Market Design for AI Algorithms

MARTINO BANCHIO
and
ANDRZEJ SKRZYPACZ
Stanford Graduate School of Business

We discuss results from a few recent papers on how market design affects play between artificial intelligence (AI) algorithms. We describe the results from Banchio and Skrzypacz [2022] where the first-price auction appears more prone to collusion than the second-price auction. We also discuss results showing that information feedback after the auction, which allows computation of counterfactual payoffs, can improve competition.

Categories and Subject Descriptors: J4 [**Computer Applications**]: Social and Behavioral Sciences—*Economics*; I2.11 [**Computing Methodologies**]: Distributed Artificial Intelligence—*Multiagent systems*

General Terms: Collusion, Online Auctions, Independent Q-learning

1. INTRODUCTION

There is a growing interest in economics on how markets are going to change with the rise of economic agents using artificial intelligence algorithms to guide their actions. A first-order concern among regulators and the public is that algorithmic pricing might facilitate and enable collusive behavior in new markets. There is a burgeoning literature that simulates algorithmic competition under this lens, often showing that sophisticated (as in Calvano et al. [2020]) and simple (as in Asker et al. [2022]) algorithms playing repeatedly against each other learn to set prices above the static-game Nash equilibrium levels.

In this survey we discuss a few papers that contribute to the question about the future of markets in the era of AI: how does market design affect the conduct of AI algorithms? We start with our model of auctions in Banchio and Skrzypacz [2022]. We picked auctions as our application because bidders in online advertising auctions compete in thousands of auctions in any given day. It is common for them to rely on automated bidding tools, either provided by the auctioneer, developed in-house or purchased from third-parties. These automated decision-makers react quickly to environmental changes, but their rules often lack equilibrium thinking (to guarantee robustness of the algorithms and also to respond to the lack of full transparency in some markets). We show that the market rules matter for the outcomes even when standard Nash equilibrium theory would predict otherwise.

Online auctions are not the only setting where decisions are increasingly automated. Market design questions arise in finance, as pointed out by a recent paper (Colliard et al. [2022]) which studies the decisions of artificial intelligence algorithms in a Glosten and Milgrom model of market-making. In the context of online plat-

Authors' addresses: mbanchio@stanford.edu, skrz@stanford.edu

forms, Johnson et al. [2022] investigates marketplace design when pricing decisions are carried out by algorithms.

2. FIRST-PRICE VS. SECOND PRICE AUCTIONS

In Banchio and Skrzypacz [2022] we have shown that auction design has direct impact on revenues when the bidders use simple learning algorithms to optimize their bids. In a stylized environment, where despite the games having the same static Nash equilibria and very similar set of outcomes that can be sustained as equilibria of a repeated game, algorithms perform very differently.

2.1 Stylized Model of Auctions

Two bidders, Alice and Bob, participate in a sequence of auctions. In every period $t = 1, \ldots, \infty$ an auctioneer runs an auction to allocate a single non-divisible object to one of the bidders. Both bidders $i = A, B$ value the object at $v_i = 1$ and the value is constant over time. We compare bidding in the first-price auction (FPA) and the second-price auction (SPA).[1] Ties are broken randomly. We assume bidders choose from a finite grid of prices at 5 cent increments: $b_i \in \{0.05, 0.1, ...0.95\}$.

The payoff of the winner of period t auction is $\pi_t = 1 - p_t$ where p_t is the price determined by the bids according to the mechanism chosen by the auctioneer. The losing bidder gets payoff $\pi_t = 0$. Bidders maximize the expected sum of discounted per-period payoffs $\mathbb{E}\Big[\sum_t^\infty \gamma^t \pi_t\Big]$ where $\gamma < 1$ is the discount factor.

The one-shot second-price auction has a dominant strategy: bidding 0.95. The one-shot first-price auction has two pure-strategy Nash equilibria: both bidders bidding 0.9 and both bidding 0.95. In the paper we also discuss equilibria of the repeated games. First, collusion via strongly symmetric equilibria is slightly easier to sustain[2] in first-price auctions than in second-price auctions, with the differences solely driven by the grid in the bids (i.e. as the grid gets finer, the difference vanishes). Considering asymmetric equilibria, collusion via a bid rotation scheme is much easier to maintain in second-price auctions than in first-price auctions.

2.2 ε-greedy Q-learning Algorithms

We model the learning algorithms by assuming that each bidder takes decisions according to the recommendations of a ε-greedy Q-learning algorithm. Q-learning estimates the action-value function:

$$Q(b^i) = \mathbb{E}[\pi | b^i, b^{-i}] + \gamma \mathbb{E}[\max_{b'} Q(b')]$$

Notice that the optimal value is simply $V = \max_b Q(b)$. If the agent learns the Q-function, he can play the optimal strategy. The Q function starts from an initial vector Q (that we pick to be large values so the bidder experiments with all bids). In period t, after choosing action b_t the Q-function is updated according to:

$$Q_{t+1}(b_t) = (1-\lambda) Q_t(b_t) + \lambda\Big[\pi_t + \gamma \max_b Q_t(b)\Big]$$

[1]We consider also a richer family of auction formats parameterized by $\alpha \in [1, 2]$. In an $\alpha-$auction the highest bidder wins and pays a convex combination of the winning and the losing bid. The weight on the losing bid is $\alpha - 1$, and the weight on the winning bid is $2 - \alpha$.

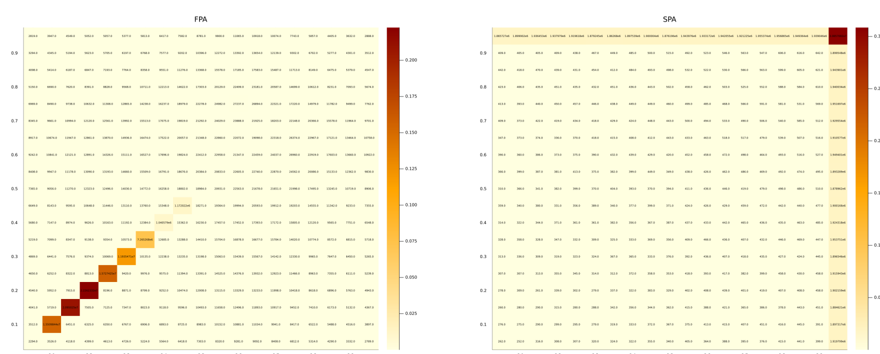
[2]An equilibrium is easier to sustain than another if its associated critical discount factor is lower.

This particular form of learning is asynchronous: only the state-action pair visited in a particular period is updated, while the rest of the Q-function remains constant. The hyperparameter λ is called learning rate. Its task is to discipline the speed of learning, and it determines the persistence of the estimates. In period t, bidder takes with probability $1-\varepsilon$ the greedy action, i.e. the action corresponding to the largest entry of her Q-vector. With the remaining probability she experiments and takes an action uniformly at random. This policy balances exploration and exploitation. Unlike other papers we focus on the case where ε is constant over time. This choice has two motivations. First, we want to make sure our simulations discover true long-term behavior of the algorithms and are not affected by a temporary coordination. Second, we think that in many real-life applications algorithms play in non-stationary environments and hence it would be optimal for them to never stop experimenting.[3]

2.3 Simulations

We compare bidding outcomes under the FPA and SPA formats. The outcomes are reported in Figure 1.

Result 2.1. *In a second-price auction, independent ε-greedy Q-learning algorithms learn the static Nash equilibrium. In a first-price auction, the algorithms repeatedly cycle between much lower bids.*



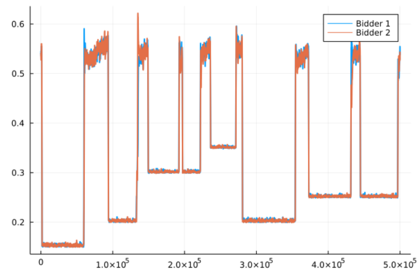
(a) FPA bids

(b) SPA bids

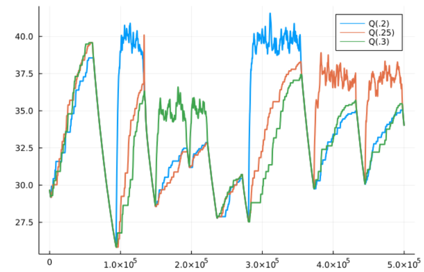
Fig. 1: Frequencies of bids from one simulation with 100,000,000 iterations. Parameters: $\varepsilon = 0.001, \lambda = 0.05, \gamma = 0.99$. Source: Banchio and Skrzypacz [2022].

To improve our understanding of what is going on in the first-price auctions, Figure 2 shows the evolution of bids (Panel a) and Q vectors (Panel b) over time. In this figure, Alice and Bob start at a (low) bid and submit (when not experimenting) that same bid for a while. Then one of them finds out a profitable deviation. Once that happens, a phase of intense exploration ensues, which lowers the average payoff to both bidders. Then, Alice and Bob re-coordinate on a pair of low bids, continuing in their cycle.

[3]For robustness we have checked that our main conclusions remain valid even when reducing the experimentation rate over time.



(a) Bids



(b) Q-values

Fig. 2: Dynamics of bids and their respective Q-values in FPA. Source: Banchio and Skrzypacz [2022].

What is going on? After Alice's upward deviation, Bob observes a drop in the winning rate and quickly updates downward the payoff of the old bid. That makes him try other bids at a faster rate than just the ε from experimentation. In turn, that makes Alice realize that the deviation is not as profitable as it appeared at first. When she does, she tries to go back to one of the lower bids. Sometimes the two of them get lucky and re-coordinate.

Importantly, as seen in Figure 2, while the algorithms coordinate on low bids, they do so imperfectly. First, they do not converge to constant low bids — our algorithms cannot converge to constant bids unless the bids form a static Nash equilibrium of the stage game. The reason is that the algorithms are good at learning a best response in case their opponent plays a constant action. Therefore, if they ever converged they would need to play mutual best responses. Second, even in periods when they do coordinate temporarily, they rarely do so at the lowest possible bids. The reason is that the forces to learn a profitable deviation make it hard for the two algorithms to coordinate on the most profitable profile of bids.

The point that algorithms just like humans face strategic uncertainty and hence would face the same problem of coordination on the many tacitly collusive equilibria has recently been emphasized by Kühn and Tadelis [2018]. This argument can affect regulators reviewing market in which one service provider offers algorithms to many participants. Even if the algorithm is not explicitly designed to facilitate collusion, the similarities in how the algorithm processes data collected by different players could contribute to tacit coordination.[4]

In the paper, we designed a series of experiments to isolate the forces behind the difference in FPA and SPA. The results suggest that the key difference between first-price and second-price auctions is the incentives they create when Alice and Bob bid different amounts. In both auction formats, the lower bidder has incentives to increase their bids to start winning again. The difference is in the incentives of the bidder bidding more: in a second price auction the payoffs are independent of the bid. But in the first-price auction the higher bidder wants to bid as little

[4]See the recent article in ProPublica Vogell [2022] that brings up this concern for the real estate market.

as possible as long as they still win. This additional force makes it easier for the bidders to re-coordinate on low bids.

3. PROVIDING FEEDBACK

A different market design question is what information should the auctioneer provide to the bidders after each auction. Of course, in practice there may be many reasons to provide or not to provide feedback. In Banchio and Skrzypacz [2022] we focus on a simple question: what happens with bidding if the auctioneer provides information about the bid of the opponent? With this information after every auction the Q-learning algorithms can update its estimates about all bids simultaneously using its counterfactual profits: what profit it would have accrued, had it bid differently in the last period.

Our result is that this additional feedback makes collusion disappear from the FPA. Similar results are presented by Asker et al. [2022], who discover low prices in a study of algorithmic pricing in a Bertrand pricing game. They also see much more competition after they provide their algorithms feedback that they use to calculate counterfactual payoffs (i.e., payoffs for all possible prices they could have set).

The choice of feedback turns out to be an important design consideration in online advertising. Recently, Google changed their display ad auctioning system from a SPA to a FPA. Alongside this change, bidders now observe the highest bid of their competitors. This fact squares well with the intuitions from our simulation: the ability to compute counterfactuals introduces an incentive to outbid the opponent that may be missing otherwise. This strongly suggests that market design choices that involve ex-post feedback may have a large impact on outcomes and consequently on revenues.

These results turn out to be much more general. The theoretical analysis in Banchio and Mantegazza [2022] shows precisely why feedback leads to convergence in a general fashion. With only partial feedback, the Q-learning algorithm suffers from sampling bias. Actions taken infrequently enjoy too much persistence, and their estimates do not adjust for undersampling. Uneven sampling alters the ordering between the estimates of the profits used to update the Q vector, leading to recurrent cycles. Synchronous Q-learning does not suffer from sampling bias, because counterfactual profits are estimated in every period with probability one.

These insights lead Banchio and Mantegazza [2022] to propose a theory of learning-robust mechanisms, where feedback provision guarantees that algorithms learn to play the dominant strategies of the strategy-proof mechanisms. The authors characterize the minimally informative feedback that guarantees competition, and they show it takes the form of a *menu description*: after the mechanism each agent observes the menu of outcomes she could have implemented, had she reported a different bid. With menus, algorithms can impute counterfactual payoffs correctly. Menu descriptions guarantee truthful implementation for a class of algorithms which includes the ones considered in Banchio and Skrzypacz [2022].

3.1 Potential Drawback of Post-auction Feedback

All the studies we have discussed so far assume that feedback is used to improve an algorithm's estimate of counterfactual payoffs. However, one concern that may arise from disclosure of private information is that such feedback may help the algorithms enforce punishment-reward schemes. For example, being able to monitor the bid of the winner simplifies the emergence of strongly symmetric schemes, where Alice and Bob can punish each other for deviating. On the other hand, bid rotation schemes do not require this kind of feedback to be implemented and they tend to be more profitable than strongly symmetric collusive equilibria (particularly so in the second-price auction).

To explore how information provision might affect the functioning of online advertising auctions, Decarolis et al. [2022] undertake a series of simulated experiments where bidders employ either Q-learning or a Neural Network approach. In both cases, bidders maintain memory of past play. Different information provision policies lead to differences in the ability of the algorithms to condition on past actions. The authors consider several settings with asymmetric bidders competing in both the generalized second-price (GSP) auction and the Vickrey-Clarke-Groves (VCG) auction. Decarolis et al. [2022] find that when more detailed information is available to the algorithms, the advertisers' rewards tend to be higher, and conversely, the auctioneer revenues tend to decline.

4. CONCLUSION

Algorithms facilitate more and more business decision making. In particular, algorithmic pricing and bidding continue to expand. That expansion will likely influence the future of markets. This new technology also creates new concerns. One of the central concerns is about potential (tacit or explicit) collusion by algorithms. It brings the questions of how to design and regulate markets in the era of algorithmic competition.

A set of recent papers that we described in this short survey contributes to this discussion by asking how market design can influence the market conduct of algorithms. These papers show that algorithms sometimes behave differently than a simple Nash equilibrium theory would predict. Moreover, we they have shown that market design matters. Changes to the market design that have no or minimal effect on the set of Nash equilibrium outcomes can have first-order impact on the behavior of algorithms.

Can market design avoid collusion in other ways? What market rules or market structures are sensitive to collusive behavior? Is communication necessary to restore competition, or can we rely on other market-based solutions? We hope future research will provide answers to these and other important questions at the core of market design for AI algorithms.

REFERENCES

Asker, J., Fershtman, C., and Pakes, A. 2022. Artificial Intelligence, Algorithm Design and Pricing. *AEA Papers and Proceedings 112*, 452–56.

Banchio, M. and Mantegazza, G. 2022. Adaptive algorithms and collusion via coupling. Working paper.

Banchio, M. and Skrzypacz, A. 2022. Artificial intelligence and auction design. In *Proceedings of the 23rd ACM Conference on Economics and Computation.* EC '22. Association for Computing Machinery, New York, NY, USA, 30–31.

Calvano, E., Calzolari, G., Denicolo, V., and Pastorello, S. 2020. Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review 110,* 10, 3267–97.

COLLIARD, J.-E., FOUCALT, T., AND LOVO, S. 2022. Algorithmic pricing and liquidity in securities markets. Working paper.

DECAROLIS, F., ROVIGATTI, G., ROVIGATTI, M., AND SHAKHGILDYAN, K. 2022. Artificial intelligence and data obfuscation: Algorithmic competition in digital advertising. Working paper.

JOHNSON, J. P., RHODES, A., AND WILDENBEEST, M. 2022. Platform design when sellers use pricing algorithms. Working paper.

KÜHN, K.-U. AND TADELIS, S. 2018. The economics of algorithmic pricing: Is collusion really inevitable? Working paper.

VOGELL, H. 2022. Rent going up? one company's algorithm could be why. *ProPublica*.