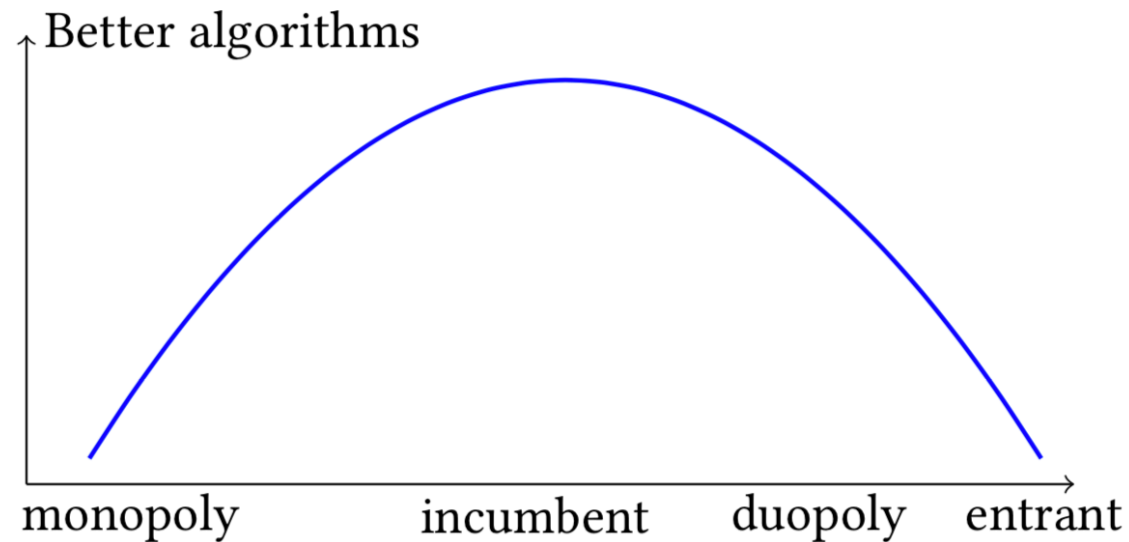


# The Perils of Exploration Under Competition: A Computational Modeling Approach

**Guy Aridor**, Kevin Liu, Alex Slivkins, Steven Wu

- Online platforms **compete** for users and **learn** from the data generated by them
  - Need to **explore** to gain information to make better decisions **tomorrow**
  - Need to **incentivize** consumers to pick me over competitors **today**
- Our model:
  - Two firms face the same K-armed multi-armed bandit instance.
  - Sequence of users choose firm to visit based on **reputation score**
  - Reputation score = sliding window average of rewards from previous users
- What bandit algorithms are incentivized under competition?
  - **Greedy** - pick what seems best based on the current information.
  - **Epsilon-Greedy** - random choice with epsilon probability, greedy otherwise
  - **Adaptive** - gradually zoom in on the best arm



- Our method: Study our model via numerical simulations
- Equilibrium algorithm selection depends on ***timing of entry***
- Simultaneous entry duopoly induces ***greedy*** algorithms in equilibrium
  - Exploration-focused algorithms fall into a ***death spiral*** vs greedy algorithms
    - Exploration -> Lower Reputation -> Fewer Users -> Lower Relative Reputation
  - Low consumer welfare
- Giving one firm a small ***first-mover*** advantage has two effects:
  - Incumbent incentivized to commit to best algorithm
  - Higher consumer welfare