What Do We Know About Algorithmic Tacit Collusion?

BY AI DENG

RECENT YEARS HAVE SEEN LEGAL scholars and antitrust agencies express interest in and concerns with algorithmic collusion. As António Gomes, Head of the Competition Division at the Organization for Economic Cooperation and Development (OECD), stated in a recent interview, developing artificial intelligence (AI) and machine learning that enable algorithms more efficiently to achieve a collusive outcome is “the most complex and subtle way for companies to collude, without explicitly programming algorithms to do so.”

The type of algorithms that are capable of collusion (tacitly or explicitly) by themselves, without human interference, may sound far-fetched. Indeed, as one scholar puts it, “AAI [Antitrust and Artificial Intelligence] literature is the closet ever our field came to science-fiction.” It has also been stated that, regarding algorithms the “possibility of enhanced tacit collusion . . . remains theoretical.”

While these statements remain true, there is growing experimental evidence that an algorithm can be designed to tacitly collude. The possibility of tacit collusion through an algorithm is, in fact, not hard to see in some stylized cases. For example, assume that you and I are the only two online sellers of a homogeneous product and we know that our procurement costs are similar. Because our prices are posted online, we also know each other’s pricing. Suppose I use an algorithm that not only monitors your price but also sets my own price accordingly.

The way my algorithm works is as follows. First, it would raise and then keep my price high until you also change your price. If it turns out that you do not raise your price in response to my price increase, my algorithm would then drop my price to the cost of the product, or even below the cost. The low price “hurts” both your and my revenue. The algorithm would keep this “low price” regime for a period of time and then repeat the process of raising and then lowering prices if you do not raise your prices as well. After several rounds of interaction, it is possible that you realize that my algorithm appears to be sending you a signal: raise price with me or suffer financial losses. At that point, you might decide to reciprocate my price increase, given your and my interest in long-term profitability. This outcome is just as likely if you also use a pricing algorithm that tries to maximize your long-term profit. Notice that during the entire interaction, there are no traditional communications between us. We do not even need to know each other as long as all the conditions are met and intended learning is somehow achieved. Note the “reward-punishment” element in my algorithm, a point to which I will return.

Many have argued that the threat of algorithmic tacit collusion is real and poses even greater challenges for antitrust enforcement than human coordination and collusion. Maurice E. Stucke and Ariel Ezrachi postulate that AI, which enables computers to make decisions and learn through experiences autonomously, “can expand tacit collusion beyond price, beyond oligopolistic markets, and beyond easy detection.” This sentiment is echoed by Michal S. Gal, who discusses “tacit collusion among algorithms, reached without the need for a preliminary agreement among them.” Dylan I. Ballard and Amar S. Naik also argue, “Joint conduct by robots is likely to be different—harder to detect, more effective, more stable and persistent.” The background note by the OECD Secretariat also states that even though “[i]t is still not clear how machine learning algorithms may actually reach a collusive outcome . . . once it has been asserted that market conditions are prone to collusion, it is likely that algorithms learning faster than humans are also able through high-speed trial-and-error to eventually reach a cooperative equilibrium.”

These concerns naturally make one wonder what we should do about the possibility of algorithms reaching a collusive outcome without companies even intending that result.

For skeptics, the first questions are probably “Is it even possible?” and “Is there any real evidence that a machine could ever achieve that in actual markets?” In fact, some have expressed doubt as to the plausibility of autonomous algorithmic collusion. For example, the Competition Bureau of Canada recently pointed out the lack of evidence of such autonomous algorithmic collusion while recognizing the constantly evolving technology and business practices. One senior U.S. Department of Justice (DOJ) Antitrust Division official recently stated that “concerns about price fixing through algorithms stem from a lack of understanding of the technology, and that tacit collusion through such mechanisms is not illegal without an agreement among participants.”
Indeed, the existing legal principle states that an agreement requires a “conscious commitment to a common scheme designed to achieve an unlawful objective.” But in the context of this debate, the evaluation of the plausibility of tacit algorithmic collusion becomes an important exercise, even before we see concrete evidence that it has passed from theoretical possibility to marketplace reality. Moreover, insights about how algorithms may or may not come to collude are invaluable in focusing attention on the key legal and economic questions, policy dilemmas, and practical real-world evidence.

In this article, I survey and draw lessons from the literature on AI and on the economics of algorithmic tacit collusion. As we will see, while we do not have all the answers to all the questions that algorithmic collusion presents, a good understanding of this literature is a crucial first step to better understanding the antitrust risks of algorithmic pricing and devising better antitrust policies to mitigate those risks.

**Cartels’ Incentive Problem**

To better understand the problems a cartel must solve to sustain an agreement to restrict competition (e.g., raise prices or reduce output), it is instructive to look at the well-known Prisoners’ Dilemma. Imagine two accomplices of a crime are being interrogated in separate rooms and they cannot communicate. They must decide whether to confess to the crime and hence expose the other accomplice. Table 1 shows the consequences of their decisions.

**Table 1: A Prisoner’s Dilemma: Understanding the Incentive Problem of a Cartel**

<table>
<thead>
<tr>
<th></th>
<th>Confess</th>
<th>Not confess (Cooperate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prisoner A</td>
<td>(0, -3)</td>
<td>(-1, -1)</td>
</tr>
<tr>
<td>Prisoner B</td>
<td>(-2, -2)</td>
<td>(-3, 0)</td>
</tr>
</tbody>
</table>

The two rows and two columns in Table 1 represent the two prisoners and their two possible choices. For example, the cell (-1, -1) tells us that if neither of them confesses, each would get one year in prison. Similarly, if Prisoner A does not confess but Prisoner B does, then Prisoner A gets three years in jail and Prisoner B goes free; this corresponds to the upper right cell (-3, 0). Since the situation is symmetric, the lower left cell is (0, -3) and the penalties reversed. Finally, if both confess, then each would get two years (as shown in the lower right cell).

Given these numbers, it is clear that from a joint-interest perspective that the best outcome is (-1, -1), a total of two years. And the prisoners can achieve that by “cooperating,” i.e., not confessing. Unfortunately for the prisoners, since confessing is the rational move regardless of what the other does, both will end up confessing, leading to two years for each, an outcome strictly worse than the “cooperative” outcome. It is not surprising that cartel members face a similar type of incentive problem. They are both better off if they cooperate (say, raise prices or reduce output). But at the same time, if I know that my competitors are raising prices, I have an incentive to lower my prices to steal the business and increase my revenue. Since cartel members are usually smart enough not to write the cartel agreement into a contract, they have to find ways to enforce their agreement.

A critical point is that solving this incentive problem is key to the success of a cartel, whether it is humans or algorithms. In other words, the use of an algorithm does not magically remove this fundamental incentive problem that a cartel faces. Of course, unlike the “one-shot” situation in the standard Prisoners’ Dilemma, competitors interact with each other repeatedly in the market. It turns out that in repeated interactions, there is “more hope” that firms can learn to cooperate. In fact, repeated interaction is an important reason that tacit collusion emerges in the stylized example discussed earlier in the article.

**The AI Literature**

Is there any evidence that computer algorithms can (tacitly) collude? Empirically, we have not seen an actual case that involves tacitly colluding robots. The most well-known case was the Topkins case prosecuted by the DOJ in 2015, but in that case computer algorithms were used as a tool to implement a cartel agreement among humans.

Interestingly, there has also been some theoretical and experimental evidence that certain algorithms could lead to tacit collusion. One algorithm that has been found to be conducive to cooperative behavior in experimental settings is the so-called tit-for-tat algorithm (TFT). This strategy starts with cooperation, but then each party will just copy exactly what the opponent did in the previous period in repeated interaction. Intuitively, if two opponents start by cooperating, then the very definition of the TFT algorithm dictates their continued cooperation. But will competitors have an incentive to deviate from cooperation? The answer is that they might not. They understand that despite the higher profit they could obtain by cheating in the current period, they will have to compete with others and hence generate lower profit in the next period. While this is not guaranteed, if the firms care enough about future profitability, they might not find it worthwhile to deviate.

The TFT algorithm, despite its simplicity, intuitive appeal, and some experimental success, has a number of limitations. For example, to implement TFT, one needs to know what the competitors have done (because TFT copies the competitor behavior) and the consequences of future interactions (because they need to assess if it pays to cooperate). In the real world, firms typically do not possess that information, except in certain special cases.

In recent years, there has been more research that aims to relax various assumptions and construct more robust algo-
rithms. In a recent study, a team of researchers designed an expert system (a type of AI technology) that enables better coordination among the opposing players in a variety of situations. The study found that although their expert system was better than many other algorithms at cooperating, the performance of that algorithm was greatly improved if it could communicate with each other (when both players adopt the same expert system). Of course, in most jurisdictions today, communications among competitors as a way to reach an agreement would be illegal.

Even more recently, two Facebook AI researchers developed algorithms that can cooperate with opponents in similar social dilemmas. One of their algorithms was, in fact, inspired by the TFT algorithm. Specifically, the researchers tried to relax the two strong information requirements of the naïve TFT algorithm. Another recent study adopted an interesting approach to design an algorithm that promotes cooperation. Their idea is to introduce an additional planning agent that can distribute rewards or punishments to the algorithm players as a way to guide them to cooperation, analogous to an algorithmic hub and spoke agreement. Another group of researchers recently proposed an algorithm that explicitly takes into account the opponent’s learning through interactions and found that their algorithm worked well in eliciting cooperative behavior. New AI research on the topic of machine-machine and machine-human cooperation continues to appear.

With the experimental evidence that algorithms can indeed be designed to tacitly cooperate, the next question naturally becomes whether such an algorithm is available for use in the real world. The answer is that despite the promising theoretical and experimental results discussed above, we have a long way to go.

Several limitations are worth keeping in mind. First, almost all studies focus on only two players (the duopoly situation). It is well recognized that everything else being equal, as the number of players increases, collusion, tacit or explicit, becomes more difficult. Second, the type of games (e.g., repeated Prisoners’ Dilemma and its variants) and the universe of possible strategies in these experimental studies are rather limited, especially when compared to the real business world. Third, most of these experimental studies assume an unchanging market environment. For example, in most studies, the payoffs to the AI agents as well as the environment in which AI agents operate are typically fixed. This is a significant limitation because demand variability and uncertainty is not just a norm in the real world, but also has been long recognized by economists to have important implications on how cartels operate. For example, with imperfect monitoring, if the market price is falling, cartel firms may have a hard time figuring out whether the falling price is due to cheating or to declining demand (“a negative demand shock”). In fact, the economic literature shows that a rational cartel would need to internalize the disruptive nature of demand uncertainty when the cartel monitoring is imperfect.

Also relevant to the question of algorithmic collusion is whether the AI agents are symmetric, in other words, whether the opposing players have identical payoffs if they adopt the same strategies. In fact, almost all studies that use the repeated Prisoner’s Dilemma or its variants focus on the symmetric case. As I will discuss in the next section, the existence of various types of asymmetry (cost, market share, etc.) tends to make reaching a cartel agreement harder. Similar to the case of time-varying demand, economists have shown that a rational cartel may also need to explicitly take asymmetry into account and adapt its pricing arrangement accordingly. So there are good reasons to suspect that the AI algorithms designed under symmetry do not necessarily fare well in more realistic situations. Given all these real-world complications, it is not surprising that empirically, as of now, there is no known case of tacitly colluding robots in the real world.

Notwithstanding these limitations, the AI literature offers several insights that inform us how best to approach the antitrust risk of algorithmic collusion. The most significant is that designing algorithms with proven capability to tacitly collude in realistic situations is a challenging technical problem. One study took ten researchers from nine universities across four continents and several years to design. The researchers started with 25 algorithms and found that, in a variety of contexts, not all of them learned to cooperate effectively (without any communication), either with themselves or with other algorithmic players. In fact, the researchers identified the more successful algorithms only after extensive experiments. They highlighted a number of technical challenges. For example, they pointed out that a good algorithm must be flexible in that it needs to learn to cooperate with others without necessarily having prior knowledge of their behaviors. But to do that, the algorithm must be able to deter potentially exploitative behavior from others and, “when beneficial, determine how to elicit cooperation from a (potentially distrustful) opponent who might be disinclined to cooperate.” In addition, the speed of learning is important. A “collusive” algorithm is arguably irrelevant to the antitrust community if it takes an unrealistically long time to learn to collude. The researchers went on to state that these challenges often cause AI algorithms to defect rather than to cooperate “even when doing so would be beneficial to the algorithm’s long-term payoffs.”

But there are even more hurdles, some of which are related to the limitations stated above. For example, some point out that developing algorithms that are able to adapt to an ever-changing environment (say, changes in consumer preferences, hence demand) and at the same time understand the complexity of real-world situations (to compute the best strategy) presents significant computational challenges. What they point out is, in fact, another significant limitation of the current AI research; that is, most of the experimental studies assume the environment is static and known to the AI agents.

What all these complications imply is that there is a lack of support for the popular belief that just any learning (“trial-
and-error”) algorithm that simply tries to maximize profit would necessarily and eventually lead to tacit collusion. This also tells us that to design an algorithm that has some degree of guaranteed success in eliciting tacit collusion, the capability to collude most likely needs to be an explicit design feature. But what are the chances that a collusive algorithmic feature is also procompetitive (that is, efficiency enhancing)? A recent article argues that the chance of this is small.

Several lessons can be drawn from these observations. First, a potentially effective antitrust policy is to explicitly prohibit the development and incorporation of these problematic capabilities in a pricing or other strategic algorithm while balancing the pro- and anticompetitive effects of algorithms. Second, there may very well be important leads that the antitrust agencies and even private litigants could look for in an investigation or a discovery process. Several types of documents are of particular interest. These include any internal document that sheds light on the design goals of the algorithm, any documented behavior of the algorithm, and any document that suggests that the developers had revised and modified their algorithm to further the goal of tacit coordination. Another type of document that should raise red flags is any marketing and promotional material that suggests that the developers may have promoted the algorithm’s capability to elicit coordination from competitors. Note that it is not necessary for the investigators to have an intimate understanding of the technical aspects of the AI algorithm to look for such evidence.

Antitrust authorities as well as researchers can also take a more proactive approach. Specifically, research can be conducted to understand “collusive” features of an algorithm. Ezrachi and Stucke called this type of research a “collusive incubator.” Joseph Harrington went a step farther and proposed a more detailed research program and discussed its promises and challenges. Specifically, he proposed creating simulated market settings to test and identify algorithmic properties in learning algorithms that tend to lead to supra-competitive prices. He also argued that any scheme that involves a type of “reward-punishment” in the spirit of the example mentioned above is a problematic feature.

Another important observation from the AI research is that the algorithms being designed are not necessarily what economists call “equilibrium” strategies. Equilibrium strategies are intuitively “stable” algorithms in the sense that, if you know that you and your competitors adopt this strategy, none of you would have the incentive to change to another strategy. That is clearly not the case for some of the algorithms recently developed by AI researchers. In a recent study mentioned earlier, despite the promising experimental findings, the researchers acknowledge that unless their algorithm is an equilibrium learning strategy, it can be exploited by others, meaning that players may have an incentive to move away from their proposed algorithm. This observation has a powerful implication: unless firms are fully committed to a “collusive” algorithm that is not an equilibrium strategy, there will be a temptation for the (rational) firms to change their strategy and hence potentially disrupt the status quo or a potentially tacitly collusive outcome.

Another question that the AI literature helps address is whether a collusive outcome is easier to achieve if competitors adopt the same algorithm. As I elaborate in the next section, it is quite likely that even if firms technically adopt the same algorithm, they would customize their version of it. In other words, absolute algorithmic symmetry is unlikely in many cases.

Putting the practical reality aside, from a technical standpoint, it is, perhaps not surprisingly, easier to design a (tacitly) collusive algorithm so that a (tacit) collusive outcome can be achieved when opposing players both adopt the same algorithm. Indeed, to design an algorithm that is capable of eliciting (tacit) collusion from an unknown opposing algorithm is a much bigger technical challenge. Therefore, if the algorithm adopted and committed to by the competitors is designed to tacitly collude, then the answer to the question posed above—whether a collusive outcome is easier to achieve if competitors adopt the same algorithm—is likely in the affirmative. In fact, studies that design such algorithms typically demonstrate that the proposed algorithm works well when both players adopt the same algorithm. These observations preliminarily suggest that in absence of further evidence, presumably at the “screening” stage, a situation where firms all adopt the same pricing algorithm deserves more attention from an antitrust agency. But is tacit or explicit collusion so likely as to justify an investigation? That is a harder question that arguably has no simple answer and requires a case-by-case evaluation.

The Economics Literature

The economics literature that explicitly examines algorithmic collusion is limited. Using computer simulations, one early study showed that a particular type of “trial and error” learning approach called Q-learning could lead to some degree of imperfect tacit collusion in a quantity-setting environment. More recently, another study reported a similar finding in a simple price-setting environment. Both are important contributions and demonstrate the theoretical possibility of algorithmic collusion, when collusion is not an explicit design goal.

It is, however, important to recognize that, as the researchers acknowledge, the strong simplifying assumptions mean that their findings are largely “suggestive.” For example, both studies allow algorithms to interact and learn for extended periods of time. The authors of the first study considered one million interactions before examining the emergence of algorithmic collusion, while the second study found algorithmic collusion, to emerge after an average of 165,000 periods. To gain some perspective, 165,000 periods is equivalent to 13,750 years if firms change prices monthly or about 115 days if firms change prices every minute. In addition, the second study assumes that the firms can only choose two price
levels (low or high), and the profits are fixed and known given the choices the firms make. Similarly, among other simplifying assumptions, the analytical result in the first study is based on the assumption that firms can only choose between two production levels. In other words, the challenges and limitations discussed in the previous section remain.

A recent article provides a set of sufficient conditions under which the use of pricing algorithms leads to tacit collusion. The author considered an algorithmic version of tacit “invitation to collude.” Three conditions must be true for the algorithmic tacit collusion to materialize in his framework. First, competitors should be able to decode each other’s pricing algorithms. Second, after decoding others’ algorithms, the competitors should be able to revise their own pricing algorithms in response. Third, firms should not be able to revise or change their algorithms too fast. Intuitively, under these conditions, a firm could essentially tacitly communicate its intent to collude by adopting a “collusive” algorithm and letting the competitor decode it. Once this tacit invitation to collude is decoded, the competitor can then choose to follow the lead or not. When making the decision, the firm on the receiving end will naturally be concerned with the possibility that the invitation is no more than a trick and that once that firm starts to cooperate, the competitor would take advantage of it by immediately reversing course (say, by immediately lowering prices to steal customers away). This is where the third condition comes into the picture. If the firms understand that changing the strategy takes time, then the receiving firm’s concern would be alleviated.

One strand of economic literature that has received much attention in the antitrust community identifies the structural characteristics that tend to facilitate/disrupt collusion. A partial list of such structural characteristics that tend to facilitate collusion includes the following:

- Symmetric competitors.
- Fewer competitors.
- More homogeneous products.
- Higher barrier to entry.
- More market transparency.
- More stable demand.
- Small and frequent purchases by customers.

Market transparency is one obvious characteristic that an algorithm could potentially enhance. Consider an online marketplace where prices are posted for everyone, including the competitors, to see. While a human can certainly check on competitors’ prices periodically, a simple web-scraping algorithm can do the same much more efficiently and at a much higher frequency. Legal scholars have argued that this supercharged transparency could prevent cartel cheating because any deviation from a tacitly or explicitly agreed-upon price could be detected immediately.

At the same time, computer algorithms are unlikely to affect some of the other structural characteristics, especially those related to demand. For example, the use of a pricing algorithm is not going to make consumer purchases smaller or more frequent.

The effect of an algorithm on some other factors is ambiguous at best. One example is asymmetry. In general, economists believe that various forms of asymmetry among competitors tend to make collusion more difficult. A leading example is one where competitors have different cost structures (i.e., cost asymmetry). In this case, firms may find it difficult to agree to a common pricing because a lower-cost firm has an incentive to set a lower price than a higher-cost firm. This tends to make the coordination problem harder.

In addition, as a research paper put it, “[E]ven if firms agree on a given collusive price, low-cost firms will be more difficult to discipline, both because they might gain more from undercutting their rivals and because they have less to fear from a possible retaliation from high-cost firms.”

I noted above that adopting the same “collusive” algorithm likely makes tacit collusion easier. In reality, firms are likely to customize their algorithms even if they technically use the same algorithms. As an example, imagine some developers are telling us that their algorithm is going to increase our profit. But what profit? Certainly, an algorithm that aims to maximize short-term profit is not going to behave the same way as an algorithm that aims to maximize long-term profit. Similarly, we also expect an algorithm to incorporate firm-specific cost information in its decision-making process. The point is that even if the algorithms adopted by competitors have the same core structure and capability, they do not necessarily or automatically eliminate asymmetry. In fact, the algorithms are typically expected to reflect existing asymmetries.

Of course, at the end of the day, it is crucial to recognize the limitations of these structural factors because they only predict which markets are more susceptible to coordination, not whether market participants are explicitly or tacitly colluding. Nevertheless, if the economic theory has any implications for antitrust authorities, it seems prudent for competition agencies to pay a close attention to markets where market information, such as price, quantity, and market share, is readily available but a computer program could collect and process that information much more efficiently than humans. Interestingly, there is some empirical evidence that increased transparency has indeed led to potential tacit collusion in real markets.

One other strand of theoretical economic literature is relevant but has been largely unnoticeable in the antitrust community. A seminal article in this literature shows that under some conditions, if market information arrives continuously and firms can react to it quickly (for instance, with flexible production technologies), the collusion becomes very difficult. Most results in this literature are derived using advanced mathematical models and are not easily accessible to a broader audience.

Recall that earlier I discussed a situation where consumer demand is volatile and a cartel, producing a homogeneous
product, can only observe the market price but not the production of individual cartel members. In that situation, when the firms observe a lower market price, they cannot perfectly tell whether it is due to someone deviating from their agreement or simply due to weak aggregate demand. Firms can deter cheating by resorting to price wars (by producing more, for example) when the price falls below a certain level. In this framework, when the time firms take to adjust their production becomes shorter, there are two countering effects on the sustainability of a cartel. On the one hand, the ability to change their production quickly means that they could start a price war as quickly as they want to. This tends to reduce the incentive to cheat and hence make a cartel more sustainable. On the other hand, when the demand is noisy and hence the market price moves due to short-term idiosyncratic factors, firms that are constantly watching the market price trying to detect potential cheating will likely receive many idiosyncratic signals of (seemingly) lower prices. Under additional assumptions, the paper shows that firms will simply commit too many type I errors (false positives), that is, start price wars too often to sustain collusion in this environment.

Some studies have also shown that one way a cartel could combat the issue is to deliberately delay the information flow. At a theoretical level, it is not hard to imagine that using algorithms to (more or less) continuously monitor the market and enabling firms to react quickly could bring a real-world situation closer to the one considered in these research studies. More broadly, the idea that technologies could potentially disrupt or destabilize cartels deserves closer study.

**Other Considerations Related to Algorithmic Collusion**

One question relates to the scope of potential algorithmic collusion. Almost all the current discussions in the legal field are effectively limited to a typical price-fixing cartel, where the price (or quantity) is the only instrument. Of course, cartels and cartel agreements come in different shapes and forms. Some cartels allocate markets, while others rig bids. Some cartels use list pricing, while others mainly rely on sales representatives. Depending on the nature of the cartel, the ways to implement an agreement, tacit or explicit, are also going to be different. And these differences have some important implications for how we think about algorithms.

For the sake of argument, let’s assume for the moment that an algorithm capable of autonomous collusion is commercially available. Now imagine a situation where an algorithm—in its attempt to (tacitly or explicitly) collude—instructs a firm to restrict output when the demand is rising, sacrificing short-term profit. This “anomaly” is something that both firms themselves could see in real time and an antitrust agency could see ex post, even without any deep understanding of the underlying algorithm itself. Alternatively, imagine that a collusive algorithm has firms maintaining the price level instead of stealing competitors’ shares, as they used to do. In an industry that relies on sales representatives, to implement this plan, firms would need to change the incentives of the representatives from “volume before price” to “price before volume.” This incentive change has been flagged as a super plus factor, that is, one that strongly indicates the presence of a cartel. Again, this change in behavior is observable to the firms in real time and discoverable by antitrust agencies ex post, without an understanding of exactly how the algorithm works. This is related to a concept I have called “outcome visibility” of a collusive algorithm.

In presence of such super plus factors, firms may have a hard time laying the blame on a computer algorithm and claim to be merely an unaware innocent bystander. As Margrethe Vestager, European Commissioner for Competition, put it in a recent speech, companies “can’t escape responsibility for collusion by hiding behind a computer program.” Strictly speaking, given that conscious parallelism or tacit collusion is not illegal in many of today’s antitrust regimes, it will be interesting to see how the agencies will approach the situation if firms provide convincing evidence that their due diligence prior to adopting the algorithm yields no evidence of the possibility of tacit collusion. On that topic, I expect there to be many discussions in the antitrust community in the coming years.

**Conclusion**

Algorithms are becoming ubiquitous in our society. They are powerful and, in some cases, indispensable tools in today’s economy. In terms of technology, we do not yet have AI sophisticated enough to, with reasonably degree of certainty, reach autonomous tacit collusion in most markets. This does not mean that we should ignore the potential risks. In fact, existing AI research has made some promising progress toward designing AI to achieve coordination among competitors. It also offers us a fundamental insight: to design a collusive algorithm, certain “collusive” design features most likely need to be explicitly incorporated into the algorithm. So just like e-mail leaves a trail of evidence when cartel uses it to coordinate, a similar trail of evidence is likely present when collusive algorithms are being designed. At the same time, future research could proactively examine algorithmic characteristics that lead to tacit collusion, as other scholars have advocated. For example, as discussed above, any algorithm that has a “reward-punishment” component has been flagged in the literature as objectionable. Of course, it is also important that we evaluate both the pro- and anticompetitive effects of the algorithms and adopt a prudent view toward algorithms so we do not exert unnecessary chilling effects on technological innovation.

Companies considering adopting strategic pricing algorithms should, at a minimum, ask: (1) how the algorithms increase their profitability and make sure that promised profitability is not achieved through intended tacit collusion, and (2) what information goes into the algorithms and make
sure that the algorithms do not provide a back door for competitors to share sensitive information autonomously. Similarly, companies should be cautious not to agree with competitors to adopt a similar pricing algorithm.

With constant advances in technology and ever-increasing computation power, new ideas will undoubtedly continue to appear, update, and even revolutionize our understanding. This is why we in the antitrust community want to keep a close eye on AI developments, so we can be proactive and prepared to address the challenges ahead. Indeed, several antitrust agencies in the world have already started taking a closer look into algorithms and algorithmic collusion. For example, UK’s Competition and Markets Authority has recently established a data unit; the French and German antitrust authorities recently launched a joint project on algorithms and their implications on competition.

I am optimistic that with additional research by and collaboration among antitrust agencies, economists, and computer scientists, we will further our understanding of the economic underpinnings of algorithmic collusion and be positioned to tackle the associated risks and challenges.

---

1 CPI Talks . . . Interview with Antonio Gomes of the OECD, ANTITRUST CHRON. (May 2017).
3 Obviously, there are many “ifs” in this stylized example as well as unanswered questions. For example, for the tacit collusion to be sustainable, the question whether the firms would at any point in time have incentive to deviate is a critical question.
5 Michal S. Gal, Algorithmic-Facilitated Coordination: Market and Legal Solutions, CPI ANTITRUST CHRON. 27 (May 2017).
8 Many discussions in the literature effectively take the proposition described above as given and examine the challenges and propose strategies to deal with them. For example, Gal discusses challenges for enforcers “when the algorithm employs machine learning based on neural networks, that is, it realizes itself the best way to behave in the market even if the coder did not model such conduct.” Gal, supra note 5. Similarly, Ballard and Naik state, “By simply allowing these bots to go to work, it is easy to imagine an effectively permanent pricing stasis settling over many markets, and not always with procompetitive effects.” Ballard & Naik, supra note 6.  
12 While we can broadly group many research studies under the umbrella of AI, the relevant literature is cross-disciplinary and involves game theory, experimental science, machine learning, and operational research.
13 To learn more about the basics of machine learning approach to AI and its implications on antitrust in general, see Ai Deng, An Antitrust Lawyer’s Guide to Machine Learning, ANTITRUST, Spring 2018, at 82.
14 In this case, David Topkins, a former e-commerce executive, was charged with price fixing in the DOJ’s first online market prosecution. Topkins pled guilty for conspiring to fix the prices of merchandise sold online. Press Release, U.S. Dep’t of Justice, Former E-Commerce Executive Charged with Price Fixing in the Antitrust Division’s First Online Marketing Prosecution (Apr. 6, 2015), https://www.justice.gov/opa/pr/former-e-commerce-executive-charged-price-fixing-antitrust-divisions-first-online-marketplace.
16 Furthermore, unless the products are completely homogeneous and firms have identical costs, firms may not find copying competitors’ pricing from the last period desirable. Equally important is that theoretically, it is known that TFT is not a robust strategy. There is much discussion on the weaknesses of TFT in the literature. For example, a single mistake in either party’s action could lead to a “death spiral.” That is, when one party defects while the opponent cooperates in just one period, the parties will end up alternating between cooperation and defection, yielding worse pay-off for both than if they had cooperated.
20 As Joseph Harrington has noted, “[A]ctual markets are far more complicated than the stark simplicity of the Prisoners’ Dilemma. Actual markets have many possible prices to be selected for multiple products, and firms that are subject to changes in cost and demand.” Harrington, supra note 18.
21 Note that some studies assume that the AI agents are not aware of the pay-offs but rather have to learn about them in the process. So the assumption of fixed pay-offs is distinct from the assumption regarding the information set of the AI agents.
22 For example, Green and Porter show that under such conditions, one way for the cartel to sustain its agreement is to agree to start a price war if the market price falls below a certain level (known as a “trigger price”), as a way to “punish” potential cheaters and reduce the incentive to cheat. Doing so is in some sense inefficient because this may start a price war simply because the demand is weak (hence lower prices) but not because of cheating. More generally, economists have argued that demand volatility tends to hinder collusion. Edward J. Green & Robert H. Porter, Noncooperative Collusion Under Imperfect Price Information 52 ECONOMETRICA 87 (1984); Robert H. Porter, A Study of Cartel Stability: The Joint Executive Committee, 1880–1886, 14 BELL J. ECON. 301 (1983). In Rotemberg and Saloner’s model, a positive demand shock (an economic boom) could disrupt collusion by increasing firms’ incentive to deviate from their agreement because they could profit more from the high demand by deviating (say, lower prices). Julio J. Rotemberg & Garth Saloner, A Supergame-Theoretic Model of Business Cycles and Price Wars During Booms, 76 AM. ECON. REV. 390 (1986); Glenn Ellison, Theories of Cartel Stability and the Joint Executive Committee, 25 RAND J. ECON. 37 (1994).
23 Consider some recent research: Lerer & Peysakhovich, Maintaining Cooperation, supra note 18; Peysakhovich & Lerer, Consequentialist Conditional Cooperation, supra note 18. Both studies aim to design algorithms to better coordinate in repeated Prisoner’s Dilemma-type situations. Their AI
agents learn using a technique known in the reinforcement learning literature as “self-play.” That is, their AI agent would essentially interact with itself to learn about its opponent’s objectives and the implications of its behavior on the opponent. If the agents are asymmetric, however, then the standard self-play would not be as informative. Incidentally, self-play is also a technique used by the well-publicized Go-playing algorithm AlphaGoZero developed by Google DeepMind.


Crandall et al., supra note 17.


Of course, this is not to say that it is theoretically impossible to have a pricing algorithm that just happens to be capable of collusion. Some preliminary evidence can be found in Waltman and Kaynak, Ludo Waltman & Üzay Kaynak, Q-Learning Agents in a Cournot Oligopoly Model, 32 J. ECON. DYNAMICS CONTROL 3275 (2008).

Anticipating an imperfect delineation, Harrington also proposed a framework to assess the antitrust liability of a particular algorithm to the extent that there is some uncertainty. Interested readers are referred to Harrington for a detailed discussion. Harrington, supra note 18.

This is important because there is a natural and understandable “fear” of complex AI/ML methods especially methods such as deep neural network. For example, Stucke and Ezrachi emphasize, “We note how, to date, most strategies discussed are powered by price algorithms and are yet to include cutting-edge neural networks. The increased use of neural networks will indeed complicate enforcement efforts.” Ariel Ezrachi & Maurice E. Stucke, Algorithmic Collusion: Problems and Countermeasures 17 (Background Note for 127th Meeting of OECD Competition Committee, June 21–23, 2017), https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?docid=DAF/COMP/WDA(2017)29&Language=En. They explain, “Due to their complex nature and ever-evolving abilities when trained with additional data, auditing these networks may prove futile. The knowledge acquired by a Deep Learning network is diffused across its large number of neurons and their interconnections, analogous to how memory is encoded in the human brain. These networks, based on non-linear transformations, are considered as opaque, black boxes. Enforcers may lack the ability to trace back the steps taken by algorithms and unravel the self-learning processes. If deciphering the decision making of a deep learning network proves difficult, then identifying an anticompetitive purpose may be impossible.” Id. at 25.

Harrington, supra note 18.

This situation is known as the “Nash equilibrium” in game theory. There are also many “refinements” to Nash equilibria, some of which are designed to be even more robust and stable.

Peysakhovich and Lerer explicitly distinguish this difference: “[T]he question of designing a good agent for social dilemmas can sometimes be quite different from questions about computing equilibrium strategies. For example, in the repeated PD [Prisoner’s Dilemma], tit-for-tat is held up as a good strategy for an agent to commit to (Axelrod, 2006). However, both players using tit-for-tat is not an equilibrium (since the best response to tit-for-tat is always cooperate). In fact, it is clear that one of their designs, the so-called antiTFT, is not an equilibrium strategy based on their Figure 5.” Lerer & Peysakhovich, Maintaining Cooperation, supra note 18.

Foerster et al., supra note 19.

It is important to note that the literature on algorithmic collusion is distinct from the large experimental literature in economics that studies human behavior in oligopolistic markets. For example in the latter literature, see Steffen Huck, Hans-Theo Normann & Jörg Dechssler, Two Are Few and Four Are Many: Number Effects in Experimental Oligopolies, 53 J. ECON. BEHAV. ORG. 435 (2004).

Ittoo & Petit, supra note 26.


For a discussion on the implausibility of “accidental tacit collusion” or “blundering into tacit collusion,” see, e.g., Edward J. Green, Robert C. Marshall & Leslie M. Marx, Tacit Collusion in Oligopoly, in THE OXFORD HANDBOOK OF INTERNATIONAL ANTITRUST ECONOMICS 2 (Roger D. Blair & D. Daniel Sokol, eds., 2014).

Interestingly, the authors also remarked, “However, Q-learning also predicts a substantial degree of collusion in Cournot games with more than two firms. This does not match experimental results. In experimental studies, firm behavior usually turns out to be quite close to the Nash equilibrium when the number of firms is larger than two.” Ittoo & Petit, supra note 26. See also Kimbrough and Murphy for another study of algorithmic collusion, Steven O. Kimbrough & Frederic H. Murphy, Learning to Collude Tactically on Production Levels by Oligopolistic Agents, 33 COMPUTATIONAL ECON. 47 (2009). Note that the limited experimental evidence reported in these papers is consistent with Wang’s more extensive experimental findings that “two Q-learning agents can learn to cooperate with each other with some low [emphasis added] probability.” Keven Wang, Iterated Prisoners Dilemma with Reinforcement Learning (Stan. U., Working Paper, 2017).


And all these are also common knowledge among the competitors.

For all technical details, see Bruno Salcedo, supra note 40. Note that the conditions described are sufficient but not necessary conditions, meaning that other conditions may also lead to tacit collusion.


See David Rahman, Information Delay in Games with Frequent Actions (Working Paper, 2013) (paper on file with author) (“I study repeated games with frequent actions and obtain a Folk Theorem in strongly symmetric strategies for the Prisoners’ Dilemma if public information can be delayed” (emphasis added)).

It is worth noting that some have argued that the result that a cartel is hard to sustain is largely driven by certain assumptions made previously. See, e.g., António M. Osório, A Folk Theorem for Games When Frequent Monitoring Decreases Noise, 12 B. E. J. THEORETICAL ECON. (2012).

As another example, even given today’s technology, one could change the content of a web page according to the origin of the web visit. It is not inconceivable that an online seller could “hide” its “deviation” from an explicit or tacit cartel by adopting such a technology.

I focus on only a couple of issues here. For additional considerations, including the role that class action plays, see Ai Deng, 4 Reasons Why We May Not See Colluding Robots Anytime Soon, Law360 (Oct. 3, 2017).


Note that although prices are always observable or at least discoverable, it may be less straightforward to discern “anomalies” because prices are driven by many factors.

Deng, supra note 51.


For a recent proposal regarding how to redefine antitrust liability in the context of algorithmic tacit collusion, I refer readers to Harrington, supra note 18.

Id.