Background:
Trick or Treat: Putting Peer Prediction to the Test

Lihu Ben-Ezri-Ravin
Useful Vocabulary

● **Common Knowledge**
  ○ Every agent knows X. They all know that they all know X. They all know that they all know that they all know X. ad infinitum.

● **Bayes-Nash Equilibrium**
  ○ Everyone’s strategy is a best response, given their knowledge about the probability of the state of the world
Useful Vocabulary

- **Uninformative Equilibrium**
  - An equilibrium which doesn’t actually reveal anything about the state of the world

- **Informative Equilibrium**
  - The opposite; an equilibrium that does reveal something about the state of the world
Peer Prediction - Motivation

- How does one incentivize people to given honest opinions?
- Three problems:
  - Forming an opinion takes time and effort, but only benefits others
  - People may have reasons to be dishonest, e.g. conflict of interest, desire to be nice.
  - There’s no baseline truth against which to verify the opinion
Peer Prediction - Methods

- Peer-prediction method developed by Miller, Resnick, and Zeckhauser in 2005
  - Called MRZ mechanism in this paper
- Incentivizes truth-telling by comparing reports against each other
Peer Prediction - Methods
Peer Prediction - Methods

- Requires a payment rule which ensures that, for any given signal, a truthful response has more utility than a lie, given that everyone else is also responding truthfully.

\[ \sum_{s \in S} \text{Pr}(s|r_i)(u(r_i, s) - u(r_i^{\text{lie}}, s)) \geq 0, \forall r_i \in S \]
Peer Prediction - Problems

- Actually, there are three pure strategy equilibria
  - Always choose A
  - Always choose B
  - Always tell the truth
- Two of these are uninformative
- The peer-prediction method had never been tested on real people
Other Literature

● Other peer-prediction methods have been developed since 2005

● Only one, named Bayesian Truth Serum, has been experimentally evaluated
  ○ Assigns high scores to answers that are more common than collectively predicted
Other Literature

- *Incentives to Counter Bias in Human Computation*
  - Develop another peer-prediction mechanism, called Peer Truth Serum
  - They test it experimentally!
  - Claim it increases accuracy by up to 80%