

Less than meets the eye: simultaneous experiments as a source of algorithmic seeming collusion

Thursday, June 20, 2024 2:00 PM (30 minutes)

This article challenges the idea of algorithmic collusion as proposed in Calvano et al. (2020) and subsequent literature. Identifying a critical mistake, we dispute the notion that supracompetitive prices result from collusive mechanisms where high prices are sustained by reward and punishment strategies. Instead, our analysis suggests that both phenomena originate from simultaneous experimentation and learning inertia inherent in reinforcement learning, without a causal link between them. Such seeming collusion can emerge rapidly in memoryless environments and with myopic agents, cautioning against misinterpreting the phenomena as collusion. Our findings advocate for simpler approaches to address algorithmic supra-competitive pricing issues.

below is an extended abstract:

Algorithmic decision-making has become ubiquitous in our lives, and its impact is increasing at an unprecedented rate. From our social media feeds to the stock market, from self-driving cars to medical diagnoses, algorithms are increasingly being used to automate decision-making processes.

In an influential paper Calvano et al. [2020b] (henceforth CCDP), show that basic and independent reinforcement algorithms, when trained simultaneously, consistently achieve supra-competitive outcomes. Furthermore, the responses of the algorithms to out-of-equilibrium stimuli resemble reward-punishment schemes that may be used to sustain collusion. The authors conclude that algorithms genuinely collude and provide policy recommendations and guidance to antitrust authorities. In particular, Calvano et al. [2020a] present tests, based on responses to stimuli, that regulators can employ to verify whether algorithms are autonomously engaging in collusion. This research has been followed up by many other studies that have used and expanded the notion of algorithmic collusion, such as Hettich [2021], Dolgoplov [2021], Banchio and Skrzypacz [2022], Klein [2021], Werner [2022], Qiu et al. [2022], Xu et al. [2023], to name just a few. Should these findings be confirmed, they could have significant implications for antitrust regulations, necessitating urgent action. It is therefore no surprise that the issue of Artificial Intelligence (AI) collusion has garnered attention from regulatory agencies, with it being a top priority on the agendas of many organizations (see e.g. OECD [2017], Autoridade Da Concorrência [2019], ACB [2019], Ezrachi and Stucke [2018], McSweeney and O’Dea [2017], Competition Bureau [2018] and Petit [2017]).

The results of CCDP and subsequent works are, however, increasingly debated. Critics raise possible methodological or design issues (Meylahn et al. [2022], Abada et al. [2022], Eschenbaum et al. [2022]), or question the interpretation of the results (Epivent and Lambin [2024], Abada and Lambin [2023]). Asker et al. [2022, 2023] emphasize the critical role of algorithmic learning protocols on supracompetitive limit prices. Calvano et al. [2023] claim that the “spurious” collusion results of Asker et al. [2023] are driven by the specific exploration mode they implement (synchronous learning), together with optimistic initialization of Q-matrices. Overall, the literature provides no formal explanation for the observational facts described in CCDP. This is mostly due to the fact it is notoriously difficult to draw theoretical results in multi-agent Q-learning processes.

This paper employs a simplified exploration procedure to elucidate the dynamics at play. It demonstrates that apparent collusion arises due to the specific learning process inherent in simultaneously-trained reinforcement learning algorithms: by construction, the initial valuations of actions are based on experiments that are performed while the other agents are also experimenting. These valuations may differ significantly from the profits observed in “play” mode when all agents play only their preferred (or “greedy”) actions. Still, the learning procedure is such that the erroneous valuations persist over time. When the rate of exploration decreases jointly, we show that agents may fail to identify profitable independent deviations and converge to prices (much) greater than Nash. Our theory is similar in spirit to that of Banchio and Mantegazza (2022), though we address much more general demand systems than the prisoner dilemma, with a specific application to the economic environment of CCDP. Our results are not restricted to cycles that possibly include cooperative actions, but also rationalize the convergence to singleton or fixed-point (supra-competitive) strategies, which represent 64 % of the simulations in CCDP. We use a mean-field assumption, eliminating the need for continuous time approximation of the learning process. Compared to Asker et al. (2022), we provide a complete characterization of the initial and final “greedy actions” in the Q-learning context, when endowed with a simple exploration procedure. We offer a comprehensive description of the underlying mechanism, including the characterization of the convergence point of the algorithm of CCDP. A notable contribution to the literature is

our explanation for the apparent reward and punishment schemes identified in CCDP and subsequent works, clarifying that these schemes are not the cause of the observed high prices.

Our theory is confronted with the results of CCDP, which we replicate faithfully. The first important step is to show that the memoryless version of CCDP also yields supra-competitive prices, which refutes the claim that the high prices are due to “genuine” collusion with high prices caused by reward and punishment schemes. In a second step, the results from our theoretical models are shown to explain the main observations of this literature: simultaneous learning causes the convergence to high prices. Finally, the apparent reward and punishment schemes are also shown to be spurious. From these observations, we note that the misinterpretation of high prices and apparent punishment schemes as evidence of genuine collusion has led to misguided policy recommendations. We propose to correct the interpretation and to implement more straightforward policy interventions against supracompetitive prices.

Primary author: LAMBIN, Xavier (ESSEC Business School)

Presenter: LAMBIN, Xavier (ESSEC Business School)

Session Classification: Parallel session: Algorithmic collusion: Foundations for understanding the emergence of anticompetitive behaviour