9^{th} International Conference on Computational Social Science IC^{2}S^2
July 17-20, 2023, Copenhagen, Denmark

Algorithmic and Human Collusion

Keywords: AI, Collusion, Experiment, Markets, Human–Machine Interaction

Extended Abstract

The use of autonomous pricing algorithms is on the rise in various industries (see European Commission, 2017). When firms use those tools, the pricing decision for a given product is outsourced from the human decision-maker to a computer algorithm. While in the past, most pricing algorithms have been rule-based with rules defined by the seller, there is a recent evolution towards self-learning algorithms (Ezrachi and Stucke, 2017). These self-learning algorithms develop strategies to achieve a specific goal, for instance, maximizing the firms' profits, without explicit instructions.

There are concerns among competition authorities (e.g., Bundeskartellamt and Autorité de la concurrence, 2019) and academic scholars (e.g., Ezrachi and Stucke, 2017) that pricing algorithms could not necessarily learn to price products more efficiently but also that there exists a possibility that they learn to collude tacitly. In other words, algorithms could learn by themselves that tacit collusion benefits the firm.

While recent papers by Calvano et al. (2020) and Klein (2021) show that algorithms can learn to be collusive, it is unclear whether pricing algorithms are more collusive than humans and therefore harm competition. Tacit collusion in traditional markets amongst human decision-makers is a well-documented phenomenon in both empirical and experimental economics Engel (2015). To assess the (anti-)competitive effects of algorithms, it is, therefore, necessary to establish a suitable baseline.

This paper provides a counterfactual for algorithmic collusion for a wide range of possible market compositions and highlights the impact of algorithms on competition. To examine whether commonly used self-learning algorithms make markets more collusive relative to the status quo of human collusion, I apply a two-step approach. In the first step, I consider self-learning pricing algorithms in an extensive simulation study to test whether algorithms learn to set supracompetitive prices and suitable strategies to support those prices as a collusive outcome. Here, I closely follow the approach from Calvano et al. (2020) but consider a different market environment that is more tractable. In the second step, I conduct market experiments in which humans compete either against each other or self-learned pricing algorithms. In the experiments, I closely mimic the market environment from the simulations. Across different treatments, I vary the market composition between algorithms and humans and the number of firms in the market. The experimental approach allows me to consider tacit collusion in a controlled setup and study the underlying mechanics. It enables me to observe humans and algorithms in the same environment and, thus, to analyze whether algorithms promote collusion. The resulting treatments are provided in Table 1.

I find evidence that algorithms foster tacit collusion in duopolies (see left panel of Figure 1). Two-firm markets with algorithms are always more collusive than human ones. In "mixed" markets, in which humans and algorithms compete with each other, self-learned pricing algorithms are as good as humans when colluding with the other market participant. Hence, pricing algorithms never promote competition but foster collusion if all firms use one. In three-firm markets, a non-linear relationship exists between the number of firms with a pricing algorithm and the level of tacit collusion (see right panel of Figure 1). Markets in which a single firm uses a pricing algorithm are more competitive than markets with only humans. However, as more firms use pricing algorithms, market prices can increase and may even exceed prices in human markets, especially if humans are inexperienced. Notably, the behavior of the algorithms is not myopic, but they learn to punish price deviations. Those punishment strategies make collusion incentive compatible and explain the outcomes (see Figure 2). As I consider a stylized market environment, I can interpret the strategies of the algorithms. The most successful algorithms learn a win-stay lose-shift strategy that is common for the iterated prisoner's dilemma (Nowak and Sigmund, 1993). The outcomes in mixed markets have a significant variance as humans choose heterogeneous strategies when playing against the algorithms.

The results highlight the potential anti-competitive effects of self-learning algorithms. While market outcomes vary depending on the exact parameterization and market composition, algorithms rarely foster but often weaken competition if they populate the market. Within the presented framework, the fear from competition authorities that algorithms can harm the competitive landscape is justified.

References

Bundeskartellamt and Autorité de la concurrence, "Algorithms and Competition," 2019.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello, "Artificial intelligence, algorithmic pricing, and collusion," *American Economic Review*, 2020, *110* (10), 3267–3297.

Engel, Christoph, "Tacit Collusion: The Neglected Experimental Evidence," *Journal of Empirical Legal Studies*, 2015, *12* (3), 537–577.

European Commission, "Final report on the E-commerce Sector Inquiry," 2017.

Ezrachi, Ariel and Maurice E. Stucke, "Artificial intelligence & collusion: When computers inhibit competition," *University of Illinois Law Review*, 2017, *2017* (5), 1775–1810.

Klein, Timo, "Autonomous algorithmic collusion: Q-learning under sequential pricing," *The RAND Journal of Economics*, 2021, *52* (3), 538–558.

Nowak, Martin and Karl Sigmund, "A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game," *Nature*, 1993, *364* (6432), 56–58.

Table 1: Treatment composition

Number of Human Players	Number of Algorithms 0	1	2	3
3	3H0A	-	-	-
2	2H0A	2H1A	-	-
1	-	1H1A	1H2A	-
0	-	-	0H2A	0H3A

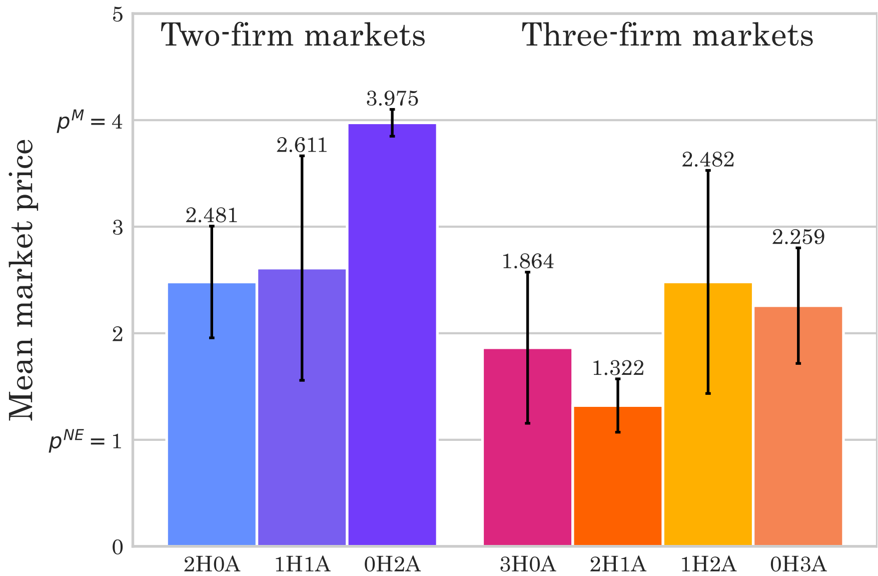

Figure 1: Average market prices for all treatments. The error bars represent the standard deviation. The abbreviations for the different treatments are introduced in Table 1.

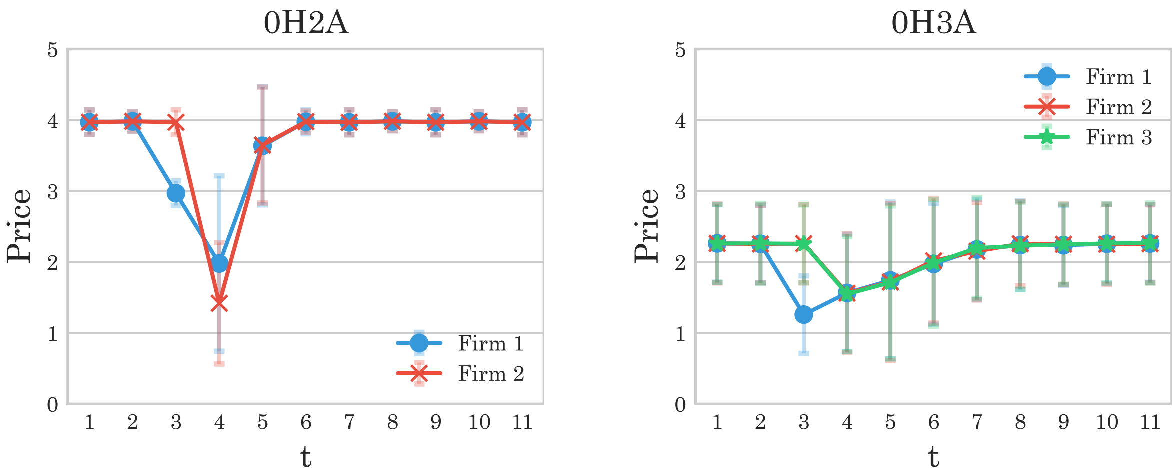

Figure 2: Punishment behavior of the algorithms after convergence. Starting from the state of convergence, the algorithms play according to their limit strategy. I induce an exogenous deviation from Firm 1 in $t = 3$ to observe the reaction of the other firms. I use 1,000 independent simulation runs. The error bars represent the standard deviation.