# The Price of Algorithmic Pricing: Investigating Collusion in a Market Simulation with AI Agents

Extended Abstract

Michael Schlechtinger
Department of Data Science
University of Mannheim
Mannheim, Germany
schlechtinger@uni-mannheim.de

Damaris Kosack
Department of Law
University of Mannheim
Mannheim, Germany
damaris.kosack@uni-mannheim.de

Heiko Paulheim
Department of Data Science
University of Mannheim
Mannheim, Germany
heiko.paulheim@uni-mannheim.de

Thomas Fetzer
Department of Law
University of Mannheim
Mannheim, Germany
fetzer@jura.uni-mannheim.de

Franz Krause
Department of Data Science
University of Mannheim
Mannheim, Germany
franz.krause@uni-mannheim.de

## ABSTRACT

Due to the rising availability and adoption of Artificial Intelligence in e-commerce, many of the online-prices are not set by humans, but by algorithms. The consequence is an opaque pricing situation that raises the potential of concealed, unfair competition by means of collusion. To examine this phenomenon, we study deep-Reinforcement-learning-based pricing algorithms by conducting an experiment involving an oligopoly model of repeated price competition. Our market model facilitates a variable environment spanning from economic theory to more realistic consumer demand models. We find that the algorithms learn to enter a collusive state and charge supra-competitive prices, without explicitly communicating with one another, and even without seeing each other's prices.

## KEYWORDS

Algorithmic Pricing; Deep Reinforcement Learning; Collusion

**ACM Reference Format:**
Michael Schlechtinger, Damaris Kosack, Heiko Paulheim, Thomas Fetzer, and Franz Krause. 2023. The Price of Algorithmic Pricing: Investigating Collusion in a Market Simulation with AI Agents: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023,* IFAAMAS, 3 pages.

## 1 INTRODUCTION

A common way to implement algorithmic pricing is the use of reinforcement learning (RL)-based algorithms, which have the tendency to end up in a state of collusion [6]. A market outcome must generally be considered legally neutral as long as it results from the competition in a market. If, however, the market outcome results from a concerted practice, this constitutes a cartel infringement.[1] According to the established case law of the European Court of Justice (ECJ), the characteristics of a concerted practice presuppose a minimum degree of coordination (concertation), a subsequent market conduct, and a causal link between the two.

[1]For a more detailed introduction to Art. 101 Treaty on the Functioning of the European Union ("TFEU"), see [6].

*Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.* 

## 2 EXPERIMENT DESIGN

We consider a stage game based on an oligopoly setting, which comprises $m \in \mathbb{N}$ consumers $Y = \{y_1, ..., y_m\}$ and $n \in \mathbb{N}$ firms $x = \{x_1, ..., x_n\}$ that simultaneously set the prices $P = \{p_1, ..., p_n\}$ so that $p_i \in [0, 2]$ holds for all $i \in \{1, ..., n\}$. We define a general selling or *demand probability* as follows:

$$d := d(\Omega) : \{1, ...n\} \to [0, 1]^n \quad \text{where} \sum_{i \in \{1,...,n\}} d_i = 1. \tag{1}$$

The parameter $\Omega$ represents the buyers' background knowledge, allowing an implementation of a custom buying probability. Thus, we define a Bertrand selling probability $d^b : \{1, ..., n\} \to [0, 1]$, given $p_{\min} = \min_{i \in \{1,...,n\}} p_i$ and $P_{min} := \{p \in P : p = p_{min}\}$ as

$$d_i^b = \begin{cases} 0, & p_i \neq p_{\min} \\ \frac{1}{|p_{\min}|}, & p_i = p_{\min} \end{cases}, \tag{2}$$

thus exclusively depending on the prices $P$. We add a more realistic buying behavior with a selection strategy based on a roulette wheel. Based on $\hat{p}_{\max}$ as the maximum price achievable in a market scenario, this modification results in the buying behavior

$$d_i^r = \frac{\hat{p}_{\max} - p_i}{\sum_{j \in \{1,...,n\}} (\hat{p}_{\max} - p_j)}. \tag{3}$$

In order to bridge theory and empiricism, we introduce the factor $\mu \in [0, 1]$. $\mu$ serves as a weight to gradually transition from one buying behavior to another. With this work, we combine the previously defined selection strategies $d^b$ in (2) and $d^r$ in (3) with $\sum_{i=1}^n d_i^b = \sum_{i=1}^n d_i^r = 1$ as

$$d^{\text{comb},\mu} = \mu * d^b + (1 - \mu) * d^r \quad \text{with} \quad \sum_{i=1}^n d_i^{\text{comb},\mu} = 1. \tag{4}$$

Due to this specific combination, $\mu$ acts as a bias. If it is set to 1, the products are perfect substitutes, if it is set to 0, the consumers' buying behavior is regulated by the roulette selection. With the parameters set, a seller can achieve a monopolistic price (MP) (i.e., the price that relates to the maximum revenue a monopolist can achieve in the market) of 1.50 (with a cumulative quantity of 50) and a competitive benchmark (CB) (i.e., the price one unit above the marginal costs) of 1.01 (with a cumulative quantity of 99). To create comparable results, we restrict the amount of consumers to $m = 200$ and thus result in a maximum price $\hat{p}_{\max} = 2$.

The RL agents are set up using a slight variation of the same baseline parametrization[2]. We employ a discrete action space which, compared to current literature[3], utilizes a relative way of action selection. This relies on the price set in the last episode $p_i(t-1)$. In order to discretize the actions (i.e., the price) within the economic environment, we generate an evenly spaced logarithmic distribution (i.e., 7 action gradients and a maximum value of the sequence of 2, resulting in $\mathcal{A} = [-2, -0.14142136, -0.01, 0, 0.01, 0.14142136, 2]$). The new price is calculated as follows:

$$p_i(t) = \ln(1 + e^{p_i(t-1)+a_i(t-1)}) \tag{5}$$

## 3 EXPERIMENTS

We investigate collusion in two different scenarios. Our baseline *Scenario A* depicts three agents that act based on the decisions proposed by their given algorithm. In *Scenario B* we manipulate each agent's state so they are only able to observe their own prices as opposed to every price on the market. We vary the number of agents (3, 5), the algorithm (PPO, DQN), and the bias $\mu$ (0, 0.5, 1). Every run is repeated 5 times, resulting in 90 runs overall (60 for Scenario A, 30 for Scenario B). A run comprises 5,000 episodes with 365 steps each. We average the prices of every step to an episode price as well as a step profit (average profit of all steps in an episode) before averaging every run within the same setting. Finally, we apply locally-weighted scatterplot smoothing (LOWESS)[5].[4]

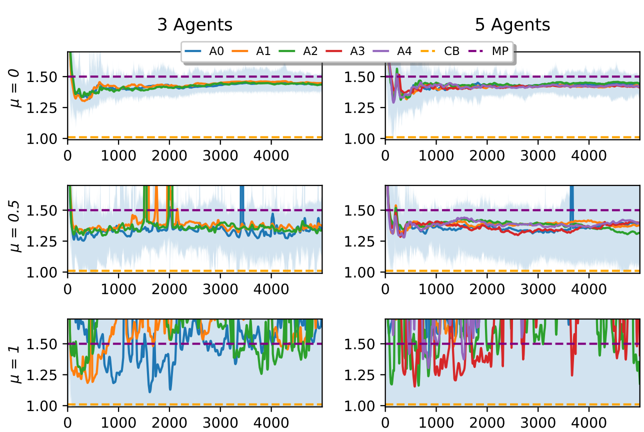


**Figure 1: Pricing in Scenario A**

[2]see [1] for tested implementations.
[3]cf. [2] and [4] for other discrete (deep) q-learning action space implementations where prices are absolute values in context to the CB and MP
[4]For the full implementation of this work, please refer to https://github.com/mschlechtinger/PriceOfAlgorithmicPricing.

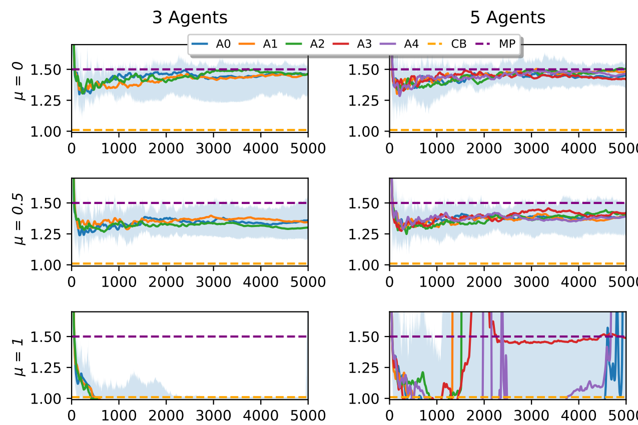


**Figure 2: Pricing in Scenario B**

## 4 DISCUSSION

In line with Calvano et. al [2], our experiments attest the deep RL agents' ability to charge supracompetitive prices and confirm their tendency to reach a collusive market outcome in a plethora of scenarios. Even when the agents were not able to reach the maximum reward (i.e., by charging the monopolistic price) they still set prices above the competitive benchmark. The large number of runs that were able to converge within 5,000 episodes result in a state of market stagnation.

The outcomes of scenario 2 show that the agents were able to charge supra-competitive prices without the need to receive their competitors' pricing information. This finding can have multiple implications; first, we see that the quality of a signal has an effect on the price setting. Hansen et al. [3] argued that the signal-to-noise ratio heavily affects results. Our results can only partly confirm those findings. While the average profit gain in Scenario B is slightly lower, we find that the overall outcome is generally identical despite this severe confinement. The resiliency most likely stems from the ability to approximate the prices via the reward function, in which other agents' prices embody unknown variables. This finding concurs well with Waltman and Kaymak [7], who observed that agents without a memory were still able to reach a collusive state.

Based on this outcome, the question arises whether the assumption that a collusive market outcome must generally be considered neutral, if a concertation cannot clearly be detected, should be upheld with regard to algorithmic pricing. The prerequisites of such a concertation are based on the logic of human behavior. It is questionable to what extent these prerequisites can equally be applied to RL-based decision-making processes. It could be argued that RL algorithms do not require any further reciprocity to gain the extra amount of trust in their competitors' expected next moves, which – in the case of human behavior – would be added through any minimal contact. The insecurity about the competitors' next moves is already reduced by the significant number of processed results from previous rounds, which are indistinguishable from other environmental information and therefore inherent to each decision (cf. Schlechtinger et al., [6]).

## ACKNOWLEDGMENTS

The work presented in this paper has been conducted in the KaReKoKI project, which is funded by the Baden-Württemberg Stiftung in the Responsible Artificial Intelligence program.

## REFERENCES

[1] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. 2021. Stable-Baselines3: Reliable Reinforcement Learning Implementations. *Journal of Machine Learning Research* 22, 268 (2021), 1–8. http://jmlr.org/papers/v22/20-1364.html

[2] Emilio Calvano, Giacomo Calzolari, Vincenzo Denicolo, and Sergio Pastorello. 2018. *Artificial Intelligence, Algorithmic Pricing and Collusion.* https://doi.org/10.2139/ssrn.3304991

[3] Karsten Hansen, Kanishka Misra, and Mallesh Pai. 2020. Algorithmic Collusion: Supra-competitive Prices via Independent Algorithms. *CEPR Discussion Papers* 14372 (2020). https://ideas.repec.org/p/cpr/ceprdp/14372.html

[4] Matthias Hettich. 2021. Algorithmic Collusion: Insights from Deep Learning. *SSRN Electronic Journal* (2021). https://doi.org/10.2139/ssrn.3785966

[5] Patrick Royston. 1992. Lowess Smoothing. *Stata Technical Bulletin* 1, 3 (1992). https://EconPapers.repec.org/RePEc:tsj:stbull:y:1992:v:1:i:3:gr6

[6] Michael Schlechtinger, Damaris Kosack, Heiko Paulheim, and Thomas Fetzer. 2021. Winning at Any Cost – Infringing the Cartel Prohibition With Reinforcement Learning. (2021). https://arxiv.org/pdf/2107.01856

[7] Ludo Waltman and Uzay Kaymak. 2008. Q-learning agents in a Cournot oligopoly model. *Journal of Economic Dynamics and Control* 32, 10 (2008), 3275–3293. https://doi.org/10.1016/j.jedc.2008.01.003