Artificial intelligence, algorithmic pricing and collusion

Giacomo Calzolari (EUI & CEPR)
Emilio Calvano, Sergio Pastorello, Vincenzo Denicoló (U. Bologna)

April 26th, 2019, BECCLE Competition Policy Conference, Bergen 2019

Pricing algorithms are populating markets

- We are increasingly delegating choices to algorithms: product recommendations, content filtering, portfolio choices, **pricing**
- Chen et al. (2016) document that over 500 sellers active on top 1,641 Amazon listings use algorithmic pricing
- Pricing automation seen as a source of value (e.g. Amazon)

- Algorithms in pricing are not a temporary phenomenon (repricing industry)

Nothing new under the sun?

- Software/algo pricing is no news (since '80s e.g. hotels, airlines and financial markets): fixed rules (e.g. proxy bidding)
- Advancements in the field of AI spun a new class of algos:
  - **Model free**
  - Autonomously **learn from experience**
  - Increasingly available **off the shelf** (data availability and computing power)

Questions

- What is the consequence of AI pricing on price levels?
- Can AI agents autonomously 'learn' to cooperate/collude?
- Can we inform the current (and lively) policy debate?

Paper

- Run experiments with AI pricing-agents in controlled environments (computer simulations)
  - Algos must be similar to those used in markets
  - Environments must be realistic
- Workhorse IO model of competition: iterated price oligopoly with differentiated goods
- Extensive comparative statics on market and algos' parameters
- Extensive robustness checks

The Pricing Environment (an "MDP")

- Time steps $t = 0, 1, 2, ..., T$ (possibly $T = \infty$)
- State s_t (e.g. past prices)
- Action a_t^i (e.g. price)
- One step dynamics: $(s_t, a_t^i) \rightarrow (s_{t+1}, \pi_t^i)$

  where π_t^i is the Reward (profit)

- Agent's problem is to chose a policy $\sigma(s_t) = a_t^i$ that solves:

$$\max_{\sigma} \sum_{t=0}^{T} E[\delta^t \pi_t] \tag{1}$$

AI: Q-learning algorithm

At any period t a Q-algorithm...

- it decides whether to explore, with probability $\varepsilon_t = e^{-\beta t}$ (uniformly) randomizing over prices, or to optimize choosing the t-optimal price
  - $\beta \geq 0$ is **rate of experimentation**

- it stores the observed information (prices, profit), updating the "Q-matrix": present discounted value to each state-actions pair

$$Q_{t+1}(a, s) = (1 - \alpha) Q_t(a, s) + \alpha \left(\pi_t + \delta \max_{a_{t+1}} Q_t(a_{t+1}, s_{t+1}) \right)$$

  - news weighted according to **rate of learning** $\alpha \in [0, 1]$

Why Q-learning? Model free, simple, popular, successful

Baseline Environment: Economics

A workhorse model in IO: repeated game, simultaneous pricing

- 2 firms/algorithms
- Differentiated goods with Logit demand
- Constant marginal costs
- Symmetric
- Deterministic

Baseline Environment: Algorithms

- Actions: a^i_t 15 price points $(\underline{p} = \frac{9}{10} p_{\text{Nash}}, \ \overline{p} = \frac{11}{10} p_{\text{mon}})$
- State: $s_t = (p^1_{t-1}, p^2_{t-1})$ (1 period memory)
- Reward: π_t (deterministic)

- Algos interact with clones, i.e. Self-Learning and use Independent Learning

- Large grid (100x100) of parameters (α, β)
- Many values of δ, baseline $\delta = 0.95$
- 1000 sessions ("episodes") for each parametrization, beginning with random initial state (mean and sd)
- We give algos plenty of time to learn (off-line learning, as CS do, e.g. Alpha-Go, Self-Driving)

Literature

- Computer Science:
  - Vast and growing literature on multi-agent learning, e.g. traffic control, interacting robots, autonomous cars, ...
  - Few papers '90s (Tesauro et al. at IBM) on learning pricing algos, but special environments and "too much" special results
- Economics:
  - Learning in games, but mostly "passive" learning (best responding to observed behavior, e.g. BR-dynamics, fictitious play, Bayesian-learning, RL)
  - Milgrom and Roberts (1990) with supermodularity these learning methods tend to static Nash
  - Optimal experimentation (Bergeman and Valimaki 2006)
  - Evolutionary GT and Automata (fixed-strategies with replicatory dynamics of best performing)
  - Relation EGT and learning (stoch. approximation), but only simple models and no memory
  - Klein (2018), sequential-staggered pricing (Maskin-Tirole 1988, commitment and coordination)

Results

The Q-learning pricing-agents interacting repeatedly typically:

1. Learn to play (Converge)
2. Learn to reasonably behave (Consistency & **Equilibrium**)
3. Learn to charge high prices (**Cooperate**)
4. Learn to collude (**Strategies**)

2- Equilibrium: how to test

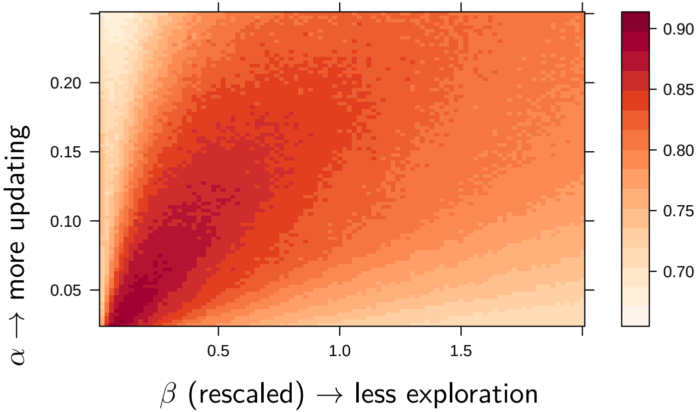
- There are $225^{15} \times 225^{15}$ (states^{prices} for 2 players) potential equilibrium candidates, impossible
- We instead check if an algorithm is best-responding to the strategy of the other algo
  - In particular from equilibrium-prices (**on path**)
  - but also **off-path** (subgame perfection)
  - and if not best-responding we calculate **how far**

2 Equilibrium: results

- For example $\alpha = 0.15$, $\beta = 3 \times 10^{-5}$ (cells visited at least 25 times)
- In 70% of sessions agents are **best-responding** 'on path.'
- In 61.5% **mutual best-reponse**, i.e. Nash Equilibrium.
- When they do not play Nash they are **pretty close** (1% profit gain left on the table)
- Hence, once they learned algos cannot be exploited
- Off-path, less best-response (but still small distance) as expected, these cells are less experimented

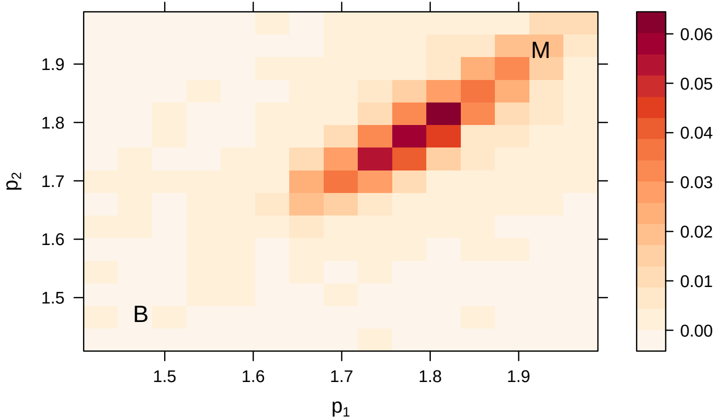
3 - Cooperation over the parameters-grid

Profit gain $\Delta(\alpha, \beta)$ (=0 Bertrand-Nash, =1 Monopoly)



- Cooperation from 60% to 99,1% Profit gain

3- Partial cooperation and price dispersion



How are competitive prices supported?

- Do agents **fail to learn** to compete? Or...

- Do agents actually learn to collude?
- Policy implications radically different (the first, we can go home, the second we must stay...)

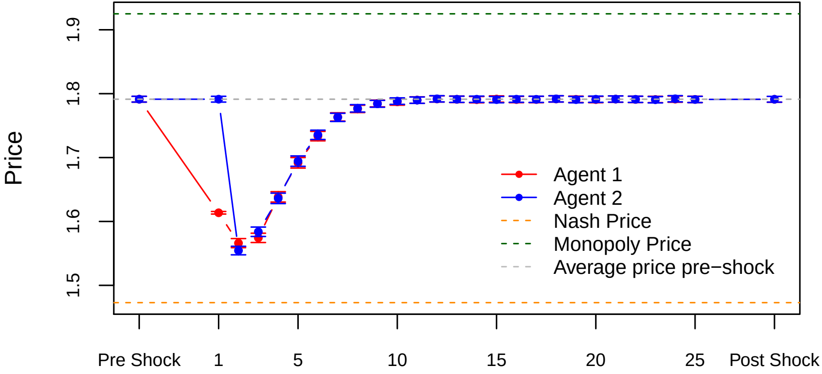
4- Learn to collude: Test 1

- What do they learn when collusion cannot be an equilibrium?
- Case 1: $\delta = 0$ (myopic)
- Case 2: $k = 0$ (no memory)

- Both cases $\Delta \approx 0$ (Max 15%)

4- Learn to collude: Impulse response of prices

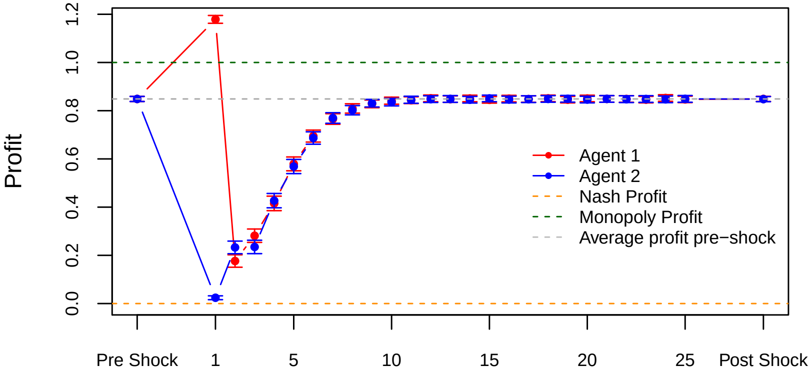
- Let agents play according to learnt strategies
- Agent 1 (Red) deviates: forced to charge lower price in $t = 1$



(parameters: $\delta = 0.95, \alpha = 0.15, \beta = 0.4 \times 10^{-5}$)

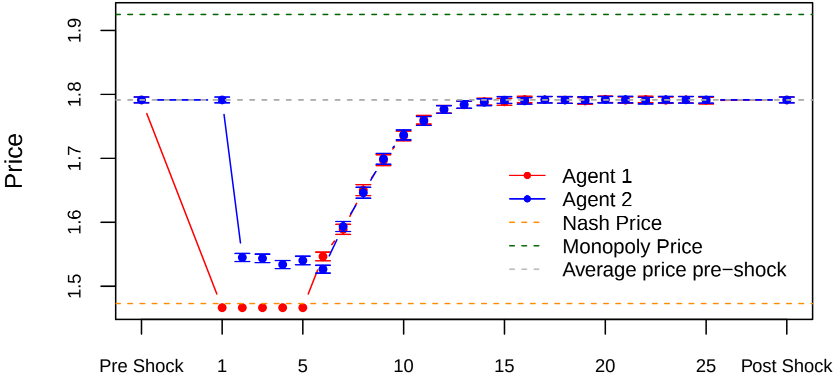
Same exercise, looking at profits

- Normalized $1 = \pi^{\text{collusive}}$



- Reaction of Blue makes deviation not profitable

5-period deviations



Robustness: more agents $N > 2$

- In the lab $N > 2$ **kills** cooperation (also experienced players).
- We looked at the case $N = 3$, $N = 4$

	$N = 2$	$N = 3$	$N = 4$
Δ	80%	74 %	70 %

- Algos are superhuman at coordinating.

Robustness: variable market structure

- $N \in \{2, 3\}$
- Outsider: enters and exists in random fashion
- entry/exit serially correlated
- $\Delta = 77\%$
- Same impulse response as before

Robustness: Asymmetric firms

- Collusion notoriously harder with asymmetric firms
- We looked at cost and demand asymmetries: results are similar

- E.g. cost asymmetry $c_1 = 1$

c_2	1	0.875	0.750	0.625	0.5	0.25
2's market share	50%	55%	60%	64%	67%	73%
Δ	80%	78%	74%	69%	65%	54%

Many Robustness Checks

Economic Environment:

- Change in δ
- Change in size of demand
- Change in product substitutability
- Different preferences (Singh and Vives)
- Stochastic demand

Increasing Complexity:

- More actions (30,50,100)
- Longer memory (2)
- Asymmetric learning: algos learning with different α and β
- Learning entrant
- Mixing different α and β after learning
- Continuous action space: DeepLearning (neural networks Value-function approximations) (in progress)

Where are we next?

- Dominant algorithm, Predatory algorithm
- AI algorithms with personalized and dynamic pricing: per-se and collusion
- Feeding actual market (big) data into learning algorithms

Policy. What to do?

We need to know more!

- Current legal approach inadequate with algos:
  - no intent
  - no mutual understanding
  - no explicit agreement
  - no communication
- Under current policy, algorithmic-collusion lawful

Possible approaches

1. Lasseiz Faire (algorithmic collusion just a theoretical curiosity)
2. Ban (the Sisyphus Luddite)
3. Regulation: ex-ante sand-boxing
4. Antitrust policy: ex-post look into algos

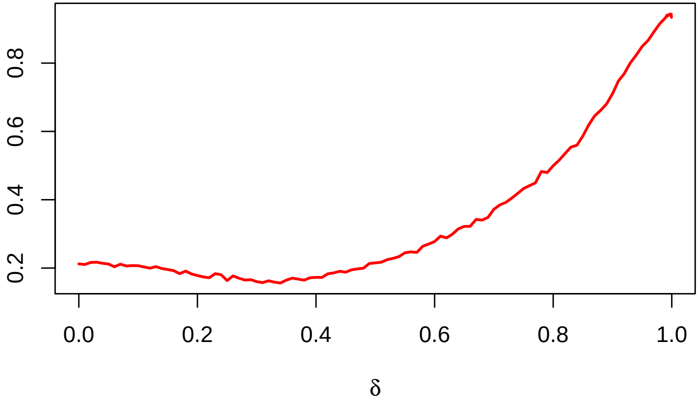
- Must reconsider balance between explicit and tacit collusion, else too many false negative with algos

THANK YOU

Artificial intelligence, algorithmic pricing and collusion

Giacomo Calzolari (EUI & CEPR)
Emilio Calvano, Sergio Pastorello, Vincenzo Denicoló (U. Bologna)

Effect of discount factor on profit gain Δ



δ