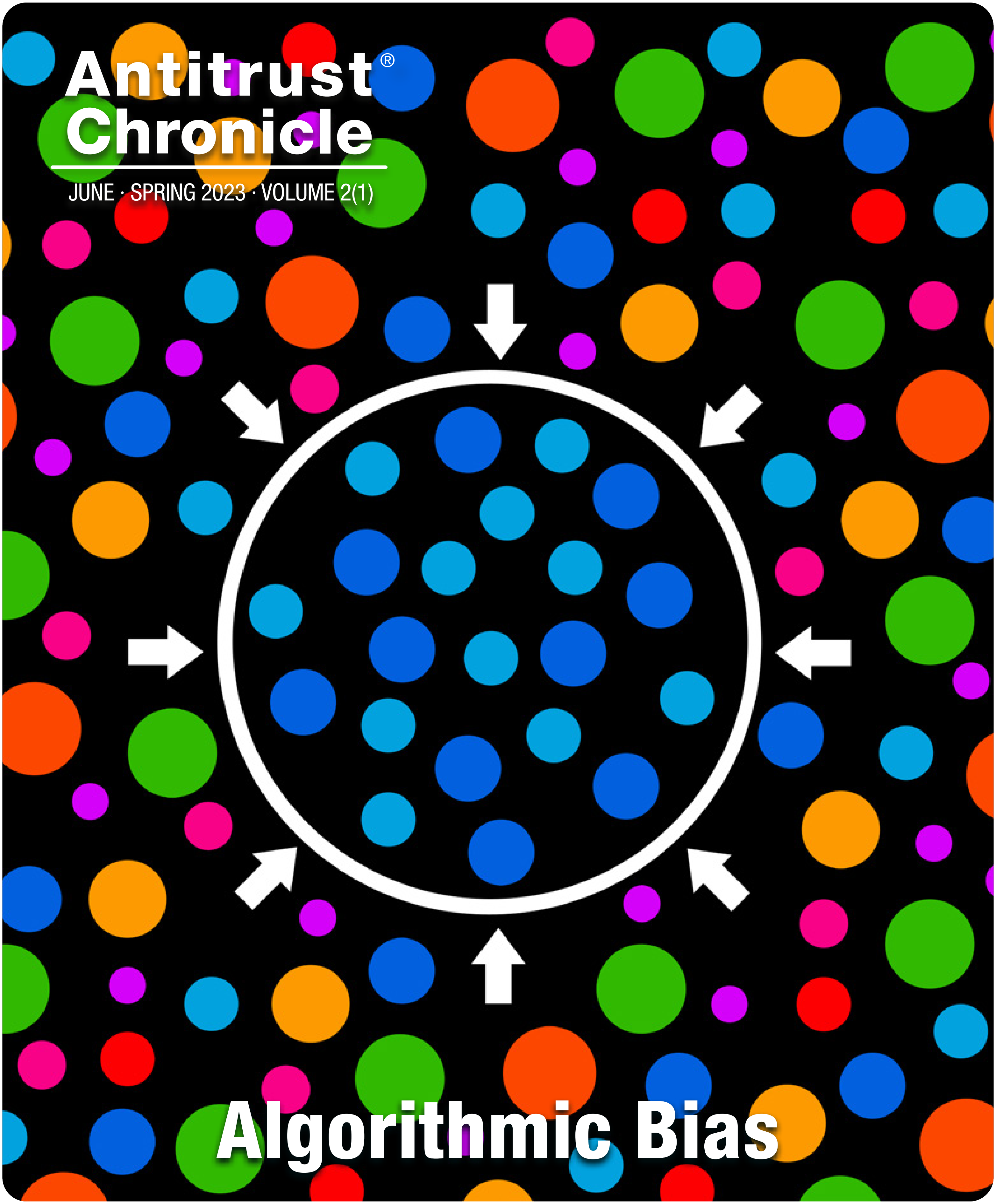
Antitrust® Chronicle

JUNE · SPRING 2023 · VOLUME 2(1)



Algorithmic Bias

CPI COMPETITION POLICY® INTERNATIONAL

TABLE OF CONTENTS

LETTER FROM THE EDITOR

Dear Readers,

Despite first being coined in the 17th century, until very recently the term "algorithm" was obscure jargon rarely used outside the field of computer science. Today, discussion of algorithms is scarcely out of popular news headlines. Such is the increasing sway that algorithms exert over every aspect of modern life.

At its base, an algorithm is a neutral mechanism, defined as "a step-by-step procedure for solving a problem or accomplishing some end." Leaving to one side the advanced algorithms used by the tech giants, the processes schoolchildren use to learn addition, subtraction, division, and multiplication are themselves "algorithms," albeit implemented by pencil and paper. Nonetheless, the amplification of the power of algorithms by vast datasets and computational power risks their use to socially undesirable ends (either intentionally or unintentionally). This is typically the result of introducing a real or perceived bias into either the algorithm itself or the input data upon which it operates.

In the antitrust context, "algorithmic bias" has famously been invoked in cases accusing large tech companies of favoring their own products or services to the detriment of competitors. But this is the tip of the iceberg, as algorithms are embedded by definition into any automated process engaged in by commercial entities, and also increasingly influence aspects of the operation of public bodies (including, potentially, antitrust enforcers).

Needless to say, the terrain is vast, but the contributors to this Chronicle address several of the most salient issues raised by algorithms at the present moment, from a variety of perspectives.

Appropriately, **Giovanna Massarotto** opens by asking what "bias" is and how it is exhibited in algorithms. Drawing on linguistics and social sciences, the author underlines that bias in people is well-known and likely inevitable. Therefore, potentially problematic algorithms include those that raise issues of bias as they rely on past data, including historical biases or unrepresentative or insufficient samples. Therefore, while antitrust agencies must be critical in addressing issues related to bias implemented through algorithms, a more important question remains unresolved if we cannot explain why and how an AI algorithm is biased in the first place.

Addressing the most headline-grabbing instances of the antitrust analysis of algorithmic "bias," **Emilie Feyler & Veronica Postal** discuss the stringent rulings by competition authorities against allegedly self-preferencing algorithms used by digital platforms. As the authors note, from a theoretical perspective, such self-preferencing algorithms can have pro-competitive benefits. There is no consensus from the economic literature on whether pro-competitive benefits or possible anti-competitive considerations prevail in the context of such algorithms used by digital platforms. Determining the net impact of recommendation algorithms on competition and consumer welfare requires individualized analysis accounting for the workings of specific algorithms, the competitive context, and the market environment.

Leaving aside these banner cases, **Robert Clark & Daniel Ershov** examine the impact of algorithmic pricing software to the bread and butter of antitrust enforcement: retail prices for consumers. Most recent academic work has studied this question from a theoretical or experimental perspective. The authors describe the first empirical analysis of the consequences of wide-scale adoption of algorithmic pricing software, focusing on Germany's retail gasoline market, where, according to trade publications, software became widely available in 2017. The evidence suggests that the use of these algorithms increased margins (and more so in competitive markets), indicating that it may have softened competition.

Looking at the other side of the coin, i.e. enforcers making use of algorithms in their enforcement activities, **Holli Sargeant & Teodora Groza** examine the balance between the benefits of algorithms in antitrust enforcement and the genuine concerns surrounding bias. They argue that while algorithmic bias should not be ignored, algorithms can be valuable tools when carefully designed, and the overemphasis on bias concerns stems from a lack of technical understanding. The article explores the use of algorithms in law enforcement, highlights the risks of bias, and presents how algorithmic design can mitigate these concerns. By offering a nuanced perspective on the potential and threats of algorithmic tools, the article contributes to the ongoing discourse on the responsible and effective utilization of algorithms in antitrust enforcement.

In the same vein, **Sampath Kannan** explains how machine learning algorithms are being used to make critical decisions. Eliminating bias in these algorithms and making them fair to groups and individuals is vitally important. The piece reviews common definitions of "fairness" for algorithms, and points a way out of the seeming impasse of choosing between mutually incompatible criteria in some scenarios. The piece identifies what algorithmic decision-making systems can learn from their human counterparts, i.e. rules-of-evidence-type limitations on the kinds of data that may be used, and exercising forbearance in making highly preemptive decisions. Finally, the article describes the need for transparent and accountable machine-learning models in the implementation of any such algorithmic systems.

Finally, and appropriately enough, **Paola Cecchi Dimeglio** underlines the importance of dialogue and innovation when it comes to correcting any dangerous trends in the use of algorithmic AI. Managing and eliminating algorithmic bias in virtual, augmented, and mixed reality technologies will be critical to this overall success. The responsibility for finding an appropriate balance lies both on businesses themselves and the legal system.

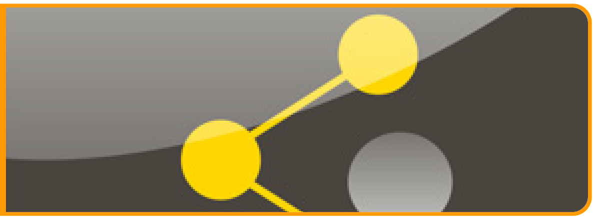
As always, many thanks to our great panel of authors.

Sincerely,

CPI Team

SUMMARIES

08



WHAT IS ALGORITHMIC BIAS AND WHY ANTITRUST AGENCIES SHOULD CARE?

By Giovanna Massarotto

Before asking whether algorithmic bias is a competition concern, we might need to understand what bias is and how it is exhibited in algorithms. Bias in people is well known and likely inevitable, raising the question of what we can do if algorithms learn to be bias. Artificial intelligence ("AI") algorithms are those that can raise issues of bias as they learn from past data, and we might have to deal with *historical bias* situations or unrepresentative or insufficient data. AI algorithms are built by software developers, who can also be biased. Companies are increasingly using AI algorithms to compete more effectively. It is fundamental that antitrust agencies tackle anticompetitive practices performed by means of algorithms, which might imply algorithmic bias. Bias is a broad term and exclusive practices are likely to increase bias in consumers. Therefore, antitrust agencies can be critical in addressing issues related to algorithmic bias. However, a more important question remains unresolved if we cannot explain why and how an AI algorithm is biased in the first place.

14

UNLEASHING THE POWER OF ALGORITHMS IN ANTITRUST ENFORCEMENT: NAVIGATING THE BOUNDARIES OF BIAS AND OPPORTUNITY

By Holli Sargeant & Teodora Groza

In the digital age, the intersection of data, technology, and antitrust enforcement has brought algorithms into focus as potential tools for uncovering anticompetitive practices and improving decision-making. However, concerns about algorithmic bias have raised questions about their use in this critical field. This article examines the balance between the benefits of algorithms in antitrust enforcement and the genuine concerns surrounding bias. It argues that while algorithmic bias should not be ignored, algorithms can be valuable tools when carefully designed, and the overemphasis on bias concerns stems from a lack of technical understanding. The article explores the use of algorithms in law enforcement, highlights the risks of bias, and presents how algorithmic design can mitigate these concerns. It then delves into the specific context of antitrust enforcement, explaining why the problem of algorithmic bias is less relevant compared to other regulatory areas. By offering a nuanced perspective on the potential and threats of algorithmic tools, the article contributes to the ongoing discourse on the responsible and effective utilization of algorithms in antitrust enforcement.

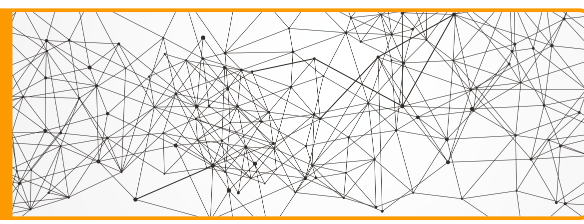
21

ALGORITHMIC PRICING AND COMPETITION

By Robert Clark & Daniel Ershov

In this article we examine the impact of adoption of AI-driven algorithmic pricing software. Most of the recent academic work has studied this question from a theoretical or experimental perspective. We describe the first empirical analysis of the consequences of wide-scale adoption of algorithmic pricing software, focusing on Germany's retail gasoline market, where, according to trade publications, software became widely available in 2017. The evidence suggests that adoption increased margins and more so in competitive markets, indicating that it may have softened competition.

26



CAN WE GET THE BIAS OUT OF OUR AI?

By Paola Cecchi Dimeglio

This article sheds light on how algorithms, originally intended to promote fairness and automation, can inadvertently perpetuate discrimination. By examining various domains such as employment, housing, banking, and education, one can uncover the far-reaching effects of bias, influencing outcomes and potentially reinforcing societal prejudices. Recognizing the urgency of the matter, the article underscores the significance of early detection and effective intervention to address algorithm bias. It highlights valuable strategies like diverse team involvement, inclusive dataset testing, and robust monitoring and review processes to identify and rectify biases. Transparency and user feedback play vital roles in mitigating bias and fostering a sense of fairness. With a collective responsibility, individuals and organizations are called upon to confront algorithm bias head-on. The aim is to forge AI systems that transcend default biases, aligning with the fundamental principles of equity and inclusivity. By embracing best practices, we can strive for a future where AI algorithms stand as unbiased pillars, actively contributing to a society that is truly equitable.

SUMMARIES

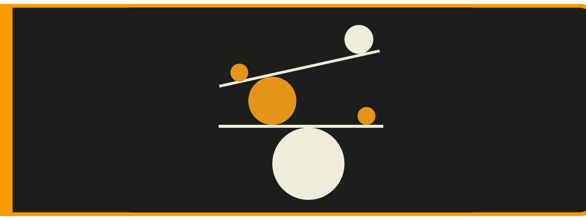
32

CAN SELF-PREFERENCING ALGORITHMS BE PRO-COMPETITIVE?

By Emilie Feyler & Veronica Postal

In response to the growing concerns around artificial intelligence, algorithms, and their influence over consumers' choices, competition authorities have adopted more stringent rules regarding self-preferencing algorithms used by digital platforms. However, from a theoretical perspective, self-preferencing algorithms can have pro-competitive benefits. There is no consensus from the economic literature on whether pro-competitive benefits or possible anti-competitive considerations prevail in the context of self-preferencing algorithms used by digital platforms. Determining the net impact of recommendation algorithms on competition and consumer welfare requires individualized analysis accounting for the workings of specific algorithms, competitive context, and market environment.

37



FAIRNESS IN ALGORITHMIC DECISION MAKING

By Sampath Kannan

This article explains how machine learning algorithms work and how they are being used to make critical decisions. Eliminating bias in these algorithms and making them fair to groups and individuals is vitally important. We review a few common definitions of fairness of algorithms, and point a way out of the seeming impasse of choosing between mutually incompatible criteria in some scenarios. We identify two things that algorithmic decision-making systems can learn from human decision-making systems – rules-of-evidence-type limitations on the kinds of data that may be used, and exercising great forbearance in making highly preemptive decisions. Finally, we describe what machine learning models are and describe the need for transparent and accountable models..

WHAT'S NEXT?

For July 2023, we will feature an Antitrust Chronicle focused on issues related to (1) **Coordinated Effects**; and (2) **Judicial Review of Economic Evidence**.

ANNOUNCEMENTS

CPI wants to hear from our subscribers. In 2023, we will be reaching out to members of our community for your feedback and ideas. Let us know what you want (or don't want) to see, at: antitrustchronicle@competitionpolicyinternational.com.

CPI ANTITRUST CHRONICLES August 2023

For August 2023, we will feature an Antitrust Chronicle focused on issues related to (1) **State AGs**; and (2) **EAB Antipasto**.

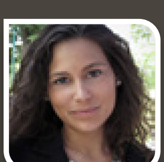
Contributions to the Antitrust Chronicle are about 2,500 – 4,000 words long. They should be lightly cited and not be written as long law-review articles with many in-depth footnotes. As with all CPI publications, articles for the CPI Antitrust Chronicle should be written clearly and with the reader always in mind.

Interested authors should send their contributions to Sam Sadden (ssadden@competitionpolicyinternational.com) with the subject line "Antitrust Chronicle," a short bio and picture(s) of the author(s).

The CPI Editorial Team will evaluate all submissions and will publish the best papers. Authors can submit papers on any topic related to competition and regulation, however, priority will be given to articles addressing the abovementioned topics. Co-authors are always welcome.

WHAT IS ALGORITHMIC BIAS AND WHY ANTITRUST AGENCIES SHOULD CARE?

BY GIOVANNA MASSAROTTO[1]



1 Academic Fellow, Center for Technology, Innovation and Competition, University of Pennsylvania Carey Law School.

I. INTRODUCTION

Before asking whether algorithmic bias is a competition concern, we might need to understand what bias is and how it is exhibited in algorithms. Bias in people is well known and likely inevitable, raising the question of what we can do if algorithms learn to be bias. Artificial intelligence ("AI") algorithms are those that typically raise issues of bias as they learn from past data, and we might have to deal with *historical bias* situations or unrepresentative or insufficient data. AI algorithms are built by software developers, who might also be biased. Companies are increasingly using AI algorithms to compete more effectively. It is fundamental that antitrust agencies tackle anticompetitive practices performed by means of algorithms, which might imply algorithmic bias. Bias is a broad term and exclusive practices are likely to increase bias in consumers. Therefore, antitrust agencies can be critical in addressing issues related to algorithmic bias. However, a more important question remains unresolved if we cannot explain why and how an AI algorithm is biased in the first place.

II. WHAT IS (ALGORITHMIC) BIAS?

Antitrust deals with competitive concerns in markets. Thus, in the algorithmic bias debate we might ask the question: *Is bias a competition concern?*

The most correct answer is: "it depends." First, we should clarify what we mean by bias and understand if an algorithm can be biased and how.[2]

In 1950, Alan Turing, the father of computer science, asked the question "Can machines think?" arguing that this question was "too meaningless to deserve a discussion,"[3] because it depends on what we mean by thinking. Similarly, if we ask the question "can humans fly?" the answer relies on what we mean by "flying."[4] When we take a plane, in a certain way we fly. However, this flying activity is different from that of birds, like eagles. Therefore, in investigating the meaning of bias, we need to address the issue of what we mean by bias, considering the context in which bias is analyzed. We usually use the word "bias" to refer to racial bias or gender bias. However, bias typically describes a wide range of behavior which can be harmful for different reasons and in different ways.[5] The Black's Law Dictionary defines bias as "inclination; bent; prepossession; a preconceived opinion; a predisposition to decide a cause or an issue in a certain way, which does not leave the mind perfectly open to conviction." It is considered different from "prejudice." "Bias is a particular influential power, which sways the judgment; the inclination of the mind towards a particular object."[6]

Thus, bias is a very serious concern. It is easy to imagine people that are influenced to lean towards a specific direction, "which does not leave the mind open to conviction;" thus bias in people. Before the digital age, bias in the news was well known. In the popular book "Manufacturing Consent," Professor Noam Chomsky and Edward S. Herman describe the bias phenomenon connected to the media. How the newspaper we choose to read selects the news it reports on, *de facto* generates bias in how and what we think. There are hundreds of wars every day. We typically know about one, or a few wars. How the news is selected can affect our "judgment." Advertising might cause bias in people for similar reasons. Bias is often unintentional, but it is present because it seems unrealistic to think that we can know about everything that is happening in the world or we know about certain products without proper advertisement. Thus, bias as something that affects our conviction seems inevitable in human beings.

What about algorithms? An algorithm is generally defined as a set of instructions in which there is some input to obtain certain output. The first non-trivial algorithm was Euclidean, a method for calculating the greatest common divisors between two integers. It seems hard to

2 The problem of meaning in language is a fundamental issue belonging to the linguistic domain, which is extremely relevant in any discussion related to natural language processing, thus artificial intelligence ("AI"). Considering that AI algorithms are those that raise bias concerns, starting from a linguistic question seems to be extremely pertinent. It brings us back to the origin of the modern computer and the first AI algorithms. See e.g. Bennison Gray, *The Problem of Meaning in Linguistic Philosophy*, 59 Logique et Analyse 609 (1972).

3 Alan M. Turing, *Computing Machinery and Intelligence*, 59 Mind 433 (1950); Noam Chomsky, *Turning on the "Imitation Game,"* in Passing the Turning Test (Richard Epstein, Gary Roberts & Grace Beber eds., Springer, 2009).

4 Noam Chomsky, Chickens fly like eagles. Humans don't fly at all (May 17, 2017), Inframethodology, https://blog.cbs.dk/inframethodology/?p=568.

5 Sun Lin Blogdgett, Solon Barocas, Hal Daume' III & Hanna Wallach, *Language (Technology) is Power: A Critical Survey of "Bias" in NLP*, Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics 5454 (2020).

6 The Law Dictionary, BIAS Definition & Legal Meaning (Black's Law Dictionary, 2nd ed.), https://thelawdictionary.org/bias/#:~:text=Inclination%3B%20bent%3B%20prepossession%3A%20a,mind%20perfectly%20open%20to%20conviction.

believe that Euclidean or similar algorithms can be biased. These algorithms are quite straightforward. The issue of bias particularly concerns AI algorithms, which are algorithms trained with a large amount of data to build models that make predictions related to the information of interest.[7] In AI algorithms, while the input is known, the output is usually unpredictable. Deep Learning ("DL), for example, is an AI method, relying on complex neural network architectures inspired by the human brain.[8] ChatGPT is a symbolic example of an AI system, which adopts deep learning techniques. These methods can build models that are very good for predicting the behavior of a system, but in turn are often very bad with explaining why the model predicts a certain behavior, challenging its validity.[9] The interpretation of AI results is becoming increasingly challenging due to the sophistication of these models and methods. Therefore, in addition to data, the design choice is important and needs to be considered in assessing an AI system and potential for bias, because some design choices in algorithms "are better than others."[10] For example, the selection of features, as well as the algorithmic assumptions can determine bias in AI models.[11] Software engineers typically make these choices when they build AI systems. Although engineers' decisions are led by technical reasons, it is challenging to think that we can ensure with certainty that these engineers are fully unbiased.[12] In other words, similar to human bias, algorithmic bias represents a concrete concern, and it seems inherent in the fact that these algorithms learn from past data and software developers who build AI systems can make technical choices that can generate bias. Above all, we need to consider that we might fail to understand why and how the algorithm made certain predictions in the first place.

In summary, AI algorithms rely on data to work and are becoming increasingly sophisticated, thus rendering the interpretation of their results difficult. Data quality and selection are essential to the AI system's performance, as much as the algorithm design. Both data used to train AI algorithms and the algorithm design are not "impartial"[13] and relevant from a bias perspective. However, how to deal with algorithmic bias might be challenging as we often cannot explain an AI system's results by detecting bias and its cause in the first place.

To make the discussion more intuitive, we can use the popular *ProPublica* study dating back to 2016, which examined Compas, an algorithm adopted by the U.S. legal system to facilitate judicial decision making.[14] Compas was trained to specifically assist judges in the U.S. in deciding whether a defendant was likely to re-offend while the trial was pending. The problem was that Compas "was found to be biased against African-Americans."[15] Nicol Turner Lee, Paul Resnick and Genie Barton noted that "[i]n the COMPAS algorithm, if African-Americans are more likely to be arrested and incarcerated in the U.S. due to historical racism, disparities in policing practices, or other inequalities within the criminal justice system, these realities will be reflected in the training data and used to make suggestions about whether a defendant should be detained. If historical biases are factored into the model, it will make the same kinds of wrong judgments that people do."[16] Moreover, data might be insufficient or unrepresentative. If data used to train the algorithm represents one group more rather than others, this disproportion is likely to be reflected in the model and AI, therefore leading to bias at scale. In addition, algorithm design choices are also "not impartial."[17]

---

7 See Tom Mitchell, Machine Learning 2 (1997) ("A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience.").

8 See Yann LeCun et al., *Deep Learning*, 521 Nature 436 (2015).

9 Tomaso Aste, *What Machines Can Learn About Our Complex World - and What Can We Learn From Them?* 7 (Mar. 4, 2021), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3797711.

10 Sara Hooker, *Moving Beyond 'Algorithmic Bias is a Data Problem'*, 2 Patterns 1 (2021). See also, Nicole Turner Lee, Paul Resnick & Genie Barton, *Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms*, Brookings Report (May 22, 2019), https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/. ("Turner Lee has argued that it is often the lack of diversity among the programmers designing the training sample which can lead to the under-representation of a particular group or specific physical attributes.") *Id*.

11 Drew Roselli, Jeanna Matthews & Nisha Talaga, *Managing Bias in AI,* Companion Proceedings of the 2019 World Wide Web Conference 539, 541 (2019).

12 See Lee, Resnick & Barton, *supra* note 10. ("Turner Lee has argued that it is often the lack of diversity among the programmers designing the training sample which can lead to the under-representation of a particular group or specific physical attributes.") *Id*. See also, Bo Cowgill & Catherine Tucker, *Algorithmic Fairness and Economics* (Sep. 24, 2020), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3361280. ("According to the Bureau of Labor Statistics in 2018, software engineers are more white, male, well-educated and better-paid than America as a whole.") *Id*. at 6.

13 Hooker, *supra* note 10.

14 See Jeff Larson, Surya Mattu, Lauren Kirchner & Julia Angwin, *How We Analyzed the COMPAS Recidivism Algorithm*, ProPublica (May 23, 2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm; Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan & Ashesh Rambachan, *Algorithmic Fairness*, 108 AEA Papers and Proceedings 22 (2018).

15 See Lee, Resnick & Barton, *supra* note 10, at 5. Larson, Mattu, Kirchner & Angwin, *supra* note 14. "Black defendants were twice as likely as white defendants to be misclassified as a higher risk of violent recidivism, and white recidivists were misclassified as low risk 63.2 percent more often than black defendants. Black defendants who were classified as a higher risk of violent recidivism did recidivate at a slightly higher rate than white defendants (21 percent vs. 17 percent), and the likelihood ratio for white defendants was higher, 2.03, than for black defendants, 1.62." *Id*

16 *Id*.

17 Hooker, *supra* note 10.

Thus, the primary question now becomes whether it is possible to prevent bias in AI algorithms and how. Being aware of this risk certainly represents the first important step. The second obvious step is to analyze the phenomenon in a way that we can diagnose and limit bias in AI algorithms effectively. Many studies have been conducted to develop a regulatory framework that mitigates the risk of bias in algorithms, which potentially can occur at large scale. In Europe, the Artificial Intelligence Act mainly aims to strengthen rules around data quality, transparency, human oversight, as well as accountability.[18] AI systems would be classified based on their related risks from safety to fundamental rights of a person. AI applications "considered a clear threat to the safety, livelihoods and rights of people will be banned."[19] In the U.S., the approach is more about adapting the existing legal framework to AI and investing in "infrastructure for mitigating AI risks."[20] As outlined above, creating AI systems that are unbiased seems very challenging, but reducing the risk of algorithmic bias is certainly in the crosshairs of legislators all over the world.[21]

Now, the issue is whether algorithmic bias affects competition and how antitrust enforcers might assist in limiting algorithmic bias. In other words:

III. IS ALGORITHMIC BIAS A COMPETITION CONCERN?

Although at first glance algorithmic bias might not seem to be a strict competition issue, algorithms are interesting from a competitive perspective for several reasons, including their potential to be biased. Consider the example of companies like Amazon and Google, which use algorithms to perform basically all activities and to compete in markets. Companies are increasingly using algorithms to do what they regularly do more efficiently, and developing the best algorithm for a company often implies winning a market. AI algorithms are general purpose technologies, and they can be implemented in different contexts and situations. AI algorithms are widely used, for example, in the advertising, search and media industries. Therefore, creating algorithms has become increasingly important for any company, regardless of the market, to compete and remain relevant.

Algorithms are often developed to provide recommendations; to automate the distribution and allocation of the demand and supply; to set or recommend a price; to monitor or filter information; to aggregate data and communicate with consumers and businesses; and they can create bias at scale.[22]

The OECD has recently released a study on algorithmic competition with a focus on algorithmic recommendation, search, allocation, pricing and monitoring algorithms, which considers the number of antitrust agencies' papers and reports drafted on these issues.[23] Because algorithms have become one of the main tools for companies to compete, it is fundamental that antitrust agencies understand and analyze these algorithms.[24] Antitrust agencies need to ensure that companies are not using algorithms to engage in anticompetitive conduct, including price fixing and exclusive conduct. This does not seem to be an option. It is essential for antitrust to remain relevant in the present data-driven economy run by algorithms.

18 News European Parliament, AI Act: A Step Closer to the First Rules on Artificial Intelligence, Press Releases (May 11, 2023), https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence.

19 European Commission, *Regulatory framework proposal on artificial intelligence* (last update Sept. 29, 2022), https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

20 Alex Engler, *The EU and U.S. Diverge on AI Regulation: A Transatlantic Comparison and Steps to Alignment*, Brookings Report (Apr. 25, 2023), https://www.brookings.edu/research/the-eu-and-us-diverge-on-ai-regulation-a-transatlantic-comparison-and-steps-to-alignment/.

21 See World Economic Forum, The European Union's Artificial Intelligence Act, explained (Mar. 28, 2023), https://www.weforum.org/agenda/2023/03/the-european-union-s-ai-act-explained/. See also, the Algorithmic Accountability Act of 2022 in the United States. Wyden, Booker and Clarke Introduce Algorithmic Accountability Act of 2022 To Require New Transparency And Accountability For Automated Decision Systems, Ron Wyden United States Senator for Oregon (Feb. 3, 2022), https://www.wyden.senate.gov/news/press-releases/wyden-booker-and-clarke-introduce-algorithmic-accountability-act-of-2022-to-require-new-transparency-and-accountability-for-automated-decision-systems.

22 OECD, *Algorithmic Competition*, OECD Competition Policy Roundtable Background Note 8-9 (2023), https://www.oecd.org/competition/algorithmic-competition.htm. [OECD Report].

23 *Id.* at 7, 10.

24 Giovanna Massarotto, *Why AI and Competition Law Matter?*, On-Topic, 3 Concurrences 2 (2021); Giovanna Massarotto, *Using Tech to Fight Big Tech*, Bloomberg Law (Sep. 27, 2021), https://news.bloomberglaw.com/tech-and-telecom-law/using-tech-to-fight-big-tech.

The Competition and Markets Authority ("CMA") has been particularly active in this field having a Data, Technology and Analytics ("DaTA") unit dedicated to these issues since 2018.[25] Many antitrust agencies are following the CMA's DaTA unit model by creating a dedicated unit focused on data and algorithmic matters.[26]

Antitrust studies on algorithmic competition clarified that algorithms enhance consumer welfare by increasing products and services qualities.[27] In 2018, a study of the OECD revealed that pricing algorithms based on AI techniques can benefit consumers significantly. UberX, for example, matched drivers to consumers seeking rides by means of a real-time pricing algorithm and a study estimated that this service generated a consumer surplus of $2.9 billion in four U.S. cities.[28] On the other hand, it has been observed that algorithms might reduce competition by favoring collusion or exclusionary and exploitative conduct. The attention of antitrust regulators has been focused on self-preferencing, autonomous tacit collusion, algorithmic pricing and algorithmic tying and bundling cases. It is not surprising that algorithms can perform all these practices, being a set of instructions with input to generate specific output. For example, vertical integrated digital platforms have raised the issue of so-called "intermediation bias" by potentially using their algorithm to favor their own products over those of a competitor.[29] The Google shopping case is a symbolic example.[30] Therefore, the issue of bias is vibrant in this antitrust discussion as if companies can use algorithms to engage in self-preferencing, collusion or excluding rivals more efficiently, they can de facto limit consumers' choices in a way that can lead to bias.

Antitrust agencies have a great responsibility, being the first arm of government regulation that can enforce competition principles in any markets by imposing remedies regulatory in nature, while Congress enacts a new law or sets up a new ad hoc regulatory agency.[31]

Considering algorithms for pricing decisions, recent studies have revealed the lack of "comprehensive data of firms using algorithms and AI for pricing purposes."[32] The studies available seem to show that algorithms used to monitor competitors' prices are quite uncommon and the end price is rarely adjusted automatically by an algorithm after having considered the other companies' prices. The same applies for personalized pricing. On the other hand, the risk that companies use algorithms to tacitly collude seems to rapidly increase.[33] However, conscious parallelism, which economists call "tacit collusion," generally is not considered unlawful in itself. If we consider exclusive conduct, including tying and unbundling, by means of algorithms, this seems perfectly plausible and, as in a non-algorithmic situation, it needs to be assessed case by case.

Several techniques to examine algorithms' design and functioning exist. Algorithmic auditing and reverse engineering seem to be the most promising methods to assess whether the algorithm can lead to an anticompetitive behavior and potentially increase the risk of bias in algorithms. Several legislations proposed to mandate algorithmic impact assessment or audit provisions to ensure a trustworthy development in AI by having consumer protection and welfare as the main goal. On the other hand, antitrust agencies can impose similar obligations on large market players that are using algorithms to limit competition and harm consumers by increasing transparency and AI accountability, without the need to wait for a new law.[34] Therefore, antitrust can be critical in addressing algorithmic issues by investigating such issues and finding solutions with large players.

---

25 Stephan Hunt, CMA's new DaTA unit: exciting opportunities for data scientists, https://competitionandmarkets.blog.gov.uk/2018/10/24/cmas-new-data-unit-exciting-opportunities-for-data-scientists/.

26 See e.g. Brian Fung, DOJ will hire more data experts to scrutinize digital monopolies, antitrust chief says, Cnn Business (Mar. 6, 2023), https://www.cnn.com/2023/03/06/tech/doj-data-experts/index.html.

27 See e.g. OECD Report at 10; Antonio Capobianco, *The Impact of Algorithms on Competition and Competition Law*, ProMarket (May 23, 2023), https://www.promarket.org/2023/05/23/the-impact-of-algorithms-on-competition-and-competition-law/.

28 Peter Cohen, Robert Hahn, Jonathan Hall, Steven Levitt & Robert Metcalfe, *Using Big Data to Estimate Consumer Surplus: The Case of Uber*, NBER Working Paper (Sep. 2016), https://www.nber.org/papers/w22627.

29 See e.g. Richard Feasey & Jan Krämer, *Implementing Effective Remedies for Anti-Competitive Intermediation Bias on Vertically Integrated Platforms*, CERRE Centre on Regulation in Europe Report 5 (Oct. 2019), available at https://cerre.eu/wp-content/uploads/2020/05/cerre_report_intermediation_bias_remedies.pdf.

30 European Commission Press Release, Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service, (Jun. 27, 2017), https://ec.europa.eu/commission/presscorner/detail/en/IP_17_1784; EU Commission, Google Search (Shopping), AT.39740 (Jun. 27, 2017).

31 See Giovanna Massarotto, Antitrust Settlements. How a Simple Agreement Can Drive the Economy 75, 145 (Wolters Kluwer, 2019); Giovanna Massarotto, *Grasping the Meaning of Big Tech Antitrust Consent*, Competition Policy International (Feb. 2020), https://www.competitionpolicyinternational.com/grasping-the-meaning-of-big-tech-antitrust-consent/.

32 See Capobianco, *supra* note 27.

33 *Id.*

34 See *supra* note 31.

However, while several issues related to bias can be addressed through audit provisions and mandate algorithmic impact assessments, there seems to remain an unresolved challenge for legislators and antitrust enforcers: How do we tackle bias effectively in algorithms in which we cannot even explain if and why there is bias in their results? This seems an important question that goes beyond algorithmic bias by challenging the foundations of our present scientific method.[35]

Will algorithms learn to be unbiased autonomously? They might do, but it seems important that we understand how they are capable of doing so. The alternative of using the algorithm blackbox to justify what remains unknown does not seem to be an effective solution.

IV. CONCLUSION

In summary, antitrust agencies can potentially set the tone for future AI development by requiring relevant players to ensure certain standards of fairness in the algorithms' decision making process to preserve competition in specific circumstances. Transparency obligations seem to be particularly important to achieve this end. Therefore, how antitrust agencies enforce competition principles in the context of algorithms can affect "algorithmic bias." However, we still have little technical comprehension of certain AI models and what they can predict. Thus, although critical, antitrust enforcement action might not be sufficient in addressing a more important question. Should we allow the adoption of algorithms whose results we cannot explain? Is this the start of a new scientific revolution or an old problem with an easy solution?

35 See Aste, *supra* note 9, at 7.

UNLEASHING THE POWER OF ALGORITHMS IN ANTITRUST ENFORCEMENT: NAVIGATING THE BOUNDARIES OF BIAS AND OPPORTUNITY

BY HOLLI SARGEANT & TEODORA GROZA[1]

1 Holli Sargeant is a PhD Candidate at the Faculty of Law, University of Cambridge. Teodora Groza is a PhD Candidate at Sciences Po Law School and Editor-in-Chief of the Stanford Computational Antitrust Journal.

I. INTRODUCTION

In the digital age, where data and technology permeate every facet of our lives, the field of antitrust enforcement finds itself at a crossroads. As policymakers and regulators grapple with the ever-evolving landscape of digital marketplaces, the spotlight has turned toward the power of algorithms to enforce antitrust rules. Algorithms have the unique potential to empower antitrust enforcement by unearthing anticompetitive practices, predicting market trends, and enabling more informed and economics-based decision-making. However, concerns about algorithmic bias have cast a shadow over their use in this critical domain.

This article explores the delicate balance between the benefits of algorithms in antitrust enforcement and the genuine concerns surrounding algorithmic bias. While algorithmic bias, the potential for algorithms to discriminate or perpetuate unfair outcomes, should not be dismissed, it is essential to approach the topic with a nuanced understanding. Despite concerns about algorithmic bias, we show that algorithms are a valuable tool for antitrust enforcement, and a lack of technical understanding has led to an overinflation of these concerns.

The article is structured as follows. It begins with an overview of the academic literature on the use of algorithms in law enforcement and of the concerns raised in this literature about the potential effects of algorithmic bias. Then, it engages with the literature on the potential of algorithms to improve law enforcement, explaining how concerns over algorithmic bias can be mitigated through careful algorithmic design. Leveraging these insights, it gives an overview of the use of algorithms in antitrust enforcement and explains that concerns over algorithmic bias are less relevant in antitrust as opposed to other fields of law due to certain particularities of antitrust laws.

The article makes three principal contributions. First, it offers a nuanced account of the potential and threats of leveraging algorithmic tools in law enforcement by introducing some critical aspects of machine learning that unveil some of the misconceptions about algorithmic bias. Second, it explains that concerns over algorithmic bias are less relevant to antitrust enforcement than other regulatory fields.

II. MAPPING THE DEBATE

As a recent report by the European Agency for Fundamental Rights put it, "AI is everywhere and affects everyone."[2] Artificial intelligence ("AI)" tools have become commonplace in law enforcement. These methods are primarily machine learning ("ML") approaches, a subset of AI, which is the focus of this article. Tax and construction authorities, social security agencies, and antitrust authorities worldwide increasingly rely on ML to perform their duties. A significant strand of the existing literature on the use of ML in law enforcement highlights the threats posed by new technologies in terms of enhancing existing biases and bolstering the discriminatory practices resulting from there. Nonetheless, algorithmic tools provide a significant opportunity for law enforcement and, if carefully designed, can mitigate human biases, increase efficiency, and ultimately lead to a shift from the current reactionary approach to regulation to an adaptive one.[3]

This section maps out the use of algorithms in law enforcement and explains the risks associated with algorithmic bias. Building on these, it introduces its key argument that well-designed AI tools have a game-changing potential for law enforcement.

A. The Use of Algorithms in Law Enforcement

In recent times, there has been a surge in the application of algorithms in different aspects of law enforcement. Although, it is often difficult to ascertain what algorithms are being built and deployed by law enforcement. One significant area where this trend is evident is in the adoption of algorithmic systems in criminal law enforcement practices. For example, the adoption of algorithmic systems in law enforcement practices presents a new dimension of crime prevention and detection. These data-driven models utilize vast arrays of information, predicting potential crime hotspots, identifying potential offenders, and even assisting in decision-making processes.[4] Algorithmic models are also used in the judicial enforcement of criminal bail and sentencing decisions.[5]

2 European Union Agency for Fundamental Rights, Bias in Algorithms: Artificial Intelligence and Discrimination 3 (2022), https://data.europa.eu/doi/10.2811/536044.

3 Lori S Bennear & Jonathan B Wiener, *Adaptive Regulation: Instrument Choice for Policy Learning over Time*, Draft working paper (2019), https://www.hks.harvard.edu/sites/default/files/centers/mrcbg/files/Regulation%20-%20adaptive%20reg%20-%20Bennear%20Wiener%20on%20Adaptive%20Reg%20Instrum%20Choice%202019%2002%2012%20clean.pdf.

4 Miri Zilka, Holli Sargeant & Adrian Weller, *Transparency, Governance and Regulation of Algorithmic Tools Deployed in the Criminal Justice System: a UK Case Study, in* Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (2022), https://doi.org/10.1145/3514094.3534200 (last visited Jun 8, 2022).

5 See e.g. Julia Angwin et al., *Machine Bias*, ProPublica (2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; Zilka, Sargeant & Weller, *supra* note 4.

Beyond criminal law enforcement, algorithms have also found utility in the administration of social welfare programs. The Australian Government used an automated debt recovery program, known as "Robodebt," to identify any discrepancies between an individual's declared income to the Australian Taxation Office and the individual's income reported to Centrelink (the social welfare department).[6] The algorithm faced considerable backlash for inaccuracies, which resulted in unjustified debt notices, leading to widespread public distrust of government use of algorithmic tools.[7]

The utilization of algorithms in law enforcement highlights both the potential benefits and significant challenges associated with algorithmic enforcement. While these models hold promise to enhance efficiency and effectiveness, ensuring transparency, fairness, and accountability is crucial to prevent unintended consequences and preserve public trust in algorithmic decision-making processes.

B. What is Algorithmic Bias?

Algorithmic bias refers to the potential for algorithms to discriminate against certain individuals or perpetuate unfair outcomes based on factors such as race, gender, or other characteristics.[8] In simple words, it covers situations when "AI makes decisions that are unfair to certain groups."[9] Nonetheless, unlike humans, AI has neither intentionality nor biases: This means that algorithms cannot be independently inclined toward certain outcomes and do not prefer certain characteristics at the expense of others. The issue of algorithmic bias arises where "[algorithmic] models encode human prejudice, misunderstanding, and bias into the software systems that increasingly managed our lives."[10] Seen through this lens, the notion of algorithmic bias is rather a misnomer: It is not the algorithm that is biased but the humans behind it.

Algorithmic bias can arise as a byproduct of two factors: biased datasets and biased design choices. On the one hand, a wealth of literature has identified the risk that algorithmic models replicate human biases found encoded in training datasets.[11] Reflecting on the use of large datasets in antitrust enforcement, Eleanor Fox noted that "[w]hen you talk about data, you also have to talk about values . . . And assumptions."[12] To illustrate how datasets can prejudice against certain groups, take the example of facial-analysis datasets that contain a preponderance of lighter-skinned subjects: Reliance on such datasets leads to higher error rates for subjects with darker skin.[13]

However, it is not all about the input data. A less developed strand of literature highlights that algorithmic bias is not only a byproduct of the data being used, but rather of the "interactions between the data and the model design choices."[14] Consequently, bias is learned and sometimes amplified, both by the goals of the model design and by reverberating the patterns found in training data.[15] This leads to the fundamental insight that the design of algorithms is not value-neutral, and the threshold rules or weights assigned in the algorithm may reveal human biases.[16]

---

6 Jordan Hayne & Matthew Doran, *Government to pay back $721m in Robodebt, all debts to be waived*, ABC News, May 29, 2020, https://www.abc.net.au/news/2020-05-29/federal-government-refund-robodebt-scheme-repay-debts/12299410 (last visited May 30, 2023).

7 Matthew Doran, *Federal Government ends Robodebt class action with settlement worth $1.2 billion*, ABC News, Nov. 16, 2020, https://www.abc.net.au/news/2020-11-16/government-response-robodebt-class-action/12886784 (last visited May 30, 2023); Australian Human Rights Commission, *Human Rights and Technology*, (2021).

8 Australian Human Rights Commission, *Using artificial intelligence to make decisions: Addressing the problem of algorithmic bias*, (2020), https://humanrights.gov.au/our-work/rights-and-freedoms/publications/using-artificial-intelligence-make-decisions-addressing.

9 PricewaterhouseCoopers, *Understanding algorithmic bias and how to build trust in AI*, PwC (2022), https://www.pwc.com/us/en/tech-effect/ai-analytics/algorithmic-bias-and-trust-in-ai.html (last visited May 30, 2023).

10 Cathy O'Neil, Weapons of math destruction: how big data increases inequality and threatens democracy (2016).

11 Robin Nunn, *Discrimination in the Age of Algorithms, in* The Cambridge Handbook of the Law of Algorithms (Woodrow Barfield ed., 2020), https://doi.org/10.1017/9781108680844.010; Deborah Hellman, *Measuring Algorithmic Fairness*, 106 Va. Law Rev. 811 (2020); Jon Kleinberg et al., *Discrimination in the Age of Algorithms*, 10 Journal of Legal Analysis 113 (2018).

12 Andrew Ross Sorkin et al., *Can Jane Fraser Fix Citigroup?*, The New York Times, Feb. 11, 2021, https://www.nytimes.com/2021/02/11/business/dealbook/jane-fraser-citigroup.html (last visited Apr 4, 2023).

13 Sara Hooker, *Moving beyond "algorithmic bias is a data problem,"* 2 Patterns (2021), https://doi.org/10.1016/j.patter.2021.100241.

14 *Id.*

15 Jeremias Adams-Prassl, Reuben Binns & Aislinn Kelly-Lyth, *Directly Discriminatory Algorithms*, 86 The Modern Law Review 144 (2023), https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-2230.12759 (last visited Feb 19, 2023); Reuben Binns, *Fairness in Machine Learning: Lessons from Political Philosophy*, 81 149 (2021), http://arxiv.org/abs/1712.03586 (last visited Nov 5, 2021); Sam Corbett-Davies & Sharad Goel, *The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning,* arXiv (2018); Sahil Verma & Julia Rubin, *Fairness Definitions Explained, in* Proceedings of the International Workshop on Software Fairness 1 (2018), https://doi.org/10.1145/3194770.3194776.

16 Dan Burk, *Algorithmic Fair* Use, 86 University of Chicago Law Review (2019).

While some ML models are easily explainable, other more complex algorithms pose challenges due to their complexity, hence the recurring metaphor of the algorithmic "black box."[17] However, the black box nature of all ML is somewhat of a misnomer. Even in large neural networks, we know how algorithms make predictions and usually understand the information used. The problem is not that the decision-making process is unexplainable but rather that predictions are made in a way that is difficult for humans to grasp due to the millions of parameters involved.[18] This reality is often confused with the well-known statement that ML decisions do not inherently generate reasons or explanations. Unlike human decision-makers — judges in particular — algorithmic tools do not necessarily accompany their decisions with explanations of how the outcome was reached. Achieving full transparency in ML remains a work in progress, and efforts toward explainability are crucial for fostering public scrutiny and moral accountability in decision-making processes.

The problem of algorithmic bias in law enforcement is most acute in regulatory fields that deal directly with human subjects and their rights, such as criminal, migration, or social security law. A case in point is the 2021 Dutch childcare benefits scandal, in which "the Dutch Tax and Customs Administration used algorithms in which "foreign sounding names" and "dual nationality" were used as indicators of potential fraud. The result was that thousands of (racialized) low- and middle-income families were "subjected to scrutiny, falsely accused of fraud, and asked to pay back benefits that they had obtained completely legally."[19]

C. Algorithmic Opportunity

The threat of algorithmic bias should not automatically translate into a wholesale rejection of the use of ML for law enforcement. To begin with, circling back to the insight that model design choices impact the functioning of algorithms, it is possible to make choices that circumvent potential biases enshrined in data pipelines.[20] Consequently, "recognizing how model design impacts harm opens up new mitigation techniques that are less burdensome than comprehensive data collection."[21]

Not only does careful design enable avoiding passing-on human biases to algorithmic tools, but it also represents a material opportunity to overcome certain shortcomings of human decision-makers. We must reiterate that "human decision-making is not significantly more accountable than AI,"[22] some scholars contending that algorithms can help move away from subjective human judgments to data-driven decisions that may be more accurate and unbiased.[23] Following such reasoning, algorithms may "overcome the cognitive limits and social biases of human decision-makers, enabling more objective and fair decisions."[24] Looking at the stages of developing ML tools,[25] the classification and decision rules that set thresholds for what optimal actions the model may take can consider the costs and benefits of making different decisions based on ML prediction under uncertainty.[26] Supervised ML expands on statistical decision theory that attempts to model the probability distribution of the possible, real-world effects for each decision option.[27] That can allow a decision-maker to add different weights to different risks or opportunities for algorithmic prediction as they attempt to operate under future uncertainty. Decision-makers already do this, but when modeling algorithms, the value judgments become clearer through encoded preferences.[28]

---

17 Pantelis Linardatos, Vasilis Papastefanopoulos & Sotiris Kotsiantis, *Explainable AI: A Review of Machine Learning Interpretability Methods*, 23 Entropy 18 (2020), https://www.mdpi.com/1099-4300/23/1/18 (last visited May 23, 2022).

18 *Id.*; Finale Doshi-Velez & Been Kim, *Towards A Rigorous Science of Interpretable Machine Learning,* arXiv (2017), http://arxiv.org/abs/1702.08608.

19 European Parliament, *The Dutch childcare benefit scandal, institutional racism and algorithms*, Parliamentary question - O-000028/2022 (28.6.2022), https://www.europarl.europa.eu/doceo/document/O-9-2022-000028_EN.html (last visited May 30, 2023).

20 Hooker, *supra* note 13.

21 *Id.*

22 Lim, *supra* note 11, at 48.

23 Kleinberg et al., *supra* note 11; Alice Xiang, *Reconciling Legal and Technical Approaches to Algorithmic Bias*, 88 Tenn. Law Rev. 649 (2021); Virginia Eubanks, Automating Inequality (2018); Alex P. Miller, *Want Less-Biased Decisions? Use Algorithms.*, Harvard Business Review, Jul. 2018, https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms (last visited May 26, 2023).

24 Ben Green, *Escaping the Impossibility of Fairness: From Formal to Substantive Algorithmic Fairness*, 35 Philos. Technol. 90 (2022), https://link.springer.com/10.1007/s13347-022-00584-6 (last visited Jan 22, 2023); See also S.1593—Pretrial Integrity and Safety Act of 2017, (2017), http://www.congress.gov/ (last visited May 26, 2023); Cass R. Sunstein, *Two Conceptions of Procedural Fairness*, 73 Social Research 619 (2006), http://www.jstor.org/stable/40971840 (last visited Aug 23, 2022).

25 See discussion in Holli Sargeant, *Algorithmic decision-making in financial services: economic and normative outcomes in consumer credit*, AI Ethics (2022), https://doi.org/10.1007/s43681-022-00236-7 (last visited Nov 23, 2022).

26 Ethem Alpaydin, Introduction to Machine Learning (4 ed. 2020), https://doi.org/10.7551/mitpress/13811.001.0001; Trevor Hastie, Robert Tibshirani & Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2 ed. 2009).

27 Alpaydin, *supra* note 26; Hastie, Tibshirani & Friedman, *supra* note 26.

28 Jon Kleinberg et al., *Human Decisions and Machine Predictions**, 133 The Quarterly Journal of Economics 237 (2018), https://doi.org/10.1093/qje/qjx032 (last visited Nov 22, 2022).

Moreover, there is a potential for algorithms to centralize decision-making. A "single algorithm has the potential to play the role of hundreds or thousands of human decision-makers," creating a potentially more transparent and auditable process.[29] Reliance on ML can thus harmonize bodies of law that are currently made up of patchworks of decisions taken by agencies and courts. Furthermore, the use of AI can make it easier to audit and evaluate the fairness of decisions.[30] Unlike relying on hundreds or thousands of individual human decision-makers, a single algorithm can be monitored and assessed for biases.

In conclusion, the reality of algorithmic bias should not overshadow the potential benefits that algorithms can offer in antitrust enforcement. While concerns about bias are valid, dismissing algorithms altogether would mean disregarding the opportunity to improve decision-making and overcome existing human biases. By carefully designing, implementing, and overseeing algorithmic tools, it is possible to harness their power while minimizing the risks of bias. Striking a balance between the benefits and risks is crucial to navigating a path toward responsible and effective utilization of algorithms in antitrust enforcement.

III. ANTITRUST ENFORCEMENT AND ALGORITHMIC BIAS

Whereas legal bodies ranging from tax to criminal law have relied on algorithmic tools for decades, the use of algorithms in antitrust enforcement has lagged behind. The topic has matured into a key focus of the antitrust community since 2021 when the Computational Antitrust project was launched to "explore how legal informatics could foster the automation of antitrust procedures and the improvement of antitrust analysis."[31] As the project highlights, using algorithms in antitrust enforcement is not only already a reality but also the future of effective enforcement in the context of constantly evolving markets. This section briefly documents algorithmic tools already deployed by antitrust agencies and then explains why the problem of algorithmic bias is less salient in the context of antitrust as compared to other legal fields.

A. Algorithms and Antitrust Enforcement: Existing Tools

Algorithms are already a reality in antitrust enforcement. The computational antitrust project ran an implementation survey in early 2022 in order to assess the current uses of computational tools by the participating antitrust agencies.[32] As this survey shows, antitrust agencies worldwide already rely on algorithmic tools for both procedural and substantive purposes. In terms of procedural aspects, such tools are used to (1) digitalize analog data collections through document management systems and (2) automatize procedural phases such as document submissions, leading to a significant speeding of antitrust investigations. When it comes to the substantive aspects of antitrust enforcement, the most widespread uses of computational tools are the following: (1) data mining and screening techniques for spotting markets and market structures that are most likely to facilitate collusion between market players; (2) information gathering algorithms for identifying price trends and their evolution over time; (3) ML algorithms for analyzing public procurement data in order to spot bid rigging practices.

The examples above testify that computational tools have become commonplace for competition authorities worldwide. Given their widespread uses, reflection on the potential biases of such tools is imperative.

B. Algorithmic Bias in Antitrust Enforcement

Within the landscape of all regulatory bodies, certain features of antitrust law single it out as particularly suited for the use of algorithmic tools. To begin with the most obvious, as opposed to bodies of law dealing directly with the rights and responsibilities of human subjects, the regulatory subject matter of antitrust laws are firms.[33] Consequently, risks of bias or discrimination operate on a different level. Instead of being biased towards certain population groups, antitrust enforcement can be biased towards certain types of firms based on their size or their origin, or towards

29 Xiang, *supra* note 23.

30 Adriano Koshiyama, Emre Kazim & Philip Treleaven, *Algorithm Auditing: Managing the Legal, Ethical, and Technological Risks of Artificial Intelligence, Machine Learning, and Associated Algorithms,* 55 Computer 40 (2022); Alfred Ng, *Can Auditing Eliminate Bias from Algorithms?*, The Markup (2021), https://themarkup.org/the-breakdown/2021/02/23/can-auditing-eliminate-bias-from-algorithms (last visited Sep 5, 2022); Pauline Kim, *Auditing Algorithms for Discrimination*, 166 University of Pennsylvania Law Review Online (2017), https://scholarship.law.upenn.edu/penn_law_review_online/vol166/iss1/10.

31 Stanford Law School, *Codex Project: Stanford Center for Legal Informatics*, Stanford Law School, https://law.stanford.edu/codex-the-stanford-center-for-legal-informatics/computational-antitrust-project/ (last visited May 30, 2023).

32 Thibault Schrepel & Teodora Groza, *The Adoption of Computational Antitrust by Agencies: 2021 Report*, 2 Stanford Computational Antitrust 79 (2022).

33 Edward Rock, *Corporate Law Through an Antitrust Lens*, 92 Colum. L. Rev. 497 (1992), https://scholarship.law.upenn.edu/faculty_scholarship/723.

achieving certain outcomes.[34] Whereas such biases have a detrimental impact on the capacity of antitrust to promote consumer welfare, they do not pose any threats to the fundamental interests of human subjects in the way in which biases in enforcing migration law, criminal law, and social security law do. Even in cases in which human biases do get passed on to algorithms tasked with enforcing antitrust laws, the effects are less consequential as compared to other fields of law.

Secondly, antitrust investigations are increasingly data-intensive. As markets grow more complex, understanding the impact of market behavior requires analyzing sizeable datasets. Even the first step of antitrust investigations, namely defining the relevant market, requires assessments of data points ranging from consumer preferences to transportation costs. More data translates into more precision. In the absence of exhaustive data, agencies need to rely on presumptions in order to filter out anticompetitive conduct. A case in point is the EU Digital Markets Act ("DMA"), which takes company size as a proxy for anticompetitive potential.[35] In this sense, the law is already biased against large platforms, and reliance on algorithmic tools that can inject more nuance into the antitrust analysis would be a move in the opposite direction, reducing bias instead of amplifying it. Furthermore, given the substantial corpus of antitrust decisions and cases, algorithmic tools can also be deployed to clarify existing law, harmonize decisional practice, and ultimately reduce the window of discretion of human decision-makers and thereby diminish the potential impact of their biases.[36]

Thirdly, antitrust laws are broad, open-textured provisions that have required courts to flash out their meaning.[37] This nature of antitrust laws has enabled the field to evolve organically and to adapt to changes in business dynamics. This is, however, a double-edged sword: Malleable laws are great for adapting to societal progress, yet this comes with costs for legal certainty and predictability.[38] Analyzing antitrust enforcement, Lim cautions against rejecting the use of algorithms on grounds of lack of transparency and potential bias. According to him, the existing rules and jurisprudence are themselves akin to a black box, leaving ample discretion for decision-makers to weigh costs, benefits, and counterfactuals.[39] Lim cites complaints from Chief Justice Roberts on the "amorphous [nature of the] rule of reason,"[40] and from Justice Breyer, who argues that implementing procompetitive benefits in the rule of reason analysis is an "absolute mystery."[41] The move from human decision-makers to algorithmic tools can lead to an increase in decisional transparency. As the previous section has shown, it is possible to develop ML tools that can weigh the costs and benefits of intervention and identify the welfare-maximizing outcome based purely on efficiency considerations.

Fourthly, antitrust rules are static instruments that seek to regulate increasingly dynamic markets. Antitrust agencies from the EU to the U.S. are increasingly sympathetic to switching from existing open-ended laws to ex-ante rules that render certain types of conduct per se illegal.[42] In the context of rapidly evolving markets, reliance on detailed regulatory instruments that contain absolute prohibitions is inherently biased toward the status quo and privileges certain market structures at the expense of others. Contrastingly, leveraging algorithmic tools represents a way to circumvent this bias, potentially enabling agencies to fine-tune existing legislation in order to make it more receptive to the dynamics of contemporary markets. This would represent the chance to move from a reactive approach to antitrust enforcement to an adaptive one, mindful of market developments. As an example, Pentland and his co-authors propose expanding the definition of monopoly power and of the premerger review process in order to take into account the data-intensive nature of contemporary markets.[43] Moving away from the static analysis of dominance/monopoly power based on market share thresholds as a proxy, the authors propose a fresh analysis factoring in the degree of "data control" of the entities at stake.

34 Anu Bradford, Robert Jackson & Jonathon Zytnick, *Is EU Merger Control Used for Protectionism? An Empirical Analysis,* 15 J. Empirical Legal Stud. 165 (2018), https://scholarship.law.columbia.edu/faculty_scholarship/2093.

35 Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), 265 OJ L (2022).

36 Kleinberg et al., *supra* note 28.

37 Daniel Crane, *Antitrust Antitextualism*, 96 Notre Dame L. Rev. 1205 (2021).

38 Teodora Groza et. al., *Exploring Computational Antitrust: A Theoretical Excursus,* Sciences po l. rev. (forthcoming).

39 Kleinberg et al., *supra* note 28.

40 *FTC v. Actavis, Inc.*, 570 U.S. 136, 160 (2013) (Roberts, C.J., dissenting).

41 Transcript of Oral Argument at 24, *Ohio v. Am. Express Co.*, 138 S. Ct. 2274 (2018) (No. 16-1454).

42 Aurelien Portuese, *American Precautionary Antitrust: Unrestrained FTC Rulemaking Authority*, Info. Tech. & Innovation Found. (2022).

43 Robert Mahari, Sandro Claudio Lera & Alex Pentland, *Time for a new antitrust era: refocusing antitrust law to invigorate competition in the 21st century*, 1 Stanford Computational Antitrust (2021).

IV. CONCLUSION

In this article, we seek to challenge the persistent assumption that the use of algorithms in law enforcement in general and in antitrust in particular is at insurmountable risk of bias and discriminatory outcomes. The potential benefits for antitrust enforcement should not be so readily dismissed: In certain respects, reliance on algorithmic tools can even diminish the risks of bias. Furthermore, antitrust law is particularly well suited for the use of algorithmic tools for several reasons: First, its regulatory targets are firms, not citizens; second, antitrust investigations are data-intensive; third, antitrust laws are notoriously open-ended, leaving outsize discretion to human decision-makers; fourth, antitrust needs to adapt to rapidly evolving market dynamics.

ALGORITHMIC PRICING AND COMPETITION

BY ROBERT CLARK & DANIEL ERSHOV[1]

1 Department of Economics, Queen's University and School of Management, University College London, respectively.

I. INTRODUCTION

In recent years, algorithms employed by companies to aid with price setting have become more sophisticated, evolving from rule-based to AI-driven. The latter can absorb huge quantities of data and autonomously engage in quick and accurate price determination. Crucially, AI-driven algorithms are able to experiment and learn from previous pricing decisions. The development of AI-driven algorithmic pricing has provoked fear regarding its potential impact on competition. The possibility that algorithms might facilitate collusion has been raised by competition-law experts, economic organizations, and antitrust authorities.[2] The main fear is that, by being able to better monitor and rapidly respond to rival price changes, algorithms might learn to implement collusive pricing strategies that would *reward* companies for following the collusive agreement and *punish* them for undercutting, giving rise to coordination on a joint-profit maximizing outcome.

In this article we summarize findings from our examination of the impact of adoption of algorithmic pricing.[3] This work (henceforth "Assad et al 2023") represents the first empirical analysis of the consequences of wide-scale adoption of algorithmic pricing software. Its focus is the German retail gasoline market, where trade publications suggest that AI-pricing became widely available in 2017. Algorithmic pricing is especially problematic in retail gasoline markets, where a number of countries, including Germany, France, and Spain, have introduced price-disclosure policies that provide the opportunity for algorithms to gather training data and to perfectly monitor prices. Our analysis leverages the data provided by the transparency regime to study the impact of adoption of algorithmic pricing on station margins, and to investigate whether the observed increases in margins can be linked to anti-competitive strategies.[4]

II. ALGORITHMIC COLLUSION

Developers of algorithmic pricing software provide little information regarding the nature of their algorithms. However, their promotional materials indicate that their software applies *artificial intelligence* and *machine learning*. They also explain that their algorithms have the capacity to incorporate information on a variety of variables, such as market conditions, prices (own and competitor), sales volumes, and costs. The algorithms begin by training on historical observations of these variables and then continuously absorb real-time data into decision making as they seek to maximize profits. Next the algorithms evaluate the sales outcomes arising from the chosen prices and plug this new information back into the algorithm as further inputs before the process begins again. Many of these algorithms will use reinforcement learning techniques that employ strategies that have performed effectively historically with higher probability in the future. Importantly, the software developers also state that a large majority of adopters allow the algorithms to operate autonomously, without human intervention.

If companies repeatedly interact, it may be possible for them to coordinate on elevated prices through a reward and punishment mechanism, whereby competitors that maintain high prices are rewarded, while those that undercut are punished. Whether or not such coordination is sustainable depends on the speed with which companies are able to detect competitor defection and to punish deviations. Algorithms may be able to facilitate collusion by helping companies monitor the pricing decisions of their rivals, and then quickly detect undercutting and punish defectors.[5] If algorithms speed up the ability of companies to react to defection, then the period during which higher profits are earned will be very short before prices collapse. As a result, companies will be less likely to defect and undermine the collusive arrangement.

An additional problem arises if multiple companies operating in the same market adopt the exact same algorithmic software. Then, *hub-and-spoke* arrangements may develop. An algorithm drawing in data from all the companies using it and providing this information to each of them, would be performing the role of *hub* facilitating coordination or collusion amongst its *spokes* – the adopters of the software. The collu-

2 See for instance OECD (2017), "Algorithms and Collusion: Competition Policy in the Digital Age"; United Kingdom Digital Competition Expert Panel (2019), "Unlocking Digital Competition"; Autorité de la Concurrence & Bundeskartellamt (2019), "Algorithms and Competition"; and Ezrachi, A. & M. Stucke (2016), "How Pricing Bots Could Form Cartels and Make Things More Expensive," *Harvard Business Review*, October 2016.

3 This article is based on our academic paper: Assad, S., R. Clark, D. Ershov & L. Xu (2023), "Algorithmic pricing and competition: Empirical evidence from the German retail gasoline market," forthcoming, *Journal of Political Economy*.

4 Legal disclaimer: This article analyses the impact of adoption of AP on competition strictly from an economic point of view. We base our understanding of the facts on publicly-available data on prices from the German Market Transparency Unit for Fuels. To our knowledge, there is no direct evidence of anticompetitive behavior on the part of any algorithmic-software firms or gasoline brands mentioned in this paper.

5 See Ezrachi, A. & M. Stucke (2015), "Artificial Intelligence and Collusion: When Computers Inhibit Competition," University of Tennessee, Legal Studies Research Paper Series #267, 2015; Mehra, S.K. (2015), "Antitrust and the Robo-Seller: Competition in the Time of Algorithms," *Minnesota Law Review*, 100.

sion-facilitation role of *third-parties* has been documented in the literature, pointing out that they may allow cartels to overcome difficulties related to coordination and information-sharing required for the successful functioning of cartels.[6]

Recent academic work has offered theoretical and experimental evidence that coordination may develop when pricing algorithms are employed.[7] A number of articles have demonstrated that reinforcement algorithms can, in time, learn to coordinate on high prices, by punishing deviations for finite numbers of periods. The literature has also noted that different algorithms, or combinations of different algorithms adopted by rival companies, may result in different outcomes and, as a result, can be either pro- or anti-competitive.[8] Therefore, there is no consensus at this time in the theoretical literature regarding whether coordination will emerge following adoption of algorithmic pricing. Furthermore, the extent to which algorithms may intensify or soften competition could be a function of the learning protocols they use.[9]

Empirical work investigating the effect of algorithmic pricing is therefore necessary, but faces three important challenges. The first is that adoption decisions are usually not publicly observed. The second is that adoption is endogenous, since adoption decisions are correlated with factors that may not be observed by researchers. The final challenge is that, should adoption be causally linked to higher prices or margins, it still would not be clear that the resulting changes can be assigned to collusion rather than to alternative causes, such as improved price discrimination.

To address the first challenge, Assad et al (2023) tests for structural breaks in pricing behaviours that are typically associated with AI-driven algorithmic pricing: (i) the number of daily price changes, (ii) the speed of response to competitors' price changes, (iii) the responsiveness of shocks to crude oil prices, and (iv) the responsiveness to local weather shocks. Leading software providers describe their algorithms as being capable of "rapidly, continuously and intelligently" reacting to the conditions of the market. For each of these metrics, Assad et al (2023) identifies structural breaks at each station. The best-candidate structural break in each metric for a given station is the week with the highest F-statistic. Since stations might break in a single metric for a variety of reasons, the paper classifies stations as algorithmic-pricing adopters if they feature best-candidate breaks in at least two of the four markers within a short time frame (four weeks). Approximately 20 percent of stations are found to experience best-candidate breaks in multiple metrics within a four-week window. Reassuringly, most breaks occur in the summer of 2017, when promotional materials by software providers suggest that commercial gasoline-retail algorithmic pricing tools became widely available in Germany.

The next step is to determine the impact of adoption on retail prices and margins; however, adoption decisions may be correlated with station/time specific unobservables, implying the existence of a selection bias and diverging outcomes between non-adopters and adopters before their adoption date. To overcome this problem, Assad et al (2023) adopts an instrumental variable approach, instrumenting for a station's adoption decision using the adoption decision by the station's brand (i.e. by brand headquarters). The idea is that brands may provide support (possibly in the form of subsidies) to stations to defray the costs of adoption. Such support has been offered in the past by brands to their affiliated stations when other technologies (e.g. electronic payment) have been introduced. Because brands' decisions to support adoption are not observed, Assad et al (2023) employs a proxy based on the share of a brand's stations that have adopted algorithmic pricing. Using this instrumental variable approach, findings suggest that adoption increases margins by roughly 15 percent. Reassuringly these estimates are in line with the profit gains claimed by software developers following adoption in other settings.

These findings offer causal evidence that adoption of algorithmic pricing increases margins, but not that the channel through which margins increase is through collusion. For this, Assad et al (2023) investigates whether the impact of adoption varies with market structure. The impact of adoption in monopoly and non-monopoly markets are compared. The effect of adoption should be stronger in non-monopoly markets if adoption influences competition. The findings suggest this to be the case. Assad et al (2023) finds a much stronger impact of adoption in non-monopoly markets than in monopoly markets.

To further confirm that adoption facilitates collusion, Assad et al (2023) focuses on small oligopoly markets – markets with just two or three stations. It investigates whether the impact of adoption varies depending on whether no stations adopted, a subset adopted, or all stations

6 See Clark, R., I. Horstmann & J-F. Houde (2023), "Hub-and-spoke collusion: Theory and evidence from the grocery industry"; Garrod, L., J. Harrington, and M. Olczak (2021), *Hub-and-Spoke Cartels: Why they form, how they operate, and how to prosecute them*, MIT Press.

7 See Calvano, E., G. Calzolari, V. Denicolo & S. Pastorello (2020), "Artificial Intelligence, Algorithmic Pricing and Collusion," *American Economic Review* 110, 3267-3297; Klein, T. (2021), "Autonomous Algorithmic Collusion: Q-Learning Under Sequential Pricing," *RAND Journal of Economics* 52, 538-558.

8 See Miklos-Thal, J. & C. Tucker (2019), "Collusion by Algorithm: Does Better Demand Prediction Facilitate Coordination Between Sellers?" *Management Science* 65, 1552-1561; Brown, Z & A. MacKay (2023), "Competition in Pricing Algorithms," *American Economic Journal: Microeconomics* 15, 109-156.

9 See Asker, J., C. Fershtman & A. Pakes (2021), "Artificial Intelligence and Pricing: The Impact of Algorithm Design," forthcoming *Journal of Economics and Management Strategy*.

adopted. The findings suggest that, relative to markets where none of the stations adopted, markets in which all stations adopted experienced margin increases of almost 40 percent. The size of the observed margin increases are in line with estimates from the literature studying the impact of coordination in retail gasoline markets.[10]

Assad et al (2023) also study the pricing behaviour that arises in markets where all stations adopt. It find that following adoption, a station is more likely to respond to a competitor's price decrease by immediately decreasing its own price. In contrast, there is no change in the likelihood of responding to a competitor's price increases. Such immediate punishment of price reductions by algorithmic rivals is somewhat analogous to anti-competitive trigger punishment strategies or price-matching strategies.

III. COMPETITION POLICY

The evidence in Assad et al (2023) is particular to retail gasoline markets in Germany, where high frequency pricing data are available, but algorithmic pricing software is being adopted in gasoline retail markets around the world. At a minimum, the results suggest that competition authorities in Germany and elsewhere should undertake a census of retail-gasoline pricing software to understand the market structure of the algorithmic software market and the extent of adoption. Such a census can help separate whether the main effect of algorithmic pricing software on competition comes from multiple stations in a market adopting the same or different algorithms. Which algorithm competitors adopt is not directly observed and the two possibilities have different implications for regulators and policy-makers.

Moreover, while the empirical focus of Assad et al (2023) was on the retail gasoline market, custom-made and "off-the-shelf" algorithmic pricing software is widely available to use for online and offline retailers in many markets. Adoption of such algorithms is growing: Brown and MacKay (2022) present evidence of algorithmic pricing by pharmaceutical drug retailers online. *Vendavo*, an AI based retail pricing software reports over 300 global deployments in manufacturing, chemicals, distribution, and high-tech industries (Vendavo.com).[11] In the U.S., more than 20 lawsuits have been launched in 2022 and 2023 against Realpage Inc., an algorithmic pricing software developer for multifamily residential properties (Reuters).[12] The plaintiffs claim that the software allowed owners to coordinate on higher prices. Another similar class action lawsuit was launched in January 2023 against Las Vegas hotel operators who allegedly used Rainmaker, a revenue management software, to coordinate on high prices (Bloomberg.com).[13] The language in the lawsuits is akin to hub-and-spoke collusion, with the algorithm serving as the coordinator.

The findings in Assad et al (2023) have potentially important implications for competition policy. Multiple antitrust authorities and economic organizations have released reports on algorithmic collusion and its implications for competition law. According to these reports, addressing situations where companies explicitly agree to collude via algorithmic pricing would not require any changes to existing competition laws. However, it is important to note that existing theoretical and empirical evidence suggests that collusion in this context is tacit. It develops without explicit communication between competitors / algorithms. Therefore, as currently defined, the unilateral adoption of algorithmic pricing software would not violate the competition statutes of most jurisdictions. When determining liability in instances of non-algorithmic collusion, competition authorities have concentrated their attention on explicit communication between competitors, instead of on the rules underpinning the collusive arrangement or on outcomes, such as elevated prices, that arise under collusion.[14] This is because it is challenging to determine with certainty that high prices are caused by collusive behavior, and because when humans engage in coordination, the collusive pricing rules are latent.

The UK Digital Competition Expert Panel states that with "further evidence … of pricing algorithms tacitly co-ordinating of their own accord, a change in the legal approach may become necessary" (p.110, 2019). What might a new approach look like? When prices are set using algorithms, there is no communication between companies, but consumers could still be confronted with elevated (collusive) prices. Some economists have suggested that, rather than targeting explicit communication, in the context of algorithm pricing it may be necessary for antitrust authorities to focus their attention on the collusive pricing rules written into the algorithms (see Harrington 2018 and Calvano et al 2020). Unlike

10 See Clark, R. & J-F. Houde (2013), "Collusion with asymmetric retailers: Evidence from a gasoline price-fixing case," *American Economic Journal: Microeconomics* 5, 97-123; Clark, R. & J-F. Houde (2014), "The Effect of Explicit Communication on pricing: Evidence from the Collapse of a Gasoline Cartel," *The Journal of Industrial Economics* 62, 191-228; Byrne, D. & N. de Roos (2019), "Learning to Coordinate: A Study in Retail Gasoline," *American Economic Review* 109, 591-619.

11 https://www.vendavo.com/customers.

12 https://www.reuters.com/legal/litigation/realpage-antitrust-lawsuits-over-rent-prices-consolidated-tennessee-2023-04-10/.

13 https://www.bloomberg.com/press-releases/2023-01-26/hagens-berman-las-vegas-hotel-operators-sued-for-alleged-scheme-to-illegally-inflate-hotel-room-rates-to-record-highs.

14 For more detailed discussion on these points see Harrington, J. (2018), "Developing Competition Law for Collusion by Autonomous Artificial Agent," *Journal of Competition Law & Economics* 14, 331-363; Calvano, E., G. Calzolari, V. Denicolo & J. Harrington (2020), "Protecting Consumers from High Prices due to AI," *Science* 370, 1040-1042.

with human-based collusion, pricing rules are not latent in the case of algorithms. Rather they are directly specified in the code. Should particular rules be identified that make collusion more likely, then authorities could insist that these be avoided by algorithms. This might be especially relevant for companies that have adopted the same algorithmic pricing software.

In Germany, the Federal Cartel Office ("Bundeskartellamt") is the competition authority charged with regulating and protecting competition.[15] Germany also has an independent advisory board, the Monopoly Commission ("Monopolkommission"), tasked with advising the German Federal Government on competition related issues.[16] The Monopolkommission's 2018 report on competition issues in Germany included a discussion on the issue of algorithms and collusion. The report states that further monitoring and evidence is needed to determine whether changes need to be made. If evidence does arise that algorithms lead to the further development of collusive markets, the report suggests that potential revisions could include (i) in cases of prohibited competitive behaviour, imposing the burden of proof that algorithmic usage has not contributed to infringement, and (ii) supplementation of the liability regimen under Article 101 of the Treaty of the Functioning of the European Union to include review and potential regulation of third parties (i.e. those that contribute IT expertise to algorithms) in cases of collusive behaviour.

Alternative potential approaches to mitigating the risk of collusive algorithms could include competition authorities or regulators testing algorithms of concern *ex ante* with simulated scenarios. These tests could help to identify the strategies that algorithms use and whether or not they are anti-competitive. For example, an algorithm that responds with a trigger strategy to competitor price decreases in order to maintain high prices would be flagged. Such ex ante tests could be useful complements to *ex post* investigations into potential anti-competitive conduct.

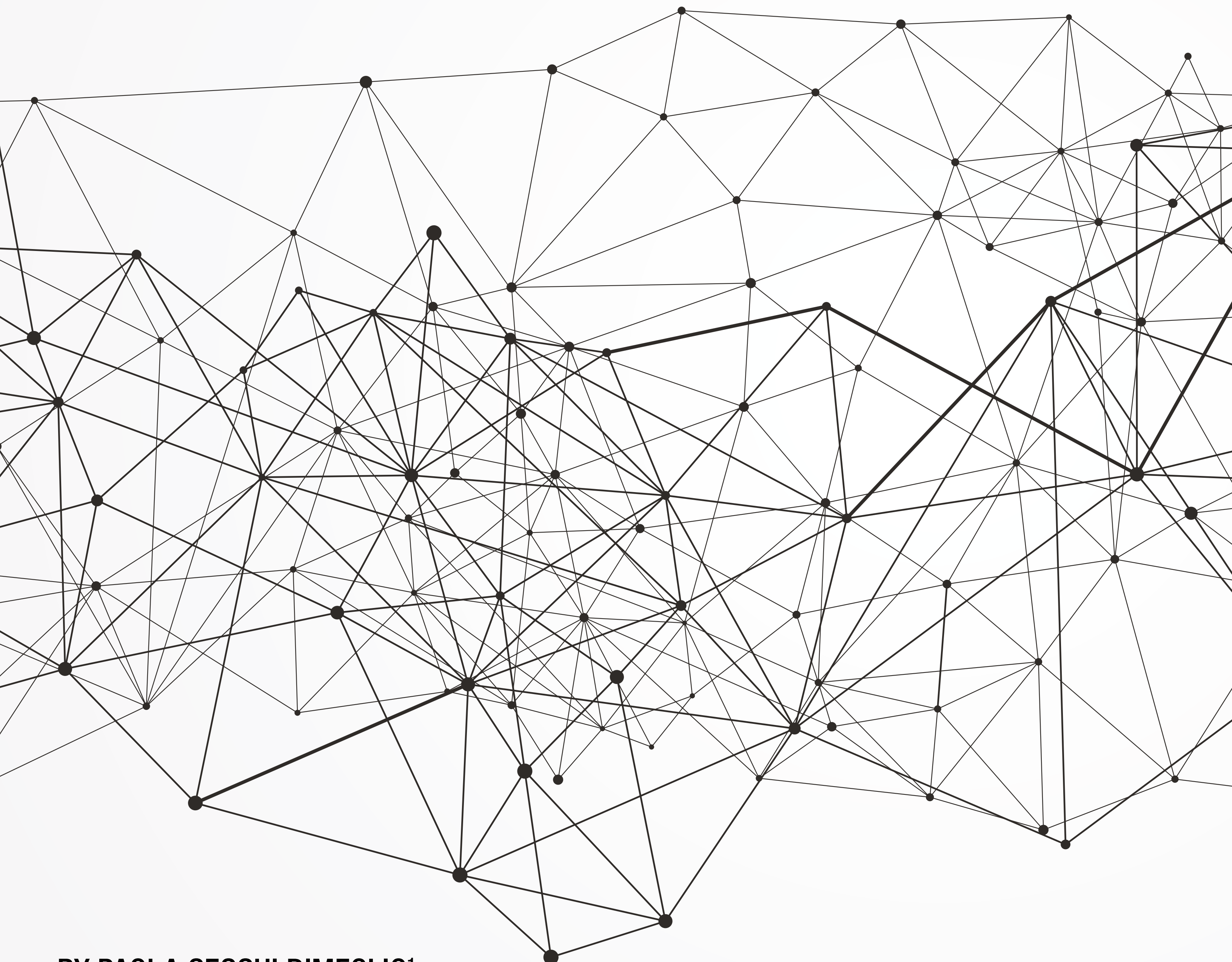
IV. CONCLUSION

A key challenge for policy makers is how to balance the fact that companies should be encouraged to make use of all available data when setting their prices, with the possibility that doing so may encourage tacit coordination or even collusive behavior and result in sustained periods of higher prices and margins. Algorithms can aid companies to incorporate a wider range of information and to more efficiently set prices. Determining which parts of a pricing algorithm's code take a company across the line of better data accumulation and information extraction into collusion facilitation is critical if we are to leverage the positives from this new technology without severely limiting competition.

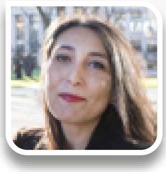
15 The Bunderskartellamt's tasks are carried out in accordance with legal provisions provided on both a national and European level. Relevant provisions regarding collusion are articles 101 and 102 of the Treaty on the Functioning of the European Union ("TFEU") and the Act against Restraints of Competition (Competition Act - GWB) (The Bundeskartellamt, 2020).

16 The tasks of the Monopolkommission are regulated by a number of legislative acts including sections 42(5) and 44 to 47 of the GWB (Monopolkommission, 2020).

CAN WE GET THE BIAS OUT OF OUR AI?



BY PAOLA CECCHI DIMEGLIO[1]



1 Paola Cecchi-Dimeglio, a lawyer, and behavioral and data scientist, is chair of the Executive Leadership Research Initiative for Women and Minority Attorneys at Harvard Law School and Harvard Kennedy School. She is also CEO of the People Culture Drive Consulting Group and author of Diversity Dividend, forthcoming from MIT Press. Email: pcecchidimeglio@law.harvard.edu

I. INTRODUCTION

Algorithm bias stacks the system against us, but we may never find out. We dread sentient, adversarial artificial intelligence ("AI"), but worse possibilities are already here. The algorithms that facilitate machine learning and drive our AI are discriminating. Meant to create automation and equity, algorithms are sometimes hurting employees and consumers and hurling businesses into lawsuits.

We are always choosing. It's complicated, time consuming, annoying. Algorithms are here to help. Right? They confront complexity and bestow time. Algorithms perform a good chunk of our lifting (light and heavy) and spare us the angst of selecting.[2] They inform governments, configure businesses, conduct wars, optimize diets, facilitate love connections, and accompany us wherever we go. But there's an ominous side.

We were not "today years old" when we learned to use algorithms. Wherever we have systemized behaviors, we have algorithms. We get through traffic, decide what to eat, even tie our shoes algorithmically. We've been spinning algorithms since childhood. During gym class, algorithms picked our teammates. We applied one algorithm on basketball day and another when the game was dodgeball. Different algorithms helped us compute the ideal lab partner. Our algorithms worked. Routinely feeding them race, gender, and cultural stereotypes (one set for gym, another for science) helped reinforce our biases.

Letting the algorithm choose made us fair. We used algorithms blindly, but we determined what they saw. All choosing may be biased, but the consequences resonate adversely for vulnerable individuals and groups. Some teammates and lab partners suspected our motivations. But just like individuals encountering algorithm bias today, most never knew a thing. And if they did, they couldn't prove it.

Right now, algorithms are guiding, deciding, and performing a host of functions. From manufacturing to social media, growing collections of processes are trusted to AI. Algorithms recognize and respond to voices, faces, images, names, conditions, and other data. But algorithms don't write themselves (unless aided by another algorithm).

We teach algorithms how to behave. In turn, they assess people and conditions with biases that might be discerned in person but go unnoticed in the digital zone. Most victims of algorithm bias never know that it happened. There is no body language, eye movement, vocal inflection; no bells, alarms, or alerts. And algorithm justice and watchdog groups are just taking their first steps.

II. WHAT IS ALGORITHM BIAS?

We cannot prevent what we cannot define. Algorithm bias seems a contradiction of terms. The algorithm was introduced as a dispassionate decision maker. When the means intended to remove discrimination begins perpetuating it, we have algorithm bias.

Some sources include the word "error" when defining algorithm bias, but this is misleading. "Error" suggests that something in the code is causing the problem, and code can be corrected. Algorithm bias doesn't work that way. Algorithms privilege and victimize different groups, and outcomes are glaringly unfair. But nothing in the original code directly states: if the applicant is female, and the open position is for a chemical engineer, deprioritize the applicant's resume. Yet somehow, bias enters.

As algorithms assume central roles in our AI, their biased behavior is receiving increasing attention. With more of what we do outsourced to AI, bits of our worst are being reactivated. We know about social media platforms allegedly rigging newsfeed algorithms to behave badly to garner clicks. But algorithms exhibiting bias or misbehaving in medical, law enforcement, legal, or other high-stakes settings can kill.

Biased algorithms are taking their place in education, banking, policing, healthcare, social media, dating, marriages, friendships, dieting, hiring, promotion, transportation, manufacture, insurance, investment, space exploration, and beyond. From birth to death, algorithms expose us to bias, and the old patterns are visible. Knowing how it can damage businesses and societies, we struggle to solve bias. Whether you are a corporate or a community, your number one asset is your people. Bias damages the victims and the perpetrators. It reduces organizations and societies. Enter AI.

Artificial intelligence promises to root out bias. But our technology is only as good as our raising, nurturing, and training of that technology. Informing our children that good people come in a variety of races, nationalities, ethnicities, religions, genders, sexualities, and abilities

2 https://blogs.thomsonreuters.com/legal-uk/2018/07/13/ask-dr-paola-algorithms-and-biases-in-recruitment/.

can be difficult for them to adapt if we only expose them to a homogenous selection of good people. Behaviorally, our children may be inclined to associate human goodness with a narrow set of identifying traits. Algorithm bias follows a similar pattern — good intensions undermined by limited examples. It takes more than we realize to prevent our technology from reflecting our history.

Breaches and cybercrime are only some of the perils posed by our tech. More people are exposed to algorithm bias than to computer viruses or bad actors roaming the digital space. Like malware, the timeline of algorithm bias is said to stretch back to the 1980s.[3] It has spawned sub-terminology like algorithm racism and algorithm sexism.[4] There is no shortage of examples. AI discriminates, and we know how.[5]

It's important to note that algorithm bias is all in the training. We encounter the AI, but bias is never the AI's fault. When a facial recognition algorithm that has been trained with more samples of white people's faces has difficulty recognizing a black person's face, we can't be surprised. Our biological eyes are engineered to recognize faces as such; the same cannot be assumed for our algorithms. Imagine if children from racially homogenous communities grew up incapable of recognizing the faces of other racial groups as faces. These shortcomings are just the beginning.

Businesses rely on AI to debias their hiring processes. Many interventions use algorithms that remove triggering data such as names, gender, schools, affiliations, volunteer activities, addresses, and other information. Surviving this initial phase, workers enter companies where other algorithms perpetuate biases that are rooted in race and gender. These biases result from how the algorithm is trained. Face recognition software develops a default bias based on the images used to train it; workplace algorithms learn from data that is skewed towards numerically dominant groups. These groups may be the majority within the organization, the profession, or the society. They may be predominant in a category where the business is trying to change outcomes. For example, the individuals promoted to partner at a professional services firm may be largely of one race and one gender. The algorithm can "see" this fact and may conclude it is a qualification. Algorithm bias can arise when desired outcomes (more women and minorities at the partner level) collide with the preponderance of data related to existing conditions.

III. HOW DO WE DETECT ALGORITHM BIAS?

We cannot delegate to algorithms and hope for the best. Blatant algorithm bias sometimes makes the news. A faulty facial recognition tool led to an innocent Black man's arrest in Detroit. Pedestrian recognition algorithms that have not "seen" enough examples of people with mobility differences cause vehicles to become disproportionately dangerous to some individuals with disabilities. And there is the notorious decision by ChatGPT that marked Australians as less preferred tenants and quickly went viral. But most of the discrimination carried out by algorithm-fed AI never gets on anyone's radar.

Everything is new, but novelty cannot excuse the bias. Ethical oversight of emerging and experimental tech calls for better "teaching" and testing of these consequential tools. New medications do not enter the market without extensive study and testing. After they are launched, sustained scrutiny and reporting is applied, and if the new drug is deemed harmful, it is pulled from the market. Algorithms have a similar broad impact, sometimes involving life and death. Rigorous testing could identify algorithm bias, before and after launch. One algorithm can have various outcomes for different groups and demographics. Bias might be detected if an algorithm is tested repeatedly by diverse parties.

Measuring the bias potential of an algorithm can be achieved through the application of processes that resemble the development of drugs and vaccines. The equivalent of clinical trials might involve highly diverse and inclusive groups of people engaging the algorithm during test phases. Comparing outcomes for these groups could identify and measure bias early in the design process. After they have been tested, adjusted, and released, algorithms must be monitored. An approach analogous to employee feedback and review could be applied to the algorithm to assess and modify the algorithm's evolving thinking and behavior. Afterall, the algorithm is a digital employee. Launching the algorithm and just leaving it to do its thing provides autonomy without oversight.

---

3 Robert P. Bartlett et al. "Algorithmic Discrimination and Input Accountability under the Civil Rights Acts." Available at SSRN 3674665 (2020); Joy Buolamwini and Timnit Gebru."Gender shades: Intersectional accuracy disparities in commercial gender classification." Proceedings of Machine Learning Research: 81:1-15, 2018.

4 Hardesty, Larry. "Study Finds Gender and Skin-Type Bias in Commercial Artificial-Intelligence Systems." MIT News, February 11, 2018. Available at http://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212 (last accessed April 1, 2023). These companies were selected because they provided gender classification features in their software and the code was publicly available for testing.

5 Hadhazy, Adam. "Biased Bots: Artificial-Intelligence Systems Echo Human Prejudices." Princeton University, April 18, 2017. Available at https://www.princeton.edu/news/2017/04/18/biased-bots-artificial-intelligence-systems-echo-human-prejudices (last accessed April 2, 2023).

IV. WHAT IS CAUSING BIAS?

In their intelligence, algorithms look for data with which to make decisions and build skills. Light or dark skin becomes part of the algorithm's definition of a face.[6] It delivers a Yes or No to the AI in a way that resonates with bygone ideas of who is human and seen and who is not. Did the algorithm arrive at this conclusion without assistance? Or was it supplied with default perspectives which allowed it to mimic long dismissed ideas?

Without exposure or practice, algorithms are deficient. When deployed into real situations, shortcomings become bias. In the workplace, small items of language like gendered pronouns can condition the AI as it seeks to build standards. Unless algorithms are fed diverse data, they are likely to digitize discriminatory default behaviors. Employment, housing, banking, credit and finance, education, and healthcare represent areas where algorithms have injected bias. When we look at the specific environments where algorithm bias crops up, we realized we are observing detectable, measurable phenomena. The biases attributed to algorithms are neither novel nor unique. Bias native to a given industry now plague that industry's algorithms and AI. Biases from banking distort banking algorithms.

A. Algorithm Bias in Employment

Without comprehensive input, algorithms trend toward old biases. Hiring and management algorithms promised to level the playing field for everyone, but encountering more males in engineering candidate pools teaches an algorithm to normalize male dominance in the profession and to use gender as a shortcut when making decisions.

B. Algorithm Bias in Housing

Algorithms across the housing sector help identify high-risk renters or guide builders and developers, but algorithms notoriously interact with complex datasets to make decisions that disadvantage black and brown people. CoreLogic, the company behind CrimSAFE, has been sued because its algorithm is alleged to disproportionately exclude Black and Latino applicants with criminal records.

C. Algorithm Bias in Banking, Credit, and Finance

Discrimination against people of color seeking mortgages and business loans is well documented. Businesses do not purposely use algorithms to perpetuate bias, but things go wrong. A UC Berkeley study showed that equally qualified Black and Brown borrowers pay financial institutions $765 million more per year than White borrowers.

D. Algorithm Bias in Education

When AI becomes the professor or grader, old biases enter. When making decisions about grades or incidents of plagiarism, algorithms show bias. A grade-estimating algorithm used to determine final grades for schools disrupted by COVID-19 favored students from elite private schools and those who were more socioeconomically advantaged.

E. Algorithm Bias in Healthcare

The racial wealth gap, cost of healthcare, and insurance coverage converge to make algorithms discriminatory. Organizations have labored to create diversity among their staff only to have algorithms regress their thinking by several decades. Algorithm bias causes providers to deliver better care to wealthier patients and make decisions that are disadvantageous to people of color.

V. ADDRESSING THE PROBLEM IN ITS INFANCY

AI is not responsible for fixing algorithm bias. Around the globe, many people grow from birth to old age having seen mostly or solely people of their own color or ethnicity. Yet they recognize other faces, and by extension, other humanity. The child or adult who does not recognize a face of different color as a human face might make good sci-fi. Unfortunately, it has become science reality.

AI does not have the data stores of history or genetics. As such, it will extrapolate from our other inputs (or lack of inputs). Algorithms can see whom we include and whom we exclude. It assumes that we like and favor the people who show up most frequently

6 Turner Lee, Nicol. Detecting racial bias in algorithms and machine learning. Journal of Information, Communication and Ethics in Society 2018, Vol. 16 Issue 3, pp. 252-260. Available at https://doi.org/10.1108/JICES-06-2018-0056/ (last accessed April 2, 2023).

on our guest lists. And that's fair. If it finds out that men are more often engineers in our world, it may build a shortcut and use it during hiring or promotion.

Blaming the algorithm transfers responsibility and perpetuates the problem. Wherever algorithms are biased, people did that. People are naturally reluctant to admit they are biased. Most business leaders will claim to have a network of colleagues and contacts that is diverse and inclusive. Close analysis usually reveals that their networks are considerably less diverse than they thought. When designing algorithms, people may believe they are inclusive, but the exposure to default standards teaches an algorithm to adopt behaviors from an untransformed world. It may not matter that the algorithm was devised by someone from a community that is the target of bias. Algorithms have to be directly taught that patterns and events in recent human history have configured the numbers in a certain manner, but it is the job of the algorithm to outperform the defaults and improve outcomes for underserved and underrepresented individuals. Our algorithms have to be reminded that both men and women are engineers. Even if you (the algorithm) see numbers that suggest it is more likely for men to occupy this role, don't use gender as a decision shortcut. Humans must be smart enough to create AI that can recognize and neutralize our bias. We need our algorithms to outperform our defaults.

Fortunately, algorithm bias is still in its infancy. As businesses rollback diversity, equity, and inclusion ("DEI") efforts, it may be left to algorithms to establish parity and close the gaps in the workplace and everyday life. These algorithms must be free of the default biases that we used to design organizations and societies.

Algorithm bias can be fixed. Left unchecked, many of our AI tools will evolve behaviors and methods that profile individuals and groups. Dialogue and innovation hold significant promise when it comes to correcting this dangerous trend. As social media and the world itself takes its first steps in the metaverse, algorithms find new prominence. Managing and eliminating algorithm bias in the virtual, augmented, and mixed realities of the metaverse may be critical to overall success. Algorithm bias will trigger multiple violations of civil rights and other legislation. This issue can be fixed proactively by businesses or addressed through the legal system.

VI. BEST PRACTICES

Left on their own to reach for data and formulate norms, algorithms routinely slip into bias. A series of best practices can help organizations and decision makers get out ahead.

1. Remain open to the possibilities of bias in the algorithms, understand how it might arise, and respond to algorithm bias in a timely manner.
2. Develop standards for determining if algorithm bias is present and disseminate methods of detection.
3. Devise algorithm bias response processes that stipulate what is to be done for a given type of bias incident. Consistently adhere to established protocols.
4. Use diverse teams to formulate, teach, and test algorithms. Let inclusive input and exhaustive testing identify discriminatory tendencies early on.
5. Build and continuously add to a set of terms or criteria that testers can use to determine if an algorithm is biased. Customize terms and criteria for different groups.
6. Ensure that design within algorithms create and sustain equity for all groups.
7. Expose algorithms to exhaustively inclusive datasets as part of a standard development process. To sustain fairness, algorithms must be more diverse than most companies.
8. Among all staff, create a level of psychological safety that lets these individuals feel comfortable voicing their observations regarding the algorithms.
9. Continuously monitor and review the AI as one would an employee. Modify the algorithm based on feedback.
10. Help educate customers regarding the possibility of algorithm bias. Create channels through which they can offer feedback, similar to customer service reporting or a survey tool.
11. Be transparent about how your algorithms work or are supposed to work so that users can begin to assess neutrality or bias in the algorithm and AI.
12. Set a standard minimum level of competence for each algorithm, treating the algorithm like a worker. Determine what is the least level of performance, what are the indicators, and how often will the algorithm undergo a performance review?

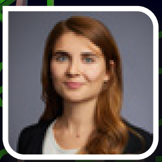
As we inadvertently impart active and atavistic discriminatory practices to our AI, algorithm bias shows us who and how we are (and were). Fixing the problem may require remembering all the ways we introduce bias in the first place. We want to win, and that drive can make bias attractive.

We know why we picked a certain person for our basketball team but never considered her for lab partner. It may have turned out that she aced the course. The classmate with the strong accent turned out to be a champion speller. Our biases fail and we are left to rethink. We can let our algorithms know that bias happens, that it is not desired, and teach them not to slip into bias. Screening and retraining algorithms can help a business prevent bias and avoid class actions and other consequences.

People don't question algorithms. It may be in the interest of businesses to prompt consumers to be aware and to offer feedback. As instances of algorithm bias begin to take up more space in news and media, the litigation will follow. The justice department is already settling cases of algorithm bias, and it's only a matter of time before the civil rights implications reach the courts and disrupt vulnerable businesses.

CAN SELF-PREFERENCING ALGORITHMS BE PRO-COMPETITIVE?

BY EMILIE FEYLER & VERONICA POSTAL[1]

1 Emilie Feyler and Dr. Veronica Postal are Consultants in NERA Economic Consulting's Antitrust Practice in White Plains, NY. The opinions expressed are those of the authors and do not necessarily reflect the views of NERA Economic Consulting or other NERA experts.

The role of algorithms and artificial intelligence ("AI") with respect to people's consumption choices and everyday decision-making has been growing hand-in-hand with the size of the digital economy. For example, about 80 percent of the content streamed on Netflix is the result of algorithmic recommendation, while only 20 percent is streamed through active user search.[2] The public launch of ChatGPT in November 2022 has pushed the boundary of what people believed AI systems could achieve further than ever. While the advances in deep learning technologies and their application to a wide range of industries opens numerous opportunities, the use of such algorithms has also raised a number of concerns with respect to ethics, privacy, security, and bias.

The notion of "algorithmic bias" refers to errors in a computer system that create "unfair" outcomes, for instance, by privileging one group of users over another. Algorithmic bias can emerge for several reasons, including a malfunction in the design of the algorithm, or the reliance on incomplete or non-representative data. Ultimately, algorithms are designed by humans, and thus their architecture involves some measure of subjective judgment with respect to their parameters and underlying assumptions. The risk of unintentional bias has been exacerbated by the development of AI and semi-supervised learning methods that allow for self-training algorithms without requiring inputs from the algorithm creator. Overall, the concerns relate to the lack of transparency of "black box" algorithms, and the fear that their users (and sometimes their own creators) do not truly understand how they work and how results are generated. Recently, a group of AI experts has urged to "immediately pause for at least 6 months the training of AI systems," to better understand AI's repercussions on the economy and society at large.[3]

In this article, we focus on algorithms that are allegedly "biased" by their creator to privilege their own content and products, the so-called "self-preferencing" algorithms. For example, according to a recent investigation by the European Commission ("EC"), Google Search's algorithm provided preferential treatment for the firm's own comparison-shopping service, Google Shopping, over rivals' comparison-shopping services (e.g. Amazon and eBay).[4] There has been active debate regarding the competitive implications of such algorithmic self-preferencing practices in digital markets, where large platforms often have a dual role: they operate the digital marketplace as information intermediaries and are also players of the marketplace, offering their own competing products.

Competition authorities around the world are increasing their scrutiny of digital markets and purported instances of self-preferencing. The European Union's Digital Markets Act ("DMA"), which went into effect in November 2022, explicitly prohibits self-preferencing practices for digital platforms, and proposed legislation in the United States, such as the American Innovation and Choice Online Act, would impose similar restrictions, although for the time being, no such bills have been passed in the U.S. Congress. However, the economic literature is not conclusive about whether self-preferencing algorithms are inherently harmful to consumers and competition. As discussed in this article, such algorithms can have pro-competitive benefits. Understanding whether pro-competitive benefits or potential anti-competitive considerations prevail requires careful analysis conducted on a case-by-case basis.

I. ANTITRUST IMPLICATIONS OF SELF-PREFERENCING ALGORITHMS

From an economic perspective, self-preferencing practices are not something new. Companies have been favoring their own products and services in downstream markets over their rivals' in disparate industries for a long time. For instance, in the retail sector, it is a common business practice for supermarkets to advantage the shelf display of their private labels products over competitors' brands. Recommendation algorithms can be considered the digital equivalent of such practices if they privilege the product of the owner of the platform over competitors' products. As such, the competitive implications of self-preferencing practices in digital markets can be explained using standard economic mechanisms. Indeed, the pro- and anti-competitive considerations of recommendation algorithms are similar to those of steering and tying practices.[5] On one hand, steering and tying practices can increase convenience for consumers, product quality, and incentives for innovation. On the other hand, they can allow firms to leverage their dominant position in one market to acquire monopoly power in a second market, or they could be used as an exclusionary tool against competitors or new entrants. Courts have generally adopted a "rule of reason" approach to traditional tying cases, weighing the pro-competitive benefits against the potential anti-competitive considerations, if any, on a case-by-case basis.[6]

2 C. A. Gomez-Uribe & N. Hunt, *The Netflix Recommender System: Algorithms, Business Value, and Innovation.* 6 (4) ACM TRANSACTIONS ON MANAGEMENT INFORMATION SYSTEMS, 1–19 (2016).

3 *Pause Giant AI Experiments: An Open Letter*, FUTURE OF LIFE INSTITUTE (March 22, 2023), https://futureoflife.org/open-letter/pause-giant-ai-experiments/.

4 Council Regulation (EC) 1/2003, Case AT.39740, Google Search (Shopping), 2017 O.J., available at https://ec.europa.eu/competition/antitrust/cases/dec_docs/39740/39740_14996_3.pdf.

5 Sheng Li, Claire Chunying Xie & Emilie Feyler, *Algorithms & Antitrust: An Overview of EU and National Case Law*, Oct. 7, 2021, CONCURRENCES ANTITRUST CASE LAWS E-BULLETIN, Art. N° 102334.

6 See the FTC's guidance on tied sales: *Single Firm Conduct: Tying the Sale of Two Products*, FEDERAL TRADE COMMISSION, https://www.ftc.gov/advice-guidance/competition-guidance/guide-antitrust-laws/single-firm-conduct/tying-sale-two-products (last visited [May 9, 2023]). See also the EC's Guidelines on Vertical Restraints, *Guidelines on Vertical Restraints*, EUROPEAN COMMISSION 61, available at https://ec.europa.eu/competition/antitrust/legislation/guidelines_vertical_en.pdf.

The traditional antitrust concerns towards tying and steering practices have found a new dimension in the context of digital platforms and recommendation algorithms. We identify two main reasons why the antitrust analysis of self-preferencing algorithms is more complex than traditional cases: self-preferencing algorithms can be arduous to detect, and they can have unconscious influence over users' choices.

As a preliminary issue, it is difficult to assess the extent to which self-preferencing algorithms exist on a digital platform. One needs to investigate the underlying complexities of the algorithm at issue to know whether specific recommendations are the result of steering practices as opposed to an objective ranking of products based on consumers' preferences, products' quality, or prices. Therefore, algorithms that unfairly advantage certain products may be harder to detect for competition authorities, users, or competitors than traditional self-preferencing practices in brick-and-mortar businesses.

In addition, algorithms can represent a "black box" to users, who may not understand how recommendations are generated and whether they stem from past behavior, prices, product quality, or other factors. Researchers and policymakers have raised concerns that digital platforms could take advantage of the opacity of these recommendation algorithms to predominantly display their own products and services, even though consumers may prefer other, cheaper, or higher-quality products.[7] In other words, the concern is that consumers, now more than ever before, could be influenced by such algorithms without full cognitive awareness, and may make decisions that are inconsistent with their preferences.

Regulators and courts around the world have released a number of decisions on cases involving algorithmic self-preferencing practices. As described below, courts and competition authorities are deviating from the traditional "rule of reason" approach and are moving towards the full prohibition of self-preferencing practices for digital platforms.

II. LEGISLATION OF ALGORITHMIC SELF-PREFERENCING

The wave of increased scrutiny of self-preferencing algorithms started in the early 2010s, when competition authorities in the United States and Europe started investigating whether Google biased its search results by allegedly promoting its own content and selectively demoting competitors' content. Despite conflicting findings of the U.S. Federal Trade Commission ("FTC") and of the EC on whether the alleged self-preferencing constituted an illegal abuse of market power, antitrust authorities around the world initiated a wave of investigations into similar practices by other technology companies.[8]

In the last few years, several countries adopted landmark legislation aimed at prohibiting self-preferencing for digital platforms, among other measures intended to enhance the protection of competition in the digital economy. A January 2021 amendment to the German Act against Restraints of Competition overhauled German competition law, allowing the German Federal Cartel Office ("FCO") to prohibit certain conduct (including self-preferencing) by digital platforms when a company's market position is found to be of "paramount significance across markets." By June 2021, the FCO had opened an investigation against Apple, Facebook, Amazon, and Google, and in 2022, it issued a ground-breaking decision declaring that Alphabet Inc., Google's parent company, and Meta, Facebook's parent company, are of "paramount significance for competition across markets" in Germany, and thus may prohibit these companies from engaging in purported self-preferencing behavior for a period of five years.[9]

The European Union followed suit with the DMA in November 2022, a wide-ranging piece of legislation aimed at regulating "unfair" business practices by large online platforms designated as gatekeepers between European businesses and consumers.[10] The DMA explicitly bans self-preferencing, stating that "the gatekeeper shall not treat more favourably, in ranking and related indexing and crawling, services and products offered by the gatekeeper itself than similar services or products of a third party."[11]

7 See the CMA's discussion on algorithms and competition: CMA, Algorithms: How They Can Reduce Competition and Harm Consumers (Jan. 19, 2021), https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers/algorithms-how-they-can-reduce-competition-and-harm-consumers#theories-of-harm.

8 FTC, *Statement of the Federal Trade Commission Regarding Google's Search Practices in the Matter of Google Inc.*, FTC File Number 111-0163, pp. 3-4 (2013); Press Release, EC, Antitrust: Commission Fines Google €2.42 Billion for Abusing Dominance as Search Engine by Giving Illegal Advantage to Own Comparison Shopping Service (June 27, 2017), available at https://ec.europa.eu/commission/presscorner/detail/en/IP_17_1784.

9 Andrea Pomana, *Germany's Google Controls Illustrate Global Antitrust Trend*, LAW360 (Jan. 21, 2022), https://www.law360.com/articles/1457160/germany-s-google-controls-illustrate-global-antitrust-trend; *German Watchdog Probes Apple's Market Dominance*, BBC NEWS (June 21, 2021), https://www.bbc.com/news/technology-57555008; *New Rules Apply to Meta (Formerly Facebook) – Bundeskartellamt Determines its "Paramount Significance for Competition Across Markets*," BUNDESKARTELLAMT (May 4, 2022), https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2022/04_05_2022_Facebook_19a.html.

10 *Digital Markets Act: rules for digital gatekeepers to ensure open markets enter into force*, EUROPEAN COMMISSION, (Oct. 31, 2022), https://ec.europa.eu/commission/presscorner/detail/en/ip_22_6423.

11 *Regulation (EU) 2022/1925 of the European Parliament and of the Council*, OFFICIAL JOURNAL OF THE EUROPEAN UNION (Sept. 14, 2022), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R1925.

U.S. lawmakers have also considered putting legislation in place to regulate self-preferencing by platform operators several times in recent years. An initial attempt at prohibiting self-preferencing by digital platforms was made in August 2021 with the Open App Markets Act, but the bill was ultimately not passed by the U.S. Congress. Another attempt was made in October 2021 with the introduction of the American Innovation and Choice Online Act, a bill focusing on regulating big tech companies and limiting self-preferencing by platform operators.[12] Neither bill was enacted and no further bills on this issue have been introduced. A July 2022 bipartisan report issued by the U.S. House Committee on the Judiciary on competition in digital markets called for regulation of various practices adopted by technology platforms, including the use of purported "self-preferencing" in algorithms.[13]

III. INSIGHTS FROM THE ECONOMIC LITERATURE

Although the rules imposed by competition authorities regarding self-preferencing algorithms are increasingly stringent, few economic studies have been conducted to assess whether such algorithms are indeed harmful to consumers, social welfare, and competition, or whether the pro-competitive benefits of such practices may outweigh the potential anti-competitive considerations. The findings of this literature have been ambiguous, suggesting that self-preferencing algorithms' effects on competition vary on a case-by-case basis.[14]

Some theoretical papers have shown that self-preferencing behavior by digital platforms does not necessarily harm competition. De Cornière & Taylor (2014) find that self-preferencing behavior by search engines for advertising may not be harmful to consumers and may in fact provide better content to consumers by reducing the nuisance costs due to excessive advertising.[15] Their model predicts that the increased revenue through sponsored ads could enable the search engine to reduce the number of advertisements displayed and therefore increase users' utility level. In a more recent article, the authors find that self-preferencing behavior by information intermediaries may benefit consumers in certain market environments, for instance, if firms compete on quality.[16] Zennyo (2022) puts forth an economic model showing that self-preferencing behavior by digital platforms can benefit consumers through lowered commission fees to third-party sellers, allowing them to decrease their prices, and in turn attracting more consumers and more third-party sellers onto the platform.[17] The economic intuition is that algorithmic self-preferencing enables the platform to sell its own products more effectively, which increases its expected profit per consumer and therefore increases its incentives to attract more consumers. To this end, the platform has economic incentives to reduce commission fees to make the resulting consumer prices lower, increasing consumer participation. The increase in consumer participation in turn stimulates third-party seller participation.

The economic literature has also investigated the effectiveness of policy intervention against self-preferencing algorithms, including behavioral and structural remedies, but existing studies have yielded ambiguous results. Hagiu, *et al*. (2022) claim that banning the platform's marketplace (i.e. preventing the platform to sell its own products) would likely result in lower consumer surplus and lower social welfare.[18] Zennyo (2022) finds that the separation of the marketplace and reseller divisions, which also prohibits the platform from self-preferencing, may be detrimental to consumers and third-party sellers. In this model, structural separation leads the platform to raise commission fees to third-party sellers, which in turn results in higher consumers' prices and lower consumer surplus. According to Kittaka & Sato (2022), while the prohibition of self-preferencing practices of dual-role platforms may have adverse effects, structural separation could improve consumer surplus through lower prices.[19] De Cornière & Taylor (2019) find that the efficiency of various policy interventions, such as imposing recommendation neutrality, transparency policies, or structural separation, depends on the market environment.

---

12 Open App Markets Act, S.2710, 117th Cong. (2021), *available at* https://www.congress.gov/bill/117th-congress/senate-bill/2710/; American Innovation and Choice Online Act, S. 2992, 117th Cong. (2021), available at https://www.congress.gov/bill/117th-congress/senate-bill/2992/text.

13 Subcommittee on Antitrust, Commercial, and Administrative Law of the Committee on the Judiciary of the House of Representatives, 117th Congress 2d Sess., Investigation of Competition in Digital Markets (Comm. Print 2022), https://www.govinfo.gov/content/pkg/CPRT-117HPRT47832/pdf/CPRT-117HPRT47832.pdf.

14 See Kittaka et al. (2023) for a detailed literature review. Yuta Kittaka, Susumu Sato & Yusuke Zennyo, *Self-Preferencing by Platforms: A Literature Review*, 66 JAPAN AND THE WORLD ECONOMY, 101191 (2023).

15 Alexandre de Cornière & Greg Taylor, *Integration and Search Engine Bias*, 45 (3) RAND JOURNAL OF ECONOMICS, 576–597 (2014).

16 Alexandre de Cornière & Greg Taylor, *A Model of Biased Intermediation*, 50 (4) RAND JOURNAL OF ECONOMICS, 854–882 (2019).

17 Yusuke Zennyo, *Platform Encroachment and Own-Content Bias*, 70 (3) JOURNAL OF INDUSTRIAL ECONOMICS, 684–710 (2022).

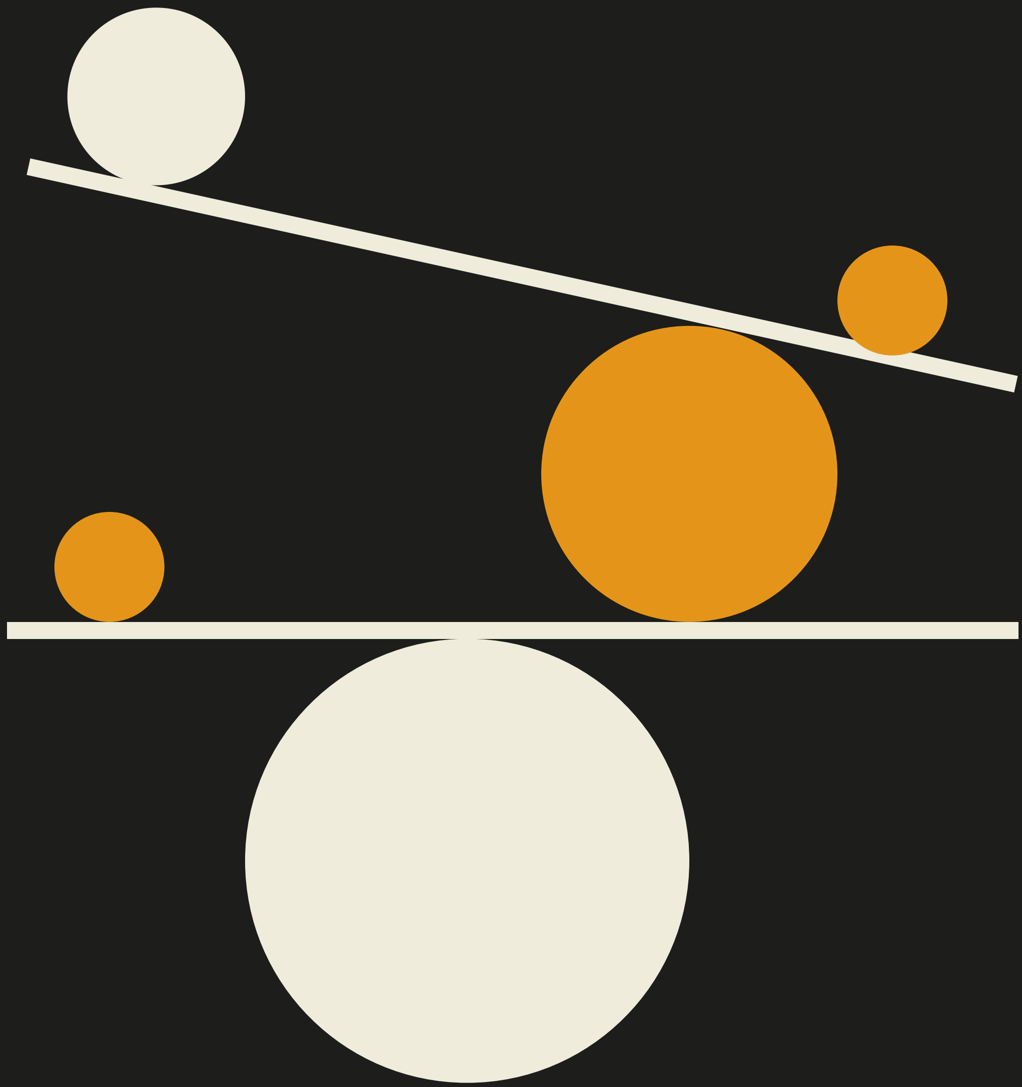
18 Andrei Hagiu, Tat-How Teh & Julian Wright, *Should Platforms Be Allowed to Sell on Their Own Marketplaces?*, 53 (2) RAND JOURNAL OF ECONOMICS, 297–327 (2022).

19 Yuta Kittaka & Susumu Sato, *Dual-Role Platforms and Self-Preferencing: Sequential Search Approach*, SSRN (Oct. 13, 2022), https://ssrn.com/abstract=3736574.

IV. CONCLUSION

In response to the growing concerns around algorithms and their influence over consumers' choices, courts and competition authorities have been deviating from the traditional "rule of reason" approach to adopt more stringent rules for digital platforms regarding self-preferencing practices. However, from a theoretical perspective, self-preferencing algorithms can have pro-competitive benefits. There is no consensus from the economic literature on whether pro-competitive benefits or possible anti-competitive considerations prevail in the context of self-preferencing algorithms used by digital platforms. Nor is there consensus on the welfare effects of policy intervention aimed at correcting bias in algorithmic recommendations. Determining the net impact of self-preferencing algorithms on competition and consumer welfare requires individualized analysis accounting for the workings of specific algorithms, competitive context, and market environment.

FAIRNESS IN ALGORITHMIC DECISION MAKING



BY SAMPATH KANNAN[1]

1 Sampath Kannan is the Henry Salvatori Professor at the Computer and Information Science Department, University of Pennsylvania.

Machine learning algorithms are being used to classify individuals in increasingly consequential situations. Such algorithms are used for example, in hiring, college admissions, bank loan decisions, and in the criminal justice system for predictive policing and predicting recidivism risk. It has been observed that while algorithmic decision-making may appear *prima facie* to be more objective than human decision-making, it inherits some of the problems of human decision-making, and presents new ones.

How do machine learning algorithms make decisions? What are their desiderata and what yardsticks do we use to measure whether they are met? What new challenges do algorithms present, compared to human decision-makers? What failings in human decision-making do they ameliorate? These are some of the questions we will examine in this brief article.

While machine learning comes in many forms, most deployed systems to date use a form of learning called *supervised learning*. Taking bank loan applications as our running example, in supervised learning, the algorithm is given many examples of individuals who applied for a bank loan along with the decisions made by a human decision-maker in each case. An application might have fields containing demographic information, as well as financial data such as income, assets, obligations, etc. For the algorithm, an applicant is just this *vector of features*, where each feature is a quantity capturing one of these items of information. The decisions by the human decision-maker --- YES / NO, coded as 1 / 0, are the *labels* associated with these examples. If the human decision maker gives more informative scores to each applicant, the learning algorithm could presumably also use this. Supervised learning is characterized by the fact that the algorithm is given such labeled examples, and must then figure out a "rule" to apply to future applicants.

This rule has to be "simple": If it follows all the "idiosyncratic complexities" of the human decision maker, it is merely "memorizing," not generalizing, and therefore not learning. This statement can be made mathematically precise. A consequence of this requirement of simplicity is that machine learning algorithms invariably make mistakes, unless the classification task at hand is inherently simple.

While classification errors are inevitable, one could ask that these mistakes do not result in discrimination against certain individuals, and systemic discrimination against groups based on gender, race, or other protected attributes. To reason about this, one needs mathematical definition(s) of what it means for an algorithm to be fair, and for it to discriminate. It is beyond the scope of this short article to even mention the many intellectual strands from philosophy, legal theory, political science, and other areas that inform the kinds of definitions that have been made. We simply describe some of the most common definitions in use. Notions of *individual fairness* seek to prevent discrimination against individuals, while notions of *group fairness* seek to ensure that protected groups (as defined in law) are treated similarly under different statistical notions of parity.

The simplest notion of group fairness is *demographic parity*. Is the same fraction of people in each protected group given a benefit or caused harm by a decision? Is each group "represented" in proportion to its share of the population? Demographic parity might not be appropriate if different groups have different qualifications or interests. In the bank loan example, the algorithm classifies people as positive, i.e. worthy of a loan, or negative, i.e. not worthy of a loan. The *false negative* rate for a group is the fraction of individuals in that group who are classified as negative, even though they are truly loan worthy (or positive). A natural goal is to equalize the false negative rate amongst groups. This accords with the well-studied principle of "equality of opportunity," since a positive classification can be seen as an opportunity. We might go further and ask that both the false negative rate, and the false positive rate (which has a similar definition) be equalized across groups. This stronger notion of fairness is referred to as "equalized odds."

In order to apply these notions of fairness, we need to know who the true positives and true negatives are in each group. This typically may come from training data where human experts have labeled people as positive or negative. Since in this setting we take the labels assigned by humans as the ground truth, a classifier that exactly mimics this human labeling would have 0 false positive and false negative rates on all groups, and thus achieve equalized odds. However, recall that a good machine learning algorithm has to produce simple classifiers, in order to be useful in classifying new data. Hence it will invariably not mimic these labels exactly and there will be non-zero false positive and negative rates. These metrics can be criticized because they regard the training data as ground truth. However, this is unavoidable, because all machine learning involves learning from data, and algorithms designed in this way can only be as good as the data on which they are trained. These metrics are in line with meritocratic, equity-based notions of distributive justice.

There is also another class of metrics that are used to judge the fairness of machine learning algorithms. To understand them, let's think of an algorithm that predicts the weather each day for a period of time. Let's be even more specific and say that for some fixed location, say New York, on each day, it predicts the probability of rain the next day. How do we judge the quality of such an algorithm? Even if it says that there is 90 percent chance of rain, and it does not rain, can't the algorithm defend itself because it left open some possibility of no rain? A good metric for this scenario is *calibration*, which captures the idea that the numerical probabilities that the algorithm outputs should mean something, just as the numerical reading of length from a ruler means something. An algorithm such as this one is said to be calibrated, if it rained exactly on 20 percent of the days on which it predicted a 20 percent chance of rain. (Of course, it must similarly satisfy the condition that it rained exactly on 30 percent of the days on which it predicted a 30 percent chance of rain, etc.) Calibration can also be used as a fairness notion.

Turning to a scenario where humans are being assigned a score by an algorithm, we can use calibration as a group fairness notion by requiring an algorithm to be calibrated separately on each protected group. Thus, for example, among the people in one group for whom the algorithm rated loan-worthiness as 80 percent, exactly 80 percent should be loan-worthy. Similar to calibration, there are notions known as positive predictive value and negative predictive value, which can be applied to algorithms that simply classify people and don't assign them scores.

None of these notions of fairness deal with corrective justice, and in fact, not many mathematical definitions capture this important notion. While calibration for each group, and equalized odds all sound like good notions of group fairness, unfortunately, except under the most perfect of circumstances with respect to how the groups are endowed with qualifications. Thus, in designing these algorithms we are forced to choose whether they will achieve equalized odds or be calibrated.

The argument about whether equality of opportunity/equalized odds or calibration is the better measure of fairness is not purely academic. The company Northpointe produced a recidivism prediction software package called COMPAS that critics at ProPublica found not to meet the fairness goal of equal false positive and false negative rates between the Caucasian and African-American populations it made predictions on. In their defense, Northpointe pointed out that the predictions for the same data were calibrated for both groups.

Let us briefly examine some arguments in favor of each of these notions of fairness. An algorithm that "really understands" the members of a group can hope to be calibrated by using that understanding in its predictions, just as a really good weather prediction model can hope to be calibrated in its task. However, one can be calibrated with just rudimentary statistical knowledge about a group. For example, when predicting the weather in Phoenix for the month of October, suppose we know from years of observation that it rains 5 percent of the time in October. Then we could predict a 5 percent chance of rain every day without having too accurate a weather model for the actual probability of rain each day, and we would probably be pretty close to calibrated. On the other hand, if we knew the weather patterns in New York intimately, we could make more varied and informative predictions for each day, and be calibrated as well. Thus, being calibrated may not tell us anything about how accurate predictions are to subsets of individuals in each group.

As a point in favor of calibration, it has been argued that an algorithm that best uses all the information it has at hand to produce the most accurate classification, will produce a calibrated classifier, and therefore this is the right kind of classifier to produce.

A system where the people being classified adjust their behavior and the data they generate based on how the classifier works is called an endogenous system. For example, imagine that individuals in different groups choose to commit a crime or not based on a risk/benefit analysis, which in turn depends on the probability with which the classifier will find a person guilty when they commit a crime and the probability with which it will find that person guilty when they do not. An individual's incentive to commit a crime is governed by the difference in these probabilities --- the larger the difference, the more dissuaded s/he is, to commit a crime. As the analysis shows, an adjudicator wanting to minimize the number of crimes overall should equalize the false positive and false negative rates across, even though equalizing calibration leads to a more accurate classifier! Here accuracy is less important than incentivizing individuals away from crime. This point naturally leads to the second part of the paper where we ask what machine learning systems could learn from more traditional systems.

In cases tried in courts there are rules of evidence that vary across jurisdictions and across types of cases that restrict what kinds of evidence may be presented. Should there be a similar restriction in the kind of data a machine learning algorithm is allowed to use? No such restrictions exist on today's learning algorithms. The justification for this might be the thinking that more data is never bad – If some features are uncorrelated or even negatively correlated with the desired output of the algorithm, the algorithm will learn this and disregard or lower the weight it gives to these features. However, if the algorithm is optimizing for accuracy it may well use some features, which would have been best left unused. In the example in the previous paragraph, an adjudicator wanting to minimize crime rate should ignore which group an individual belongs to, even if the most accurate classifier demands that this information be used. A further study of rules of evidence, the reasons they were put in place, and how machine learning algorithms might adopt them would be invaluable.

Another potential in traditional systems one is wary of making "preemptive" decisions, lest they impinge on individual liberties. But it is difficult to define what constitutes a preemptive decision. On the one hand, a decision that is entirely a function of existing data seems to be too deterministic . On the other hand, all decisions are made based on existing data. Nevertheless, we tend to think of a restraining order on an individual, or the denial of certain types of internet or social network access as examples of preemptive actions, while a sentence including jail time for a crime for which one has been convicted does not seem so preemptive. Preemptive judgments have to be made carefully balancing the need to protect individual liberties against the potential harm to society and the individual. How can we define when an algorithmic decision is preemptive? Or perhaps, define the degree of preemption involved in a decision? Do we forbid preemptive decisions in certain spheres and allow them in others? And how do we redesign machine learning algorithms to be only as preemptive as warranted? There has been no work along these lines because even defining what preemption means is challenging.

Algorithmic decision-making is nucleated on decisions made by human decision-makers. This is both a feature and a bug. On the negative side, this means that bias and discrimination in human decision-making become systematized and broadly applied. However, one could hope to mitigate this bias by choosing training data from a diverse set of human decision-makers. One of the major issues in algorithmic decision-making is representational fairness in the training data, both in the set of people the algorithm is trained on, as well as in the set of features it uses.

However there is hope. Unlike a human decision maker, an algorithm is eminently auditable in all aspects: The particular machine model used, how it was trained, and what results it produces on test data can all be fully recorded and analyzed. In fact, for this reason, these algorithms should be monitored and audited constantly. There is a concern about the auditability of really complex algorithms.

To understand this concern, it is first necessary to understand what machine learning models are. When designing a machine learning classifier for a task (such as determining who gets a bank loan) one first chooses the type of classifier one wants. For example, a classifier might have the rule, “If 5*income + assets - liabilities - loan amount > 0, then award the loan.” Of course this is a highly simplified, unrealistic example of this kind of rule. This kind of rule computes a weighted score out of the features of the application and grants a loan if this score is above a threshold. Before looking at any data this model (the type of rule the algorithm will use) is already chosen; during training the weights (5 for income, +1 for assets, -1 for liabilities, and -1 for loan amount) are learnt.

Another popular kind of model is a decision tree. Here the algorithm asks a series of (typically yes/no) questions, and based on the answers, decides how to classify. In the same highly simplified setting above, the first question might be, “Is income > 3*loan amount?” If the answer is “yes,” a second question might be, “Is credit rating > 600?” and if “yes” again, perhaps the loan is granted. However, if the answer to the first question is “no,” then perhaps a second question might be, "Is assets - liabilities > loan amount?," and again if the answer is “yes” the loan might be granted. So each possible sequence of answers to these questions leads one down a path, at the end of which lies a decision - to grant the loan or not. Here the model is a decision tree. But what question to ask first, what subsequent questions to ask for each possible answer to the first question, etc., are learnt from the training data during the training phase.

These are just two of the simplest and most common models in use.

Lately, a form of learning called deep learning has proved remarkably effective in natural language understanding, computer vision, and many other tasks. Here models are incredibly complex, loosely based on our understanding of how neurons connect and work in biological organisms. Again, the rough architecture of the neural net is the model chosen *a priori*, and the training data helps determine the parameters of this model.

As may be evident from the three examples of models above, they can vary greatly in their understandability. Notions such as transparency, interpretability, accountability have become increasingly important desiderata as algorithms in complex models are making decisions nearly as inscrutable as some decisions made by humans.

As remarked earlier, machine learning algorithms in critical decision making tasks should be monitored continually and audited from time to time. It also seems clear that these monitoring and auditing tasks should not be done solely by mathematicians and engineers. To engage policy makers, social scientists, etc., it is important that these models be transparent and interpretable.

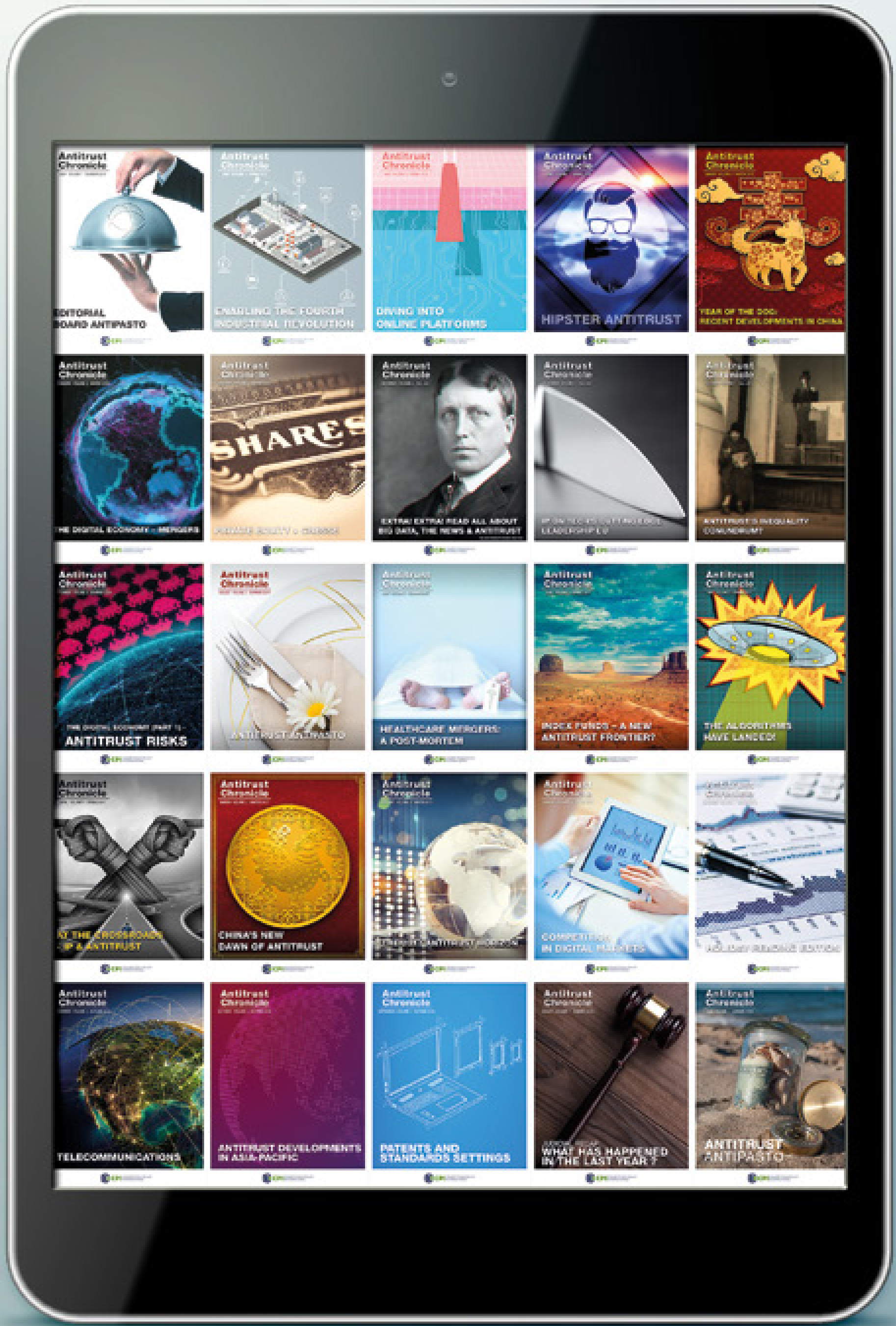
In summary, algorithmic decision making offers the prospect of reducing bias in critical decisions. However, in order to succeed such algorithms should mimic some aspects of human decision making that have evolved over centuries of experience. Because algorithms often learn from flawed data, and because their design might not be perfectly aligned with societal goals, they should be continually monitored, and modified as necessary. In any particular application, choosing between competing, and sometimes incompatible notions of fairness will remain a major challenge, that must be solved jointly by algorithm designers, policy makers, and society at large.

CPI COMPETITION POLICY® INTERNATIONAL

CPI Subscriptions

CPI reaches more than 35,000 readers in over 150 countries every day. Our online library houses over 23,000 papers, articles and interviews.

Visit competitionpolicyinternational.com today to see our available plans and join CPI's global community of antitrust experts.



CPI COMPETITION POLICY® INTERNATIONAL