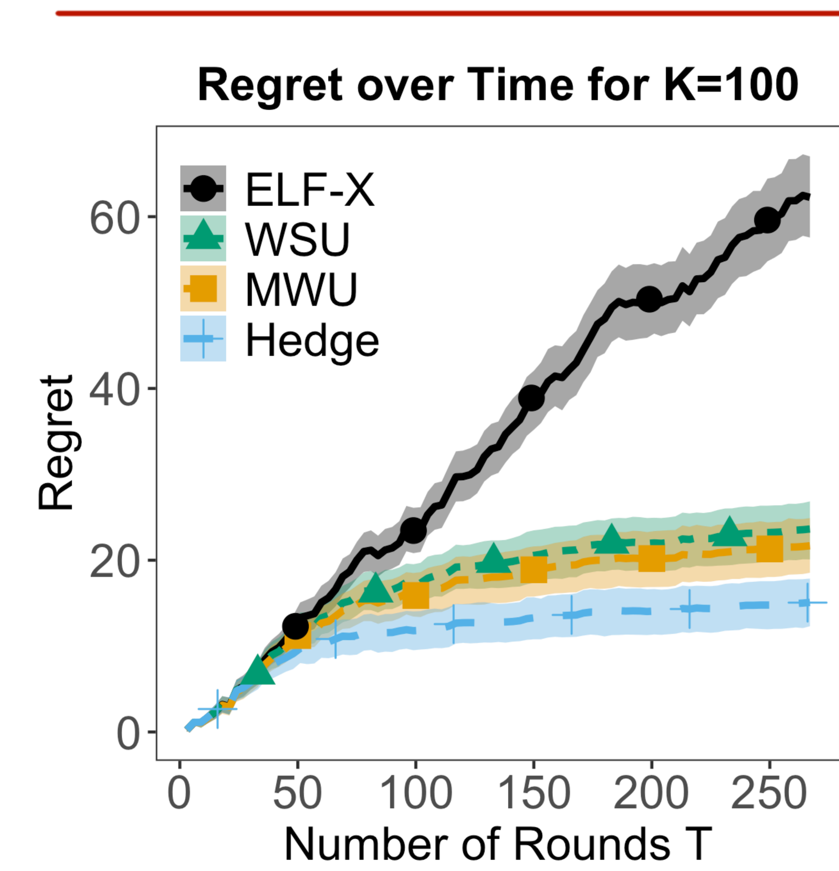
But they are not incentive compatible!

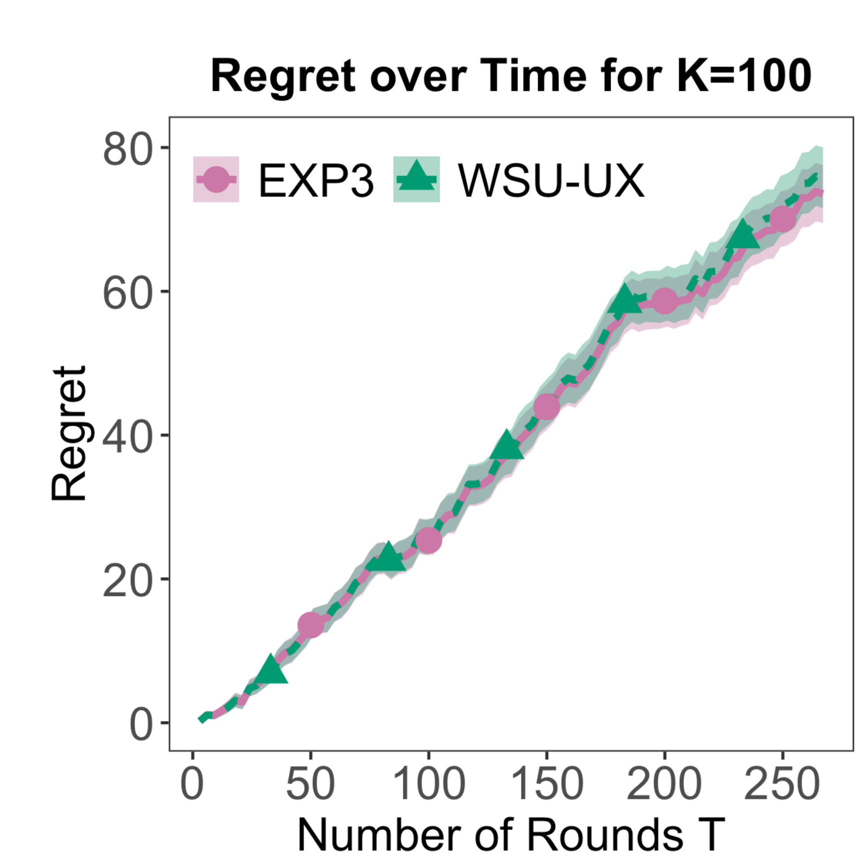
WSWMs Made No Regret



✓ truthfulness ✓ $R(T) \le \sqrt{T \ln K}$

Experiments on FiveThirtyEight NFL Datasets





ELF-X: A variant of ELF [Witkowski et al. '18] that is incentive compatible for forward-looking agents, but no proven regret guarantee.

References

[Lambert, Langford, Vaughan, Chen, Reeves, Shoham, Pennock, '08]. Self Financed Wagering Mechanisms for Forecasting, EC'08.
[Roughgarden and Schrijvers, '17]. Online Prediction with Selfish Experts. NeurIPS'17.
[Witkowski, Freeman, Vaughan, Pennock, Krause '18]. Incentive Compatible Forecasting Competitions. AAAI'18.