From the Dark Side to the Bright Side: Exploring Algorithmic Antitrust Compliance

Introduction

If pricing algorithms could autonomously collude, can they be made automatically antitrust compliant as well? That is the question many have started pondering after a series of public comments by EU competition officials in recent years. Explaining this concept, the EU Competition Commissioner Margrethe Vestager stated in a recent speech that “(w)hat businesses can—and must—do is to ensure antitrust compliance by design. That means pricing algorithms need to be built in a way that doesn’t allow them to collude.” She later elaborated on her view at another conference: “Some of these algorithms will have to go to law school before they are let out. You have to teach your algorithm what it can do and what it cannot do, because otherwise there is the risk that the algorithm will learn the tricks…. We don’t want the algorithms to learn the tricks of the old cartelists…. We want them to play by the book also when they start playing by themselves.”

Another senior EU official echoed the view that firms should program “software to avoid collusion in the first place” and that “[R]espect for the rules must be part of the algorithm that a company configures and for whose behavior the company will be ultimately liable.” The concept of compliance by design appears intuitive enough that Gosselin, Jones, and Martin wrote that “[t]he software is always a product of its programmers—who of course have the ability to (affirmatively) program compliance with the Sherman Act.”

As desirable as antitrust compliance by design is, as Professor Simonetta Vezzoso put it in her 2017 article, “[w]hile the idea of competition compliance by design might be gaining some foothold in the mind-sets of some competition authorities, there are currently no clear indications how it could be integrated into the already complex competition policy fabric.” Indeed, what does it mean to “program compliance with the Sherman Act?” That is the question that Professor Joseph Harrington asked in a recent paper. He concluded...
that all that the current jurisprudence tells us is to make sure algorithms do not communicate with each other in the same sense that human managers are prohibited from communicating under the Sherman Act. But as both Professors Vezzoso and Harrington suggested, there is more we could do.

In this paper, I discuss several potential pathways to algorithmic compliance and argue that a robust compliance program should take this holistic and multi-faceted view. Specifically, I will look at a monitoring approach to compliance, then venture into the harder problem of designing compliant algorithms from the ground up. I will also discuss some existing proposals, draw additional lessons from recent AI literature, and finally present potential technical frameworks, inspired by the current machine learning literature, for compliant algorithmic design. I will not take a position on what should be illegal under antitrust law. For this, I refer interested readers to the careful analysis by Harrington. Instead, my discussion will be focused on preventing algorithms from engaging in coordination, explicit or tacit, that leads to supra-competitive pricing or other types of collusive behavior.

Algorithmic Compliance: A Monitoring Approach

The first approach we discuss is the use of automated monitoring as a compliance tool. Despite not being the type of competition by design that would immediately come to mind, these types of algorithmic tools can and should be an important component of a compliance program. Instead of trying to dictate the design process, these tools monitor the behavior of humans, as well as the algorithms. The main advantage of this approach is that it does not attempt to open the black box of complicated algorithms; it focuses instead on the relevant firm behaviors that can be observed and interpreted.

The first approach is to directly monitor the "symptoms" of an antitrust violation. Some of these "symptoms," which allow one to draw a credible inference of (explicitly) collusive conduct, are referred to as "plus factors." More formally, Kovacic, et al., define plus factors as "economic actions and outcomes, above and beyond parallel conduct by oligopolistic firms, that are largely inconsistent with unilateral conduct but largely consistent with explicitly coordinated action." They further define the super plus factors as the strongest of such factors. For example, unexplainable price increases or other types of abnormality in prices have been recognized as such plus factors. There is by now a robust "cartel screen" literature that studies empirical approaches and designs algorithms to detect such price anomalies. And with adequate data, these empirical screening algorithms could be an important addition to an algorithmic compliance program.

What appears to be somewhat underappreciated in the literature of cartel screening is perhaps how plus factors beyond prices can also be helpful when it comes to antitrust screening. For example, it has been noted that output restriction when demand is strong and prices and profits are high is a (super) plus factor. It is not hard to imagine designing an algorithm that keeps an eye on output level in relation to firm’s capacity, prices, and profits, all of which are common performance metrics firms look at during their normal course of business. Similarly, when prices are relatively high or rising and firms have excess capacity, stable market shares and the lack of customer churn also indicate collusive conduct by competitors. An algorithm could be designed to keep an eye on such market information.
As a final example, imagine that a collusive algorithm has firms maintaining price level instead of stealing competitors’ shares, as they used to. In an industry that relies on sales representatives, to implement this plan, firms would need to change the representatives’ incentives from “volume before price” to “price before volume.” This incentive change has also been flagged as a super plus factor and can be efficiently monitored by an algorithm.

The second approach to algorithmic compliance has seen increasing adoption and success in the RegTech (Regulatory Technology) industry where AI technologies are being deployed to help companies meet their regulatory compliance needs. RegTech is often touted as the new FinTech (Financial Technology) and has seen rapid growth in recent years. Several RegTech companies currently offer AI technologies based on natural language processing (NLP) and natural language understanding (NLU), which capture and understand voice and text communications. One company offers a technology that can “detect intent, emotion & entities” in communications. This type of technology could also be used to monitor communications among competitors to detect and ultimately prevent collusive behavior. Indeed, there has been a wealth of evidence of competitor communications in a number of high-profile international cartel cases, and such communications played a critical role in the investigation. With AI technology from the NLP and NLU fields, machines can flag problematic communications in real time in a cost-effective manner.

Robust algorithmic monitoring allows firms to respond to regulatory agencies’ compliance inquiry and to meet their compliance requirements efficiently. While such monitoring does not directly tell us how to design a compliant algorithm, it reduces the likelihood of non-compliance by giving the right incentives to the decision makers. Next, we turn to the challenging question of how to make sure the algorithms are antitrust-compliant. As I will argue below, monitoring algorithms can play a role in that effort.

Algorithmic Compliance: Compliance by Design

Looking Back: Lessons from the Literature
If we could send pricing algorithms, or more realistically the algorithm developers, to law schools, what would we teach them in Antitrust Compliance 101, other than simply “do not design a collusive algorithm”? The existing literature provides some ideas.

Our first lesson is inspired by a recent AI research study in which a team of researchers designed an expert system (a type of artificial intelligence) that enables better coordination among opposing players in a variety of situations. The study found that although their expert system was better at cooperating than many other algorithms, the performance of the algorithm is significantly improved when it can communicate with other algorithms through costless, non-binding messages (“cheap talk”). This research shows that, just like humans, the ability to signal can be a key to forging a cooperative relationship among competing algorithms. The lesson here is plain and simple: your algorithms should never open a secret backdoor channel for communications between competitors, even if the communications are written in machine-readable syntax.
The second lesson is also straightforward. You should never design a pricing algorithm with the explicit goal of eliciting or facilitating coordination or collusion. Interestingly, there is both a technical and an economic/behavioral component here. From a technical perspective, as I have noted elsewhere, even though we have a long way to go until there are commercially available autonomously colluding robots, researchers have been able to successfully devise algorithms that elicit tacit coordination among competitors in laboratory environments.\(^\text{18}\) For example, two Facebook AI researchers recently developed algorithms that can cooperate with opponents in social dilemmas, including the repeated prisoners’ dilemma.\(^\text{19}\) Another recent study adopted an interesting approach to design an algorithm that promotes cooperation, introducing an additional planning agent that can distribute rewards or punishments to the algorithmic players as a way to guide them toward cooperation, analogous to an algorithmic hub-and-spoke agreement.\(^\text{20}\) Yet another group of researchers recently proposed an algorithm that explicitly accounts for the opponents’ learning through interactions; they found that their algorithm worked well in eliciting cooperative behavior.\(^\text{21}\)

From a behavior perspective, we note that a recent research paper provided a set of sufficient behavioral conditions under which the use of pricing algorithms leads to tacit collusion. The conditions the article identifies are essentially those underlying an algorithmic version of tacit “invitation to collude.” Specifically, three conditions must be true for algorithmic tacit collusion to materialize in this 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 quickly. 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 about 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 (for instance, by immediately lowering prices to steal customers). 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. Note that the result does not require the algorithms themselves be designed in a sophisticated way. As such, the most immediate lesson is that, from an algorithm design perspective, the developers should not include a component that enables switches between collusive and non-collusive pricing conditional on what the competitor does in response to a signal (e.g., an invitation). An obvious corollary is that one should never intentionally expose these pricing algorithms to the competitors.\(^\text{22}\)

While intuitively straightforward, this corollary is worth more discussion. First, modern algorithms are often capable of setting prices for thousands of products. While one may argue that coordination on so many prices tends to make collusion difficult, the tremendous amount of price variations as a result of the algorithms price-setting could potentially allow competitors to “decode” the algorithms. For example, researchers have attempted to decode Uber’s black-box surge pricing algorithm with some success.\(^\text{23}\) What this means is that we may need to accept the reality that competitors have some visibility into each other’s algorithms.
The compliance protocols above should eliminate or at least make it easier to detect many cases of algorithmic tacit collusion. A much more challenging task is identifying specific algorithms or algorithmic features that should or should not be built in, a question that many may have in mind when thinking about compliance by design. To answer this question, more research is needed. Professor Harrington proposed a research program and discussed its promises and challenges. Specifically, he proposed the use of simulated market settings, or “collusive incubators,” as Professors Ariel Ezrachi and Maurice Stucke called them, to test and identify algorithmic properties that tend to lead to supracompetitive prices. Once the problematic algorithms or algorithmic features are identified, they become part of a prohibited set. As an example, Professor Harrington underscored the importance of the reward-punishment scheme in eliciting and sustaining collusion. In fact, a basic teaching of the repeated strategic game (e.g., the repeated prisoners’ dilemma) is that firms could sustain collusion and resist the temptation to cheat if the long-term gain (from colluding) offsets and even exceeds the short-term loss (from cheating). To do this, there must be rewards for colluding and punishments for cheating.

If this research program sounds ambitious, here is some good news: We are not starting from scratch. While the economics literature that explicitly examines algorithmic collusion is limited, there has been work that adopts this framework. One early study showed that a type of reinforcement learning appears to lead to some degree of imperfect tacit collusion in a quantity-setting environment. Two recent studies reported similar and even stronger evidence of tacit collusion by reinforcement learning agents in a price-setting environment. As discussed above, in the AI field of multi-agent learning, researchers study precisely how AI agents could learn to cooperate with each other (or with humans) when their individual interests are in conflict. Regardless of whether algorithmic tacit collusion is illegal under antitrust law, it is advisable for companies considering adopting pricing algorithms to, at a minimum, know whether their algorithms are based on those known to elicit tacit collusion. One may take a step further and argue that it is a good compliance practice to test whether their pricing algorithms lead to tacit collusion in experimental environments.

In the next section, I discuss potential ways to implement compliance by design when we do not necessarily have knowledge about the problematic features beforehand or when it is difficult to isolate the properties that lead to supra-competitive pricing.

Looking Forward: A Research Proposal

Professor Vezzoso highlighted the significant challenges in programming antitrust law directly into algorithms. She noted that “programmers are capable of developing and implementing effective, concrete solutions for complying with norms that are specific and detailed enough. Programmers must articulate their objectives as ‘a list of instructions, written in one of a number of artificial languages intelligible to a computer’…. The flexibility of human interpretations, meaning the possibility that legal practitioners interpret norms and principles differently and that legal interpretation evolves over time, may conflict with the apparent stiffness of computer language…. The degree to which competition law is, or should be, suitable for automation is an interesting yet neglected topic.” Indeed, given how concise the US Sherman Act is, most legal scholars, if not all, would agree that turning the Sherman Act into a set of specific if-then type instructions is a tall order, if not outright impossible.
But there is one particularly interesting thing about how programmable our antitrust laws are: The lack of programmability is exactly the problem that modern machine learning and AI technologies are designed to circumvent. Consider the task of automatically recognizing and distinguishing cats and dogs in images. The traditional computer programming approach may be to enumerate all the physical differences between cats and dogs. This is known as “rule-based” programming. But given how many subtle physical differences and similarities there are between cats and dogs, it gets difficult very quickly to improve classification accuracy. The standard machine learning approach circumvents this problem by providing a large number of examples that consist of inputs (images) and associated outputs (the “label” describing whether the image is a cat or a dog) to a statistical model. A large number of such examples, also known as training data, allows the model to search for the most predictive inputs, as well as the best way to map these inputs to the correct output, all without relying on rules that humans must painstakingly write down.\(^{31,32}\) This tells us that, perhaps, we also do not need to write down all the explicit instructions of antitrust compliance. Fortunately, economists and courts have identified a set of indicators predictive of collusive conduct.\(^{33}\) These are the plus factors we have proposed to use in a monitoring approach to antitrust compliance. The question we address in this section is whether these powerful predictors and the algorithms designed to monitor them can contribute directly to the design of antitrust-compliant algorithms, and if so, how.

Before drawing inspirations from the existing AI literature, we should note that the first principle of designing an algorithm is to specify its objective function. The most likely objective of a pricing algorithm is to maximize profit. But firms face various constraints when they maximize their profits. They may have limited capital or limited production capacity. There are also regulatory and ethical constraints on a firm’s pursuit of profits. Conceptually, antitrust compliance can be thought of as a similar constraint. Therefore, the technical challenge of compliance by design can be seen as one of implementing compliance as a constraint in the training/learning process of the algorithm.

Once we cast the technical challenge in this framework, several strands of AI literature offer inspiration for possible paths forward, allowing one to directly incorporate compliance into the algorithm design. In reinforcement learning (RL), the so-called actor-critic approach relies on both an actor who tries to figure out a strategy that leads to the best outcome (e.g., highest profit) and a critic who examines the desirability (e.g., antitrust compliance) of an action dictated by the strategy, given the circumstances, and provides the feedback to the actor for adjustment.\(^{34}\) Generative adversarial networks (GANs), a conceptually similar idea, also draw strength from two algorithms. In this type of model, while one algorithm tries to generate some content (say, an image), the adversarial algorithm tries to identify it as a computer-generated fake.\(^{35}\) A compliant pricing algorithm could have a similar actor-critic/adversarial structure in which, as the actor tries to maximize firm profit, the critic could look at a compliance score of a pricing decision taken by the actor and provide feedback so the actor could learn to steer away from problematic actions. The compliance score can take a variety of values to “discipline” the pricing algorithms. For example, the compliance score could be negative (i.e., a penalty) if, given the actions taken, cartel behaviors, as evidenced by plus factors including collusive prices,
arise at the end of the training; and the score could be positive if such evidence does not arise. More sophisticated scoring method could consider the strength of the plus factors as shown in the literature. The compliance score may also be explicitly treated as a constraint in the pricing algorithm’s profit maximization problem, as in the statistical literature on regularization methods. What is different about these approaches is that the compliance component is an integral part of the algorithm design.

I emphasize that even though the inspiration come from the existing literature, the conjectured approaches are nontrivial deviations from it. There are numerous technical and practical problems to be resolved. For example, in a standard actor-critic approach, the objective of the two algorithms is typically the same, whereas in our conjectured application, the critic would adopt a very different objective. Another obvious challenge includes the proper ways of specifying the compliance score. Despite the aspirational nature of this discussion, the key message is that the existing literature offers plausible frameworks for designing compliant pricing algorithms.

**Explainable AI**

In the recent years, there has been a rapidly growing interest in explainable AI from both academia and the private sector. As the name suggests, explainable AI aims to make algorithmic decision-making understandable to humans. Notably, the Defense Advanced Research Projects Agency (DARPA) sponsors a program called XAI (Explainable Artificial Intelligence). The organization FATML (Fairness, Accountability, and Transparency in Machine Learning) also aims to promote the explainable AI effort. While we still have a long way to go in explainable AI research, the industry and academic interest is a promising sign.

Some of the commercial interest in explainable AI comes from the commercial lending industry because of the regulation and the need to explain lending decisions to consumers, especially when the decision is made by machine learning models. It should be no surprise that the same need for explainability goes well beyond the lending industry. For example, interpretability of algorithms can be equally important in the medical and health care domains. It should also be an important part of the research program for antitrust compliance by design. If Commissioner Vestager’s warning—that companies cannot hide behind an algorithm—is to be taken seriously, with explainable AIs firms would have no reason not to understand the pricing behavior of their algorithms.

The AI research community has proposed several ideas to help achieve interpretability and explainability of AI. Two common approaches, rooted in the technical aspects of AI, are (1) the use of inherently interpretable algorithms (known as “white-box algorithms”) and (2) the use of clever backward engineering. Naturally, there is a multitude of explainability, and different domains may find different definitions acceptable. In the context of algorithmic (tacit) collusion, the ability to explain and answer why or what-if questions is helpful when it comes to understanding the algorithmic pricing decision. Suppose your pricing algorithm is setting a price that you think might be too high. A helpful what-if question could be, “What if we lower the price?” or “Would we generate higher immediate profit by doing that?” The answer might
be, “Based on demand forecasts and our customers’ price elasticity, this is the optimal price we should set” or “We have no reason to lower our price because we know that the competitor’s algorithm is not going to lower theirs, and we know that because we have determined that this is the best course of action that benefits both of us in the long term,” or even “We should not lower our price because the last time we lowered our price, the competitor started a price war.” Whether or not the last two responses suggest problematic algorithmic conduct, having this knowledge can be extremely helpful.

But how would one go about tackling such algorithmic explainability? An AI study published in 2018 titled “Contrastive Explanations for Reinforcement Learning in terms of Expected Consequences” is an important step toward achieving the type of explainability discussed above. In the framework of standard reinforcement learning, the study’s researchers developed a method that enables an RL agent to precisely answer the what-if questions similar to those we posed above. Suppose we are curious about why the autonomous RL agent takes action A, instead of another action B. Their RL agent will answer this why question by contrasting the expected outcomes or consequences of the two actions. This type of contrast underlies our conjectured answers above.

It is important to keep in mind that multi-agent learning, the more relevant technology to algorithmic collusion, is outside the scope of this particular AI research study. It is safe to say, however, that it will be only a matter of time until we see progress in that area as well. Meanwhile, because RL algorithm learn through “trial and error” and are trained not to make the same mistakes, but rather to exploit the correct decisions, documenting the learning process of the (pricing) algorithms can be helpful. To see why, consider the following example. If the algorithm is able to figure out on its own to use reward and punishment to elicit and sustain tacit collusion, then the learning process should reflect that. Specifically, it will reflect how the payoffs changed as the algorithm adopted different actions and how the actions taken by the algorithm have changed as a result. A recent study nicely demonstrated this idea. In their experimental setting, the researchers demonstrated that when one firm deviated from the supra-competitive prices, the other RL pricing algorithms punished the deviation as the way to restore and sustain supra-competitive prices.
Summary

Professor Vezzoso nicely summarized what it takes to achieve compliance by design when she wrote that “competition compliance by design can become an effective tool in the enforcer’s kit only if it is based on open, constructive and ongoing dialogues and exchanges with all interested stakeholders, including the enforcers themselves, firms, computer science experts, designers and providers of algorithms, academia, and consumers. Otherwise, a serious risk is that ‘competition by design’ remains an enticing slogan or, even worse, an ex-ante prophylactic measure.”

As I have argued elsewhere, there are significant technical hurdles in explicitly designing a collusive algorithm in the first place, and there are practical reasons why we may not see colluding robots any time soon. But as more market players develop pricing and other strategic algorithms, it is time to start thinking about antitrust compliance as an integral part of automated decision-making. In this paper, I discussed several potential pathways to algorithmic antitrust compliance. There are many open questions. It is my hope that this paper stimulates more discussion, and even debates, on this important issue, one that is guaranteed to be increasingly relevant as we march forward in this fourth industrial revolution powered by constant AI advances.
Notes

1. Ai Deng, PhD, is an Associate Director at NERA Economic Consulting, an Editor at American Bar Association’s (ABA) Antitrust Source, and a lecturer at the Advanced Academic Program, Johns Hopkins University. His previous work on the topic of algorithmic collusion has appeared in the ABA’s Antitrust Magazine and featured on Competition Policy International. In November 2018, he testified live on the risk of algorithmic collusion at US FTC’s Hearings on Competition and Consumer Protection in the 21st Century. He has also spoken on the antitrust implications of technologies at CLE-credited events sponsored by the ABA and the Knowledge Group and at the Competition Bureau Canada. He can be reached at ai.deng@nera.com. The views expressed in this paper are those of the author and do not necessarily reflect the opinions of NERA or its clients, the ABA, Johns Hopkins University, or its affiliates.


3. Interview at the 2017 Web Summit, see video at https://youtu.be/90OhCfyYQOk.


9. For readers who wish to have a quick overview of algorithmic collusion, see, for example, Ai Deng, “What Do We Know About Algorithmic Tacit Collusion?” Antitrust, 33(1), 2018.


12. Ibid.


14. Deloitte curates a list of RegTech companies and, by the time of this writing, there are 263 companies on the list. Eighty percent of these companies were started in the past 10 years. Companies started in the five-year span between 2012 and 2016 account for 63% of all RegTech firms. The UK Financial Conduct Authority is also actively looking to RegTech. See, for example, Alison Noon, “UK Finance Cop ‘Aggressively’ Pursuing Robo Regulation,” Law360, 17 January 2019, available at https://www.law360.com/securities/articles/1119869/uk-finance-cop-aggressively-pursuing-robo-regulation?nl_pk=c40ee15e-1509-4120-8a09-e3fcecfaa0&utm_source=newsletter&utm_medium=email&utm_campaign=securities.


17. This is consistent with the point made by Professor Harrington. He writes that “[j]urisprudence tells us it means that AAs [Algorithmic Agent] cannot communicate with each other in the same sense that human managers are prohibited from communicating. Thus, AAs would not be in compliance if they coordinated their conduct using arbitrary messages unrelated to the competitive process, but would be if coordination was achieved through their prices, as that is an example of legal conscious parallelism.” Harrington, 2019, Ibid, p. 19.

18. For a discussion of some of these algorithms, see Deng, 2018, Ibid, and Deng, “4 Reasons Why We May Not See Colluding Robots Anytime Soon,” Law360, 3 October 2017.


22. One extreme case of complete disclosure is when competitors purchase and implement the same third-party pricing algorithm. A great deal of concern has been raised in such a situation. One mitigating factor is that, as I have argued elsewhere, firms are likely to customize their version of the same algorithm and the resulting asymmetry tends to make algorithmic tacit collusion difficult, everything else being equal. See Deng, 2018, Ibid. In addition, one question that a rational firm or a vendor considering adopting a pricing algorithm is expected to ask is how it would compete with its competitors if everyone is using the same algorithm.


24. For a discussion of these technical difficulties, see Deng, 2018, Ibid.

25. In essence, this experimental approach is quite similar to the gene-knockout experiment in genetic research. The problematic collusive features, to the extent that they can be identified, are analogous to the genes that make the plant prone to certain disease.
More broadly, Professor Harrington argued that algorithmic pricing decisions should not “be dependent on rival firms’ responding in a particular manner; for example, a price increase is not contingent on rival firms subsequently matching that price, or maintaining price is not contingent on rival firms conducting a price war if price were to be reduced.” Harrington, 2019, Ibid, p. 27.

In a recent article, four economists adopted this approach and suggested an interesting way to detect (tacit) price fixing by algorithms in a laboratory setting. Specifically, the researchers forced their algorithm to “cheat” in one period by lowering its own price. They found that the (algorithmic) competitor immediately lowers its price in the next period by an even larger amount, suggesting a “punishment.” Furthermore, it appears that the “cheating” algorithm anticipates the punishment because it also immediately cuts its price further. After that, the researchers find that both algorithms gradually raise their price to a pre-cheating supra-competitive level nearly in lock-step. See Emilio Calvano, Giacomo Calzolari, Vincenzo Denicolo and Sergio Pastorello, “Artificial Intelligence, Algorithmic Pricing and Collusion,” working paper, April 2019.


Of course, we should always keep in mind the pro-competitive effects of a given algorithm or algorithmic feature and understand that we may need to tackle the harder problem of trading off the pro- and anticompetitive effects.

In fact, one could argue that it is this “work-around” that creates the black box we are now struggling to open.

For an application of machine learning in the legal context, see, Daniel Martin Katz, Michael J. Bommarito II, and Josh Blackman, “A general approach for predicting the behavior of the Supreme Court of the United States,” PLOS One, 12(4), 2017.

I emphasize these are predictive but not causal factors.


See, for example, Marshall, et al., 2011, Ibid.

For readers with a technical background, the regularization methods are closely related to the concept of “shrinkage.” The basic idea behind this type of approach is to introduce constraints on the model parameters. For an example of how regularization may be used to balance different modeling objectives, see Mohammad Taha Bahadori, Krzysztof Chalupka, Edward Choi, Robert Chen, Walter F. Stewart, and Jimeng Sun, “Causal Regularization,” Cornell University, working paper, 2017, available at https://arxiv.org/abs/1702.02604.

In this sense, our conjectured application mimics the statistical concept of regularization.


See, for example, Patrick Hall and Navdeep Gil, An Introduction to Machine Learning Interpretability, 2018; and Joel Vaughan, Agus Sudjianto, Erind Brahim, Jie Chen, and Vijayan N. Nair, “Explainable Neural Networks based on Additive Index Models,” Cornell University, working paper, June 2018, available at https://arxiv.org/abs/1806.01933. It is important to note that explainable AI does not mean the algorithms must be self-explanatory. Additional tools, such as what’s known as “surrogate models,” will often be necessary to open the black box.


It is worth noting that several economic studies have investigated reinforcement learning for its potential to lead to tacitly collusive outcomes in simulated oligopolistic markets. See, for example, Waltman and Kaymak, Ibid, and Calvano, et al., Ibid. For an introduction to various machine learning algorithms, including RL, see Deng, 2018, Ibid.

As the researchers put it, their method “calculates contrasts between the consequences of the user’s query-derived policy, and of the learned policy of the agent.” In addition, the researchers also developed a procedure that translates the environments and the actions to a description that is easier to understand for human users. See van der Waa, et al., at 1.

Calvano, et al., Ibid.

Vezzoso, Ibid.

About NERA

NERA Economic Consulting (www.nera.com) is a global firm of experts dedicated to applying economic, finance, and quantitative principles to complex business and legal challenges. For over half a century, NERA’s economists have been creating strategies, studies, reports, expert testimony, and policy recommendations for government authorities and the world’s leading law firms and corporations. We bring academic rigor, objectivity, and real world industry experience to bear on issues arising from competition, regulation, public policy, strategy, finance, and litigation.

NERA’s clients value our ability to apply and communicate state-of-the-art approaches clearly and convincingly, our commitment to deliver unbiased findings, and our reputation for quality and independence. Our clients rely on the integrity and skills of our unparalleled team of economists and other experts backed by the resources and reliability of one of the world’s largest economic consultancies. With its main office in New York City, NERA serves clients from more than 25 offices across North America, Europe, and Asia Pacific.

Contact
For further information and questions, please contact the author:

Dr. Ai Deng
Associate Director
Washington, DC: +1 202 466 9210
ai.deng@nera.com

To receive publications, news, and insights from NERA, please visit www.nera.com/subscribe.

The opinions expressed herein do not necessarily represent the views of NERA Economic Consulting or any other NERA consultant. Please do not cite without explicit permission from the author.