AUDITING THE RANKING STRATEGY OF A MARKETPLACE'S ALGORITHM IN THE FRAME OF COMPETITION LAW COMMITMENTS WITH SURROGATE MODELS: THE AMAZON'S BUY BOX CASE

Documents de travail GREDEG
GREDEG Working Papers Series

Jeanne Mouton
Benoit Rottembourg

GREDEG WP No. 2024-27
https://ideas.repec.org/s/gre/wpaper.html

Auditing the Ranking Strategy of a Marketplace's Algorithm in the Frame of Competition Law Commitments with Surrogate Models: The Amazon's Buy Box Case

Jeanne Mouton

(Université Côte d'Azur, CNRS, GREDEG, France & EU Commission[1])

Benoit Rottembourg

(Inria, Regalia)

GREDEG Working Paper No. 2024-27

Abstract: *In a global context where competition authorities are investigating and sanctioning Amazon's marketplace for practices of self-preferencing at the expense of their business users and consumers (Italian AGCM 2021, EU Commission 2022, UK CMA 2024, US FTC on-going since 2023), we observe a trend of imposing remedies on dominant players in digital markets. In addition, the Digital Market Act, shifting from an ex-post enforcement approach to ex-ante obligations on designated gatekeepers, is strengthening auditing power over these gatekeepers, which risk heavier penalties in the event of non-compliance. Therefore, competition authorities and regulators need tools to audit the compliance of these dominant players in the e-commerce sector over the obligations and remedies they are imposing on dynamic, and personalized algorithms. Most of these algorithms embed Machine-Learning components, introducing opacity and potentially biases in the decision-making process. The aim of the paper is to explore the benefits of using black-box auditing techniques to provide insights into the behavior of these online algorithms. We anchor our research in the literature of product prominence from vertically integrated players, of choice ranking, and of the specific literature related to Amazon search ranking, automatic pricing and Buy Box's algorithms. Through a study of the pricing and ranking of several thousand products on Amazon, from 2017 to 2023, we illustrate the potential of surrogate models. While our dataset only covers some categories on Amazon.fr, the large number of competitions allowed us to demonstrate, with a 94% accuracy, that the variable is Amazon, or variables correlated to it, had a positive effect on winning Buy Box before mid-2022, and that this positive effect has decreased after mid-2022. In our research, the machine learnings models revealed a significantly higher degree of accuracy and sensitivity compared to a logistic regression, opening the discussion on the added value and role of surrogate models based on machine learning techniques in guiding the auditor, as well as raising the question of their probative value in the regulatory context.*

Keywords: algorithms, ranking algorithms, digital markets, online marketplace, competition law, audit, machine learning.

JEL codes: K21, L41, L51, L81.

[1] The information and views set out in this article are those of the author and do not necessarily reflect the official opinion of the European Commission.

Introduction

This paper proposes a methodology to audit, ex-post in the regulatory framework, concerns of self-preferencing of a marketplace' ranking algorithm. We rely on the European (EU) Commission decision in December 2022, imposing remedies on Amazon's Buy Box, as a case study.

Amazon is a vertically integrated online marketplace platform on which third-party sellers meet buyers, and on which Amazon also markets its own-brand products. In addition, the platform has developed complementary offerings for third-party sellers (in terms of logistics, commercial data, and cloud services) as well as for consumers (a fast delivery service, video and audio streaming, etc.). There are several algorithms at work on Amazon marketplace, among which, a search engine algorithm (Subhradeep et al., 2022), a pricing algorithm (Fry and Manna, 2016), an algorithm to treat reviews (Seele et al., 2021). In this paper we focus on the Amazon Buy Box algorithm, meaning, the algorithm that determines the seller who wins the Buy Box for a determined product, when there are several sellers. The Buy Box is a privileged ranking position on the product detail page and which only displays the top recommended offer.

The EU Commission opened a formal investigation in November 2020 under the concerns over a bias in granting sellers' access to its Buy Box, leading to a preferential treatment of Amazon's own retail business and sellers using Amazon's logistics and delivery services. Amazon submitted a first set of commitments on the 8th of July 2022[2], to address the Buy Box concern, which included: treating all sellers equally when ranking the offers for the purposes of the selection of the Buy Box winner, displaying a second competing offer to the Buy Box winner if there is a second offer from a different seller that is sufficiently differentiated from the first one on price and/or delivery. This first set of commitments were subject to a Commission's market test, in which all interested third parties were consulted to verify whether these commitments would remove all the competition concerns. Following the results of the market test,[3] Amazon amended its initial proposal on the 22nd of November 2022. On the 20th of

[2] Press release, Commission seeks feedback on commitments offered by Amazon concerning marketplace seller data and access to Buy Box and Prime, 14 July 2019, available at: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_4522

[3] The market test can be framed as a game of revelation of information and lobbying in which the competitors and third parties have the incentive to ask for better commitments from the firm under scrutiny, and the Commission aims to find an equilibrium between the firm's proposal and competitors' concerns.

December 2022, the EU Commission accepted commitments from Amazon to treat equally all sellers when ranking the offers for the purposes of the selection of the Buy Box winner.[4]

This case study triggered our attention in a global context of an increase of commitment decisions following concerns of abuse of dominance in digital markets.[5] The Amazon's Buy Box is prone to generate a distortion of competition for complementors while exploiting the constrained rationality of the consumers. This potential distortion of competition occurs with the vertically integrated platform preferencing its own services. In this particular case study, the self-preferencing operates through discriminating conditions in the ranking algorithm. Auditing a recommendation algorithm *per se* already raises challenges (Metaxa et al., 2021; Mullainathan et al., 2012; Gaddis, 2018), but in this case study, the remedies aim to restore the competitive process, adding an extra challenge. Once the regulator identifies the concerns and accept the commitments by the undertaking, the trustee is in charge of monitoring the implementation of the commitment decision. It is in this very technical context, of imposing remedies on a ranking algorithm requiring ex-post monitoring, that we are comparing the pros and cons of potential methodological framework through surrogate models.

We propose a methodology framework to audit a ranking algorithm of a digital market, demonstrating the contribution of surrogate models. In this paper, we collect data from an API price tracker and then compare a logistic regression with modern machine learning methods to demonstrate the added value of using more sophisticated technologies. We claim that the latter provide sensitivity analysis offering richer insight to the auditor. Finally, we try to identify whether there is an evolution of the key explanatory variables over time and whether we can identify a tipping point. In terms of expected results, a tipping point in our models would be a differentiation event in the platform's algorithm, which, in our case analysis, could be interpreted as an attempt to comply with the remedies imposed by the regulator.

We anchor our research in the literature of product prominence from vertically integrated players (Cure et al., 2022), of choice ranking (Grobvic and Chen, 2018), and of the specific

[4] Press release, Commission accepts commitments by Amazon barring it from using marketplace seller data, and ensuring equal access to Buy Box and Prime, 20 December 2022, available at: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_7777

[5] The OECD issues a report in 2022 on remedies and commitments in abuse cases and found around 30% of abuse of dominance cases with settlement or commitments in OECD countries in 2020, and around 40% for Europe countries16. The issuance of the report was fuelled, among other reasons, by the growing number of enforcement actions against digital platforms. Digital markets characteristics can be specifically challenging when it comes to designing remedies. Indeed, the digital sector is deemed to be fast-moving and intensely innovative which can render remedies ineffective if market conditions quickly alter over time.

literature related to Amazon search ranking, automatic pricing and Buy Box's algorithms (Chen et al., 2016; Gómez-Losada et al., 2022; Hunold et al., 2022; Farronato et al., 2023; Raval 2023; Waldfogel, 2024; Dash et al., 2024).

First, we review the literature. Then, we present our research method, from the data collection to the three steps research method applied. Finally, we discuss our results and provide policy recommendations for remedies imposed on ranking algorithms in digital markets.

Review of literature

There are several trade-offs between using a commitment decision versus sanctioning an infringement in EU competition law competition law. Both the EU Commission and the companies investigated benefit from procedural economy gains of settlement: it is cheaper and faster to settle rather than litigate, settlements are pragmatic (Dunne, 2014). However, commitments decisions results in costs: loss of enforcement of Article 102 (in terms of clarification of the law, public censure, deterrence, punishment, disgorgement of illicit gains) (Wils, 2008) and cost of supervision of the commitments. Settling comes at the cost of learning from enforcement of competition law (Steviev, 2014).

Choné et al. (2014) and Gautier and Petit (2018) discuss the optimality of commitments decisions and its interaction with formal infringement decisions in EU competition law. Feasey & Krämer (2019) study how to remedy "intermediation biases" by vertically integrated companies. Hoehn (2020) discusses the challenges of designing and implementing behavioural remedies in the digital economy. Petit and Gal (2021) consider three radical restorative remedies for digital markets: mandatory sharing of algorithmic learning, subsidisation of competitors and temporary shutdown.

With specific regards to Amazon's algorithm, Janger & Twerski (2019) challenge the myth that Amazon is a neutral intermediation platform between sellers and buyers. On the contrary, Amazon has implemented several strategies, including the Buy Box, enabling it to control each sale and maximise its profits. Yet, these mechanisms are not transparent to the consumer nor to the seller. This deconstruction of the platform's neutrality is also illustrated by its pricing algorithm. Chen et al (2016) gathered four months of data on all merchants on the Amazon platform selling one of the 1,641 top-selling products of their sample with the aim of detecting automated pricing algorithm strategies. The results of their study does not show a clear advantage for cases where Amazon is a seller but demonstrated that sellers who might use pricing algorithms are more frequently winners of the Buy Box, which would have the effect of exacerbating a disparity between algorithmic sellers and non-algorithmic third-party sellers.

Specific to self-preferencing practices, Farronato et al. (2023) study whether Amazon is engaging in self-preferencing practices in search ranking products looking at real consumer search data. The sample studied by the authors covered 3,019 unique searches by 184 users. Looking at the data on search users, it confirms that ranking matters since 72,1 percent of search users did not click past the first page. The authors found that Amazon brand was a meaningful predictor in search results, suggesting self-preferencing practices from the platform. Dash et al., 2024 further researched a series of empirical investigations as to whether Amazon engages in self-preferencing in relation to Amazon's Buy Box, offer listings pages, Alexa Search and recommendation systems.

Specific to the Amazon's Buy Box, Etumnu (2022) studies the performance of third-party sellers against Amazon by distinguishing between products that are sold and shipped by Amazon, those sold by the third-party seller and shipped by Amazon and those sold and shipped by a third-party merchant. Etumnu observes equivalent performance between the first two categories, depending on whether the product is sold by Amazon or by a third-party seller using Amazon's logistics, and weak performance for the last category, but cannot draw any conclusions about the implementation of anti-competitive practices. Gómez-Losada et al. (2022) studied twenty-two products on the Amazon page in Italy over a period of ten months and carried out two experiments with the aim of estimating the characteristics considered by Amazon in order for a seller to appear in the purchase box. One of the experiments involved predicting which seller would appear in the Buy Box. The authors noted that the consumer's experience and the price dynamics of the sellers were criteria taken into account when ranking the Buy Box. Raval (2023) studies Amazon's algorithm rules to determine which seller wins the Buy Box and to measure Amazon's self-preferencing practices. The interpretation of his results shows self-preferencing practices in favor of Amazon's retail and Amazon fulfilment. There is also a preference for Amazon Retail over third party FBA offers. Waldfogel (2024) empirically studies Amazon's own search ranks on the marketplace in the context of the entry of the DMA which prohibits self-preferencing. The author finds that shortly after the EU Commission designated Amazon as a "gatekeeper", the Amazon rank differential fell from a

30 position advantage to a 20 position advantage, compared to other major brands' rank positions which were unaffected.

Summarizing from an economics risk assessment perspective[6]: Amazon is a vertically integrated platform, with the resources to implement self-preferencing practices. If the product presented in the Buy Box is the result of a biased ranking, and not the best product (according to determined criteria), there is manipulation of the information provided to the consumers and a diversion of demand from potential winning sellers. Furthermore, the Buy Box has been qualified as "unavoidable" for consumers, they are nudged to select the seller winning the Buy Box over any other sellers.

Method

1.1. Framework of the research methodology

The closest paper to our research in terms of method is the paper from Raval (2023) that aims to study Amazon's algorithm rules to determine which seller wins the Buy Box and to measure Amazon's self-preferencing practices. Raval collected data from several countries (US, UK, Germany, France) and product categories between December 9, 2020 and January 29, 2021. To predict which offer wins the Buy Box, the author estimates a multinomial logit model. The empirical model from Raval is an example of an explanatory linear regression. In this essay, differentiating from the model from Raval, we use data collected on Amazon.fr only, we collected less products and sellers' data but more granular data. Raval compared the approximations of his empirical model with the actual winners of the Buy Box, and his model correctly predicted the Buy Box offer 88% of the time across all categories and 80% of the time for products with multiple offers.

We use the same source of data collection as Raval. However, we restrict both geographic scope to our data set, only focusing on France, and we collect data from less categories, with the aim to have for each ASIN[7] at least three sellers on average over the last 90 days prior to the data collection. Hence, we restricted the scope but with the aim to achieve higher

[6] For a survey on the economics literature of hybrid marketplaces with a focus on Amazon and self-preferencing, see Etro (2023).
[7] An ASIN is the Amazon Standard Identification Number, to identify a product.

granularity. Keepa maintains price histories for all products on Amazon, and allows searches by categories of products, ASIN or seller identification.

We aim to test the following hypotheses. First, we complement a logistical regression with more sophisticated machine learning methods, relying on the hypothesis that machine learning methods will improve the accuracy of our model and capture more sensitivity to variable changes. Second, we make the hypothesis that before the implementation of the remedies our model will show the importance of the variable *isAmazon* as reflecting self-preferencing practices. This finding will be aligned with the literature (Raval, 2023). Third, we expect the importance of the variable *isAmazon* to decrease in the course of the negotiation and pre-implementation of the remedies, which could either indicate a decrease or a cease of self-preferencing from Amazon.

We built and compared four models in order to identify the influence of variables in the Buy Box winning decision. The logit regression and the decision tree optimize for a set of parameters, with either a set of coefficients or one tree. The random forest and gradient boosting methods are ensemble methods, hence iterative approaches. In the random forest method, 10% of the data or a column are removed from the first tree, and this is repeated 100 or 150 times. In the gradient boosting method, the second tree learns from the "mistakes" of the first tree and this is repeated 100 or 150 times. All four models use the same characteristics to describe an offer by a seller for a given product:

- Year and Month of the BuyBox competition
- Price of the offer
- isPrime (when the offer is compatible with the prime program of Amazon): True/False
- isMAP[8]: True/False
- isShippable[9]: True/False
- isAddonItem[10]: True/False
- isPreorder: True/False
- isWarehouseDeal[11]: True/False

[8] From Keepa API description: is MAP means "if the price of this offer is hidden on Amazon due to a MAP (minimum advertised price) restriction.
[9] From Keepa API description: False to is Shippable might mean that the offer is temporarily out of stock or a pre-order.
[10] From Amazon's website: "Add-on items" https://www.amazon.com/gp/help/customer/display.html/ref=hp_468520_buyaddon?nodeId=200876660
[11] From Amazon's website: "Warehouse deals" for second-hand products

- isScam[12]: True/False
- isPrimeExcl (exclusive for Prime customers): True/False
- isFBA[13]: True/False
- shipsFromChina: True/False
- allTimePositiveRating: 0-100, in percent
- allTimeRatingCount: number of ratings

Our approach consists in building every pair of offers, o1 and o2 (we have close to one million such pairs in our dataset) that compete for the Buy Box for a given product, when at least one of them wins the Buy Box, and try to predict which offer is the most likely to win the Box. For each characteristic of the offers we compute the difference between o1 and o2 and the model "learns" the effect of this set of differences on the result when o1 beats o2 (and win the Buy Box) or the converse.

As shown above, the variables at stake are mainly categorical and even binary variables (with values taken in the set {True, False}). Their difference for a couple of offers is still categorical. Only the price, the positive rating percentage and the rating count can be considered as continuous variables, so is their difference. As frequently mentioned in the literature of e-commerce classification algorithms, tree-based methods are known to out-perform both linear models and neural network techniques when large tabular (and often categorical) data is concerned. For a survey on application of Machine Learning techniques in e-commerce we refer to (Mice et al., 2019; Policarpo et al., 2021; Bertolini et al., 2021).

Within the family of Machine Learning techniques, tree ensemble methods and boosting techniques - like Random Forests or Gradient Boosting models -, have performance and robustness properties that make them the solution of choice in e-commerce classification and ranking solutions (Breiman, 1996 & 2001; Friedman, 2001; Chen and Guestrin, 2016; Prokhorenkova et al., 2018). More recently, a comparison between tree ensemble methods and neural networks (or deep learning techniques) for tabular data classification problems can be

https://www.amazon.fr/Offres-Reconditionnees-Toutes-Les-Offres/b?ie=UTF8&node=8873224031

[12] From Keepa API description: is Scam is a "Boolean value indicating whether or not our system identified that the offering merchants attempts to scam users".

[13] Frm Keppa API description: is FBA "whether or not this offer is fulfilled by Amazon".

found in Shwartz-Ziv and Armon (2022) and Grinsztajn et al. (2022). They assess the performance of tree ensemble methods.

It is likely that Amazon itself relies on such techniques to produce its ranking and elect the Buy Box winner, the one that maximizes the economic value for the platform. Thus, we have developed and trained four classic classification models, two explainable ones - a multinomial regression and a decision tree (Breiman et al. 1984) - and two machine learning models, a Random Forest (Breiman, 2001) and a Gradient Boosting (Chen and Guestrin, 2016). In the following we will name those models respectively: “logit”, “decision_tree”, “random_forest”, and “gradient_boosting”. We deliberately limited ourselves to the basic characteristics of the offers and did not add external variables or knowledge. Similarly, we did not optimize the parameters of the four models intensively to compare them in a fair manner.

1.2. Data collection

The data was collected from Keepa, a third-party API for Amazon, tracking all products available in US, Germany, UK, France, Italy, Spain, Canada, Japan, India, Mexico, and Brazil. Each product is identified by an Amazon Single Identification Number (ASIN).

However, we restrict both geographic scope to our data set, only focusing on France, and we collect data from less categories, with the aim to have for each ASIN at least three sellers on average over the last 90 days prior to the data collection.

We decided to arbitrarily focus on the top 7 categories of Raval’s sample in number of offers (median and average): books, CDs, DVDs, videogames, toys, tools, and sport.

We collected the ASINs on these seven categories whenever the new offer count was greater than 3 on average on the last 90 days. We also collected ASINs on several sub-categories of these six categories with the same rule of collecting ASIN whenever the new offer count was greater than 3 on average on the last 90 days. The table below summarizes all the categories of products on which ASIN data was collected as well as the number of ASINs:

Category	***Sub-category***	***90dav.: New offers count greater than 3 (number of ASINs)***
Books		50
Books	Bandes dessinées	54
Books	Roman et littérature	17
Books	Tourisme et voyage	80
CDs		81
CDs	Musique classique	91
CDs	Pop	88
CDs	Jazz	84
DVDs & Blu-ray		54
DVDs & Blu-ray	Films	58
DVDs & Blu-ray	Séries TV	76
DVDs & Blu-ray	Actions et aventures	63
Videogames		64
Videogames	Xbox ones: jeux consoles accessoires	51
Videogames	Jeux vidéos PC: jeux et accessoires	57
Toys		72
Toys	Jeux de société	77
Toys	Jouets radiocommandés	24
Toys	Maquette et modélisme	72
Tools		36
Tools	Quincaillerie	38
Sport		30
Sport	Cyclisme	28
Sport	Vêtements de sport	32
Sport	Sports nautiques	84

The dataset created has 91,000 offers for 9,486 sellers. For 11% of the offers Amazon is the seller. The oldest offer is from the 22/06/2016 and the most recent on the 06/09/2023. On the dataset we found around 80,000 observations in which with have a unique Buy Box's winning offer. Therefore, we could create about 1,6 million competitions between the winning offer and every non-winning offer. Unfortunately, due to the lack of access to rankings, for offers not winning the Buy Box, we had to limit our competitions to competitions with the winner. When competitions had no winners at all, we eliminated the corresponding offers from our learning procedure.

Before diving in prediction models, we observed the change in the percentage of wins (2019 2023) depending on whether you are Amazon, FBA, or Merchant (6-month moving average).

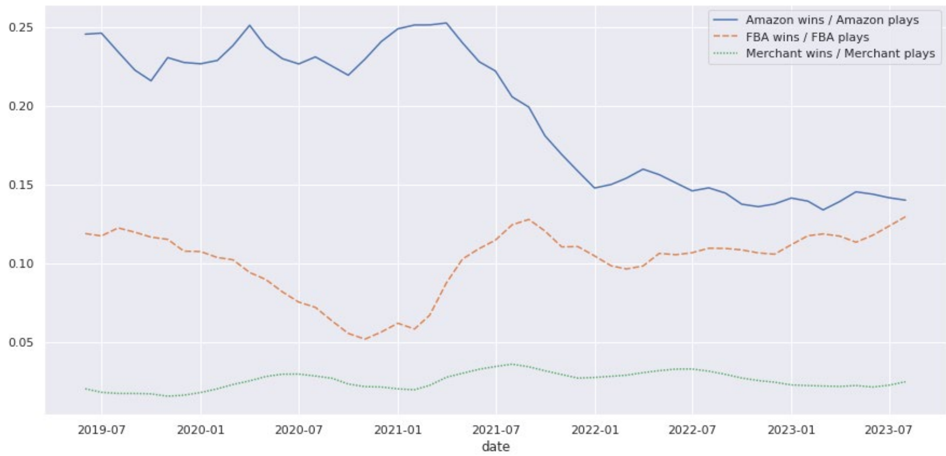


Figure 4: Change in percentage of wins, 2019-2023 depending on whether the seller is Amazon, FBA, or Merchant.

While this figure does not prove anything (all sellers could have changed their behaviour independently from the Buy Box), it helps describing the dataset depending on the category of the seller.

Additionally, we observe from July 2019 to July 2023 the variation in the percentage of times Amazon wins versus Amazon plays, by product family, over time.

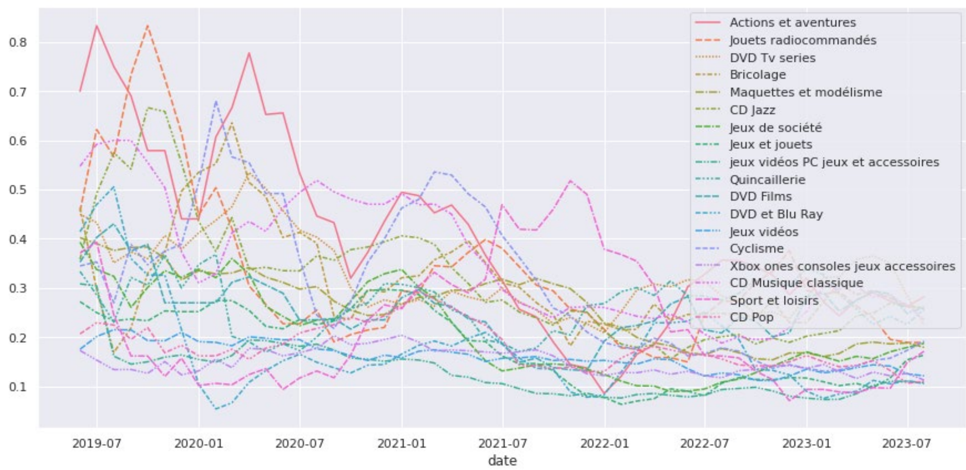


Figure 5: Variation in the percentage of times Amazon wins versus Amazon plays, by product family, over time.

1.3. Accuracy of the models' predictions

Overall, the four models proposed good to excellent prediction accuracies. About 83% accuracy for the "logit" model, which is very close to the results found in Raval (2023), about 90% for the "decision_tree" approach, 93% for the "random_forest" and 94% for the "gradient_boosting".

The figure bellow shows the respective accuracy of the four models for varying instances of the learning data sets (subtracting or adding samples in the learning data of the models).

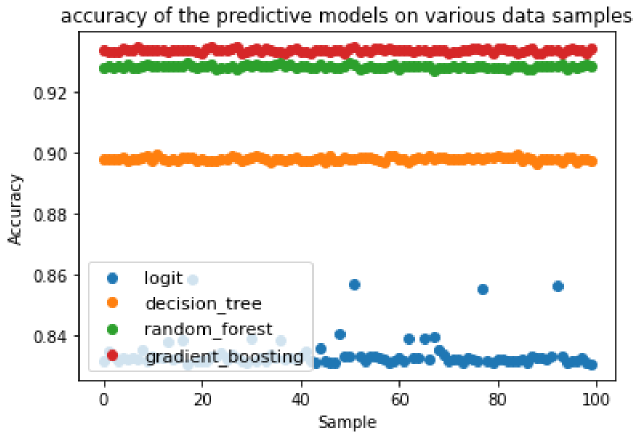


Figure 6: Accuracy of the predictive models on various data samples.

It is important to note that the four models are better at predicting the Buy Box winner when a product history is known. Previous winners offer a good indication of future winners, particularly when the competitor set is stable. A loss of about 2% performance occurs when a product is unknown to the prediction engine. Funnily, the accuracy does not vary significatively between the 23 family of products that we studied, giving a strong indication that Amazon probably uses the same family of algorithms for these sets of products.

Results

1.4. Variables' influences

All models highlight a strong importance to the four variables:

- Price
- *isAmazon*
- isFBA
- allTimeRatingCount

For our best model, the "gradient_boosting", studying Shapley values is a classic way to observe the relative importance of the variables in the prediction. A prediction can be explained by considering that each feature value of the instance is a "player" in a game where the prediction is the payout. Shapley values – a method from coalitional game theory – tells us how to fairly distribute the "payout" among the features. The Shapley value is considered as one of the rare methods to deliver a full explanation because it is based on a solid theory and distributes the effects fairly (Shapley, 1953; Sundararajan and Najmi, 2020; Albini et al., 2023).

The Shapley values can, in theory, be computed for any kind of Machine Learning algorithm. We show the results on 1000 data points in Figure 8. Even if the (Shapley) value of a variable is hard to interpret (it represents its impact on the probability function of winning the Buy Box), their relative importance emphasizes their respective role in the classification model.

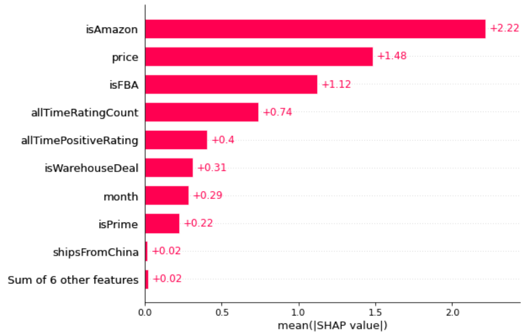


Figure 7: Shapley value computed on 1000 data points.

1.5. Differentiation event in the Buy Box's algorithm

From our model, can we detect a shift in the model over time, a tipping point?

We try to identify a change, relying on the hypothesis that Amazon started to make changes in the process of the negotiations of the commitments offered to the EU Commission. One of our hypotheses was to add "time" as a variable and to check if this variable plays a role in the explicability of the model. However, we are not sure that we will be able to identify a "clean break" over time in Amazon's algorithm. Indeed, on large platforms, A/B testing may be progressive, some family of products being privileged over others, in order to minimise risks and learn on the fly. This would make even more sense in the context of the commitments offered to the EU Commission since there were five months between the first set of commitment and the final set accepted by the EU Commission. Therefore, it might be challenging to identify a before and after the commitments over one day, but we expect to identify a shift in the model over two periods of time, amounting to several months.

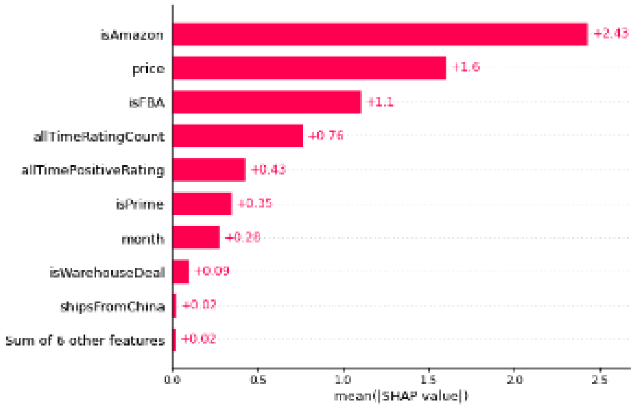
The first insight on a change of "regime" for the Buy Box algorithm, is the change in chances of success of Amazon's offers, for every family of products as shown on figure 4.

For a given month m, say June 2022, our approach consisted in building two models. One model for all competitions happening before m, and another model for all competitions happening after m. We used the same gradient_boosting algorithm for the two models, with different learning data sets.

Three conclusions could be drawn:

- first, the two models were of good to excellent quality, close to 92% accuracy,
- applying the model_before to competitions after June 2022 (for which it had not been trained) gave significantly poorer accuracies, showing a strong difference in behaviour. A loss of 7% of accuracy was observed,
- comparing Shapley values for the two models offered significant differences. The contribution of the variable *isAmazon* was considerably diminished in the model after June 2022 compared to the model before June 2022.

The figure highlights the change of importance of variable *isAmazon*, before and after June 2022.



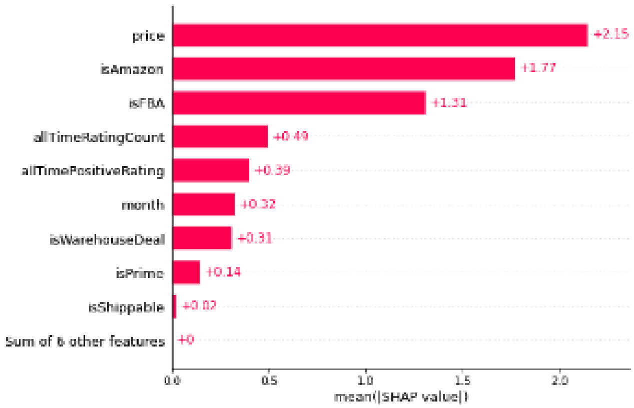


Figure 8 (above) and Figure 9 (below)

<u>Figure 8:</u> Shap values of the model before June 2022 (variable importance on the chance of winning).

<u>Figure 9:</u> Shap values of the model after June 2022 (variable importance on the chance of winning).

To summarise our results: we created a "surrogate model" that matches the behaviour of Amazon's Buy Box, and which is capable of reproducing the behaviour of the real Buy Box within 94% of accuracy. Our model, based on machine learning technology has learned from a sample of 1,600,000 seller 1 and seller 2 pairs, which one has the best chance to win the Buy Box for two given product offers. Our model showed even good results for a competition involving a completely unknown ASIN (without longitudinal learning).

When we analyse the explanatory variables in our model, relying on Shapley method, the variable *isAmazon* clearly stands out. The same result can be obtained with a logistic regression, but with a weaker predictive quality. When creating a counterfactual on the *isAmazon* variable, we observe a significant deterioration in the model's performance. The interpretation is the following. Our surrogate model replicates the real Amazon's Buy Box which needs the *isAmazon* variable to perform. It is possible that:

- Amazon uses variables other than those to which we have had access,
- The behaviour of Amazon as a seller, is so unique (by its pricing practice, its footprint or other selling tactics) that it can be inferred from the data without explicitly using the *isAmazon* variable, so we cannot be sure that the variable *isAmazon* is used,
- However, it does have an effect in our fairly precise model with the set of visible variables we had access to.

Finally, around mid-2022, we can observe a change in the model. The 'before' model (trained on competitions before June 2022) and the 'after' model (respectively) are not identical, not interchangeable, with different qualities in precision. The values of the explanatory variables (as defined by Shapley) are not the same before and after. In particular, the price variable (more important than before) and the *isAmazon* variable (less important than before). The impact of the counterfactual is significantly greater before than after. Furthermore, our descriptive analysis also shows that before mid-2022 Amazon won proportionately more competitions than after mid-2022.

This leads us to believe that Amazon heavily changed its algorithm during 2022 and that it seems that a similar algorithm is at work for all the product categories observed in our sample. The change occurred gradually because the effect of the change is not sudden and is identical for all categories observed (see Figure 5). Only looking at the variables available, Amazon products have an advantage over the competition, but this advantage has significantly decreased in the course of 2022.

Discussion

While it seems that the commitments had a positive effect, suggesting some efficiency, at which costs? In other words, what can our results say about the auditability of these commitments?

To obtain our results we first had to collect data through an API of a third party (which is costly and requires programming skills). This was only possible because of the fact that there are several API for Amazon, but this degree of data collection is not available for all online platforms that may be under antitrust scrutiny. The ideal situation would be for an auditor or a trustee to be able to "open the black box" and have access to both the data and the code. Since we could not have access to any of them, we relied on surrogate models. While the first one, a logistic regression, is rather simple, the machine learning models needed the expertise of data scientists. We chose to run four predictive models in order to compare the accuracy of each model from a simpler to a more sophisticated one. It would not be necessary for an auditor to run several models, however, their choice in which model to rely on, would depend on a trade-off between the explicability of the model and its accuracy. In our research, the machine learnings models revealed a significantly higher degree of accuracy and sensitivity, supporting the need to have a cooperation of lawyers, economists, and data scientists on complex antitrust cases in digital markets.

From our results, around mid-2022, we can observe a change in the model. The 'before' model (trained on competitions before June 2022) and the 'after' model (respectively) are not identical, not interchangeable, with different qualities of precision. We made the hypothesis that the remedies should be implemented by the company on or around the date of implementation of the commitment decision. The remedies are supposed to restore competition from a situation in which, in this case, the company may have abused its dominance (since there is no formal decision on the finding). Since remedies are supposably correcting for a situation in which the firm is diverting profits and their implementation should cause a loss of profits, therefore the firm should delay as much as possible their implementation, i.e. 6 months from the date of the commitment decision. However, remedies are proposed in an information asymmetry setting, reinforced with the opacity of the algorithm of the Buy Box. Hence, when we observe a change in our model months before the date of the decision, it seems that the company is proposing remedies that they know they can implement, that they have already tested, suggesting the aim for the company to minimize their losses in their remedies' proposal. This minimization of the losses would consider the need that a trustee could observe a variation in the algorithm model, as an indication of compliance. Such a strategy could be questioned for *fairwashing*.[14]

While this observation could challenge whether a commitment decision was the right enforcement tool for this case, it is again necessary to recall that a commitment decision involves trade-offs. Commitment decision are cheaper and faster to implement (Dunne, 2014), and a systematic use of commitment decisions is optimal for the most harmful and difficult cases to deter (Choné et al., 2014). The Amazon's Buy Box would have unarguably been a difficult case to litigate, therefore a situation in which competition would have been restored under a formal infringement procedure could have taken years. However, the commitment decision comes at several costs (Wils, 2008; Steviev, 2014), including that, potential claims for damages from sellers on Amazon will be not able to rely on a formal infringement decision. All sellers that have been evicted or exploited, through these discriminatory practices, would need to demonstrate in Court that Amazon infringed competition law, establish the causal link between the infringement and their damage, and evaluate the monetary damage.

Our model also triggers broader discussions in law & economics, among which the trade-off between the explicability of the methods chosen to assess the compliance of an algorithm and their accuracy. In our paper we compare the results of machine learning models, that are more sophisticated than a regression. However, the interpretation of the explanatory variables is not as straightforward than with a regression since we need to implement the Shapley method. If

[14] Yaghini et al. (2024) illustrate the situation between a regulator wishing to design penalties that enforce compliance with their specification without disincentivizing machine learning builders from participation: they call it the regulation game for trustworthy machine learning. In their paper the authors model the relationship between machine learning builders and fairness and privacy regulators. They authors introduce ParetoPlay, acting as an equilibrium search algorithm, seeking social optimum solutions to the game.

Shapley values have already been used as a presumption and are part of the state of art when it comes to machine learning models, it adds a degree of complexity in explaining the results while increasing the accuracy of the results. Hence, there is a clear trade-off between the explicability (which difficulty increases with the sophistication of the method) and the accuracy of the results (which, in our case, positively increases with the sophistication of the method). Which method is the most solid statistically to yield conclusions and self-preferencing prejudices? This question should be framed in the debate of what a proof is and what is the precision required for a technical proof advanced to a regulator. Should we be satisfied with a simple proof that is less explanatory, because it is easier to explain to a judge and therefore facilitates the regulatory discussion as well as the judicial review? Digital markets raise complex issues, making it more difficult to establish a counterfactual. There is a real need for technical expertise and inter-disciplinary dialogue between data scientists, lawyers, and economists (Fletcher et al., 2021). There is also the issue of the acceptance of AI-generated proofs in a context where AI applied to law seems to be poorly accepted and it is morally more complex to accept the risk of AI error, as opposed to a human error.

Conclusion

This paper aimed to provide a methodology for surrogate models in the frame of competition law commitments decision auditing algorithm. We relied on the Amazon's Buy Box case study. On the side of sellers, since Amazon is vertically integrated, the marketplace has an incentive to manipulate the ranking and engage into self-preferencing. In the case of the Amazon's Buy Box, the practice can either consist in preferencing Amazon, the retailer, in the winning offers or, preferencing retailing subscribing to the FBA program (fulfilment by Amazon) since it also increases Amazon's profits. Therefore, if consumers are manipulated into purchasing products offered by the platform - because of the self-preferencing practices of the vertically integrated platform - complementors suffer from loss of profits. On one hand consumers are biased in their decision making by the recommender system and on the other hand the shift of demand distorts the competition between suppliers (Fletcher et al.; 2023). In the EU Commission's commitment decision[15], the Commission expressed preliminary concerns that the anticompetitive practices involving the Buy Box "prejudices consumer choice and directly

[15] Antitrust procedure, article 9 Regulation (EC) 1/2003, case AT.40462 – Amazon marketplace and case AT.40703 – Amazon Buy Box, 20 December 2022.

harms consumers by driving them to view and transact offers whose selection and display does not mirror the outcome of competition on the merits".[16]

We audited the commitment decision from the EU Commission in December 2022 to impose remedies on Amazon's Buy Box. We add to the literature both on the frame of our research, auditing an algorithm in the context of a commitment decision, as well as on the methodology, by complementing a logistical regression with more sophisticated machine learning methods.

Our results fit in the literature on product prominence in marketplaces, specifically when it comes to the Amazon's Buy Box. While previous papers studied the criteria to win the Buy Box, we add to the literature by testing four methods (logit, decision tree, random forest, gradient boosting) and by focusing on the compliance of Amazon with the commitment decision from the EU Commission. The choice of the categories in the dataset relied on a criterion of a minimum number of sellers for a same offer (at least 3), therefore, when we observe that the variable *isAmazon* contribution is diminished, it is not economically trivial. The more offers and sellers there are in our database, the greater the potential damage resulting from an infringement of competition law. Needless to also remind here that Amazon.fr is the most popular marketplace in France, with 312,36 million visits per month during the last semester of 2022.[17]

In this essay, we relied on a "surrogate model", to replicate the Buy Box algorithm. While our dataset only covers some categories on Amazon.fr, the large number of competitions, allowed us to demonstrate, with a 94% accuracy, that the variable *isAmazon*, or variables correlated to it, had a positive effect on winning Buy Box before mid-2022, and that this positive effect has decreased after mid-2022. We can therefore conclude, from our model, on our limited dataset, that being a product sold by Amazon was conferring an advantage and that this advantage has decreased. If it is not possible to state whether the change in the value of the variable *isAmazon* overtime can be interpretated such as Amazon stopping practices leading to unduly favouring its own services, it does seem that the commitments had a positive effect on the intra-marketplace competitions for the other sellers. However, the wording from the commitment decision, "non-discriminatory conditions", would suggest the need to establish a counterfactual, of what would non-discriminatory conditions mean compared to the situation in which discriminatory conditions were applied to select the featured offer. Regretfully, the

[16] Ibid, para 208.
[17] Source is Statista: https://fr.statista.com/statistiques/475107/sites-e-commerce-les-plus-visites-france/.

commitment decision is opaque on the assessment of the effects of the compliance which rests in the hands of the designated trustee.

One of the limits of our models is the behavioural dimension that is not considered. Consumers are allowed to like and even prefer Amazon. If consumers like better Amazon, should not they benefit from the self-preferencing practices and aren't the complementors not efficient enough? This discussion can be framed on the very limited literature on the welfare effects of self-preferencing practices (Bougette et al., 2022). However, one could argue that consumers preferring Amazon products are not expressing a "real preference" but rather a shaped preference. Consumers or rather, algorithmic consumers[18] might have been trained to prefer Amazon's products. There would be a learning experience effect resulting in biased decision-making in favour of Amazon's products, which would not reflect a balancing of each qualitative dimension when purchasing a given product. While this strong hypothesis would need to be tested through an experiment, it would be aligned with the findings that AI has the potential to shape and manipulate consumers' preferences.

Finally, our paper fits in the discussion of all the others use cases of algorithms in antitrust. There is a strand in the empirical literature, in which data science techniques are relied upon to detect ex-post distortions on the market with massive datasets. Huber and Imhof (2019) as well as Wallimann et al. (2022) propose to combine statistical screens with machine learning techniques to detect bid-rigging cartels.

Bibliography

Albini, Emanuele, Shubham Sharma, Saumitra Mishra, Danial Dervovic, and Daniele Magazzeni. "On the Connection between Game-Theoretic Feature Attributions and Counterfactual Explanations." Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (2023):411-431.

Bertolini, Massimo, Davide Mezzogori, Mattia Neroni, and Francesco Zammori. "Machine Learning for Industrial Applications: A Comprehensive Literature Review." *Expert Systems with Applications* 175, (2021): 114820.

[18] The expression "algorithmic consumers" comes from a paper from Gal and Elkin-Korren (2016) in which the authors emphasised both the virtues and downsides for consumers in relying on algorithms in their decision-making processes.

Bougette, Patrice, Oliver Budzinski, and Frédéric Marty. “Self-Preferencing and Competitive Damages: A Focus on Exploitative Abuses.” *The Antitrust Bulletin* 67, no. 2 (2022): 190-207.

Breiman, Leo, Jerome Friedman, Charles J. Stone, and R. A. Olshen. *Classification and Regression Trees*. Taylor & Francis ed. (1984).

Breiman, Leo. “Bagging Predictors.” *Machine Learning* 24, (1996): 123-140.

Breiman, Leo. “Random Forests.” *Machine Learning* 45, (2001): 5-32.

Chen, Le, Alan Mislove, and Christo Wilson. “An Empirical Analysis of Algorithmic Pricing on Amazon Marketplace.” Proceedings of the 25th international conference on World Wide Web (2016):1339-1349.

Chen, Tianqi and Carlos Guestrin. “Xgboost: A Scalable Tree Boosting System.” Proceedings of the 22nd ACM sigkdd international conference on knowledge discovery and data mining. (2016):785-794.

Choné, Philippe, Saïd Souam, and Arnold Vialfont. “On the Optimal use of Commitment Decisions Under European Competition Law.” *International Review of Law and Economics* 37, (2014): 169-179.

Cure, Morgane, Matthias Hunold, Reinhold Kesler, Ulrich Laitenberger, and Thomas Larrieu. “Vertical Integration of Platforms and Product Prominence.” *Quantitative Marketing and Economics* 20, no. 4 (2022): 353-395.

Etro, Federico. “The Economics of Amazon.” *Available at SSRN 4307213* (2022).

Etumnu, Chinonso E. “A Competitive Marketplace Or an Unfair Competitor? an Analysis of Amazon and its Best Sellers Ranks.” *Journal of Agricultural Economics* 73, no. 3 (2022): 924-937.

Farronato, Chiara, Andrey Fradkin, and Alexander MacKay. “Self-Preferencing at Amazon: Evidence from Search Rankings.” *NBER Working Paper 30894,* January, 2023. doi:10.3386/w30894. https://www.nber.org/papers/w30894.

Feasey, Richard and Jan Krämer. *Implementing Effective Remedies for Anti-Competitive Intermediation Bias on Vertically Integrated Platforms* Centre on Regulation in Europe asbl (CERRE), 2019.

Fletcher, Amelia, Gregory S. Crawford, Jacques Crémer, David Dinielli, Paul Heidhues, Michael Luca, Tobias Salz, Monika Schnitzer, Fiona M. Scott Morton, and Katja Seim. “Consumer Protection for Online Markets and Large Digital Platforms.” *Available at SSRN 3923588* (2021).

Fletcher, Amelia, Peter Ormosi, and Rahul Savani. “Recommender Systems and Supplier Competition on Platforms.” *Journal of Competition Law & Economics* (2023). doi:https://doi.org/10.1093/joclec/nhad009.

Friedman, Jerome H. “Greedy Function Approximation: A Gradient Boosting Machine.” *Annals of Statistics* (2001): 1189-1232.

Fry Chantal and Manna Sukanya. “Can We Group Similar Amazon Reviews: A Case Study with Different Clustering Algorithms” *2016 IEEE Tenth International Conference on Semantic Computing (ICSC)*, Laguna Hills, CA, USA, 2016, pp. 374-377, doi: 10.1109/ICSC.2016.71.

Gaddis, S. M. (2018). Audit studies: Behind the scenes with theory, method, and nuance. Vol. 14. Springer.

Gal, Michal S. and Nicolas Petit. “Radical Restorative Remedies for Digital Markets.” *Berkeley Tech.LJ* 36, (2021): 617.

Gautier, Axel and Nicolas Petit. “Optimal Enforcement of Competition Policy: The Commitments Procedure Under Uncertainty.” *European Journal of Law and Economics* 45, (2018): 195-224.

Gómez-Losada, Álvaro and Néstor Duch-Brown. “Competing for Amazon’s Buy Box: A Machine-Learning Approach.” Springer, 2019.

Gómez-Losada, Álvaro, Gualberto Asencio-Cortés, and Néstor Duch-Brown. “Automatic Eligibility of Sellers in an Online Marketplace: A Case Study of Amazon Algorithm.” *Information* 13, no. 2 (2022): 44.

Grinsztajn, Léo, Edouard Oyallon, and Gaël Varoquaux. “Why do Tree-Based Models Still Outperform Deep Learning on Typical Tabular Data?” *Advances in Neural Information Processing Systems* 35, (2022): 507-520.

Hoehn. “Challenges in designing and implementing remedies in innovation intensive industries and the digital economy.” In Gerard, Damien and Assimakis Komninos. *Remedies in EU Competition Law: Substance, Process and Policy* Kluwer Law International (2020).

Huber, Martin; Imhof, David. “Machine learning with screens for detecting bid-rigging cartels.” *International Journal of Industrial Organization* Vol. 65 (2019).

Hunold, Matthias, Ulrich Laitenberger, and Guillaume Thébaudin. “Bye-Box: An Analysis of Non-Promotion on the Amazon Marketplace.” *CRED WORKING PAPER no 2022-4* (2022).

Janger, Edward J. and Aaron D. Twerski. “The Heavy Hand of Amazon: A Seller Not a Neutral Platform.” *Brook.J.Corp.Fin.& Com.L.* 14, (2019): 259.

Marty, Frédéric and Reis, Patrice. “Perspectives juridiques et économiques sur les procédures négociées en droit de la concurrence.” *SciencePo Working papers* Main hal-00614942, HAL. (2011)

Metaxa, Danaë, Joon Sung Park, Ronald E. Robertson, Karrie Karahalios, Christo Wilson, Jeff Hancock and Christian Sandvig (2021), “Auditing Algorithms”, Foundations and Trends® in Human-Computer Interaction: Vol. 14, No. 4, pp 272–344. DOI: 10.1561/1100000083.

Mullainathan, S., M. Noeth, and A. Schoar. (2012). “The market for financial advice: An audit study.” Tech. rep. National Bureau of Economic Research.

OECD. *Remedies and Commitments in Abuse Cases*. OECD Competition Policy Roundtable Background Note (2022).

Policarpo, Lucas Micol, Diórgenes Eugênio da Silveira, Rodrigo da Rosa Righi, Rodolfo Antunes Stoffel, Cristiano André da Costa, Jorge Luis Victoria Barbosa, Rodrigo Scorsatto, and Tanuj Arcot. “Machine Learning through the Lens of E-Commerce Initiatives: An Up-to-Date Systematic Literature Review.” *Computer Science Review* 41, (2021): 100414.

Prokhorenkova, Liudmila, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. "CatBoost: Unbiased Boosting with Categorical Features." *Advances in Neural Information Processing Systems* 31, (2018).

Raval, Devesh. "Steering in One Click: Platform Self-Preferencing in the Amazon Buy Box." (2023) https://deveshraval.github.io/buyBox.pdf

Seele Peter, Dierksmeier Claus, Hofstetter, Reto *et al.* "Mapping the Ethicality of Algorithmic Pricing: A Review of Dynamic and Personalized Pricing." *J Bus Ethics* 170, 697–719 (2021). https://doi.org/10.1007/s10551-019-04371-w

Shapley, Lloyd S. *A Value for N-Person Games. Contributions to the Theory of Games*. (1953).

Shwartz-Ziv, Ravid and Amitai Armon. "Tabular Data: Deep Learning is Not all You Need." *Information Fusion* 81, (2022): 84-90.

Maitra, Subhradeep, Laxminarayan Sahoo, and Kalishankar Tiwary. 2022. "Study, Analysis and Comparison Between Amazon A10 and A11 Search Algorithm". *Journal of Computer Science Research* 4 (4):1-6. https://doi.org/10.30564/jcsr.v4i4.5111.

Sundararajan, Mukund and Amir Najmi. "The Many Shapley Values for Model Explanation." *Proceedings of the 37th International Conference on Machine Learning* (2020).

Waldfogel, Joel. Amazon Self-preferencing in the Shadow of the Digital Markets Act. No. w32299. National Bureau of Economic Research, 2024.

Wallimann, Hannes; Imohf, David; Huber, Martin. "A Machine Learning Approach for Flagging Incomplete Bid-Rigging Cartels." *Comput Econ* 62, 1669–1720 (2023). https://doi.org/10.1007/s10614-022-10315-w

Wils, Wouter PJ. "Settlements of EU Antitrust Investigations: Commitment Decisions Under Article 9 of Regulation no. 1/2003." *World Competition* 29, no. 3 (2006).

Wils, Wouter PJ. "Ten Years of Regulation 1/2003-a Retrospective." *Journal of European Competition Law & Practice* 4, no. 4 (2013): 293-301.

Yaghini, Mohammad; Liu, Patty; Boenisch, Franziska; Papernot, Nicolas. "Regulation Games for Trustworthy Machine Learning". (2024) https://arxiv.org/abs/2402.03540

DOCUMENTS DE TRAVAIL GREDEG PARUS EN 2024
GREDEG Working Papers Released in 2024

2024-01	Davide Antonioli, Alberto Marzucchi, Francesco Rentocchini & Simone Vannuccini *Robot Adoption and Product Innovation*
2024-02	Frédéric Marty *Valorisation des droits audiovisuels du football et équilibre économique des clubs professionnels : impacts d'une concurrence croissante inter-sports et intra-sport pour la Ligue 1 de football*
2024-03	Mathieu Chevrier, Brice Corgnet, Eric Guerci & Julie Rosaz *Algorithm Credulity: Human and Algorithmic Advice in Prediction Experiments*
2024-04	Mathieu Chevrier & Vincent Teixeira *Algorithm Delegation and Responsibility: Shifting Blame to the Programmer?*
2024-05	Maxime Menuet *Natural Resources, Civil Conflicts, and Economic Growth*
2024-06	Harald Hagemann *Hayek's Austrian Theory of the Business Cycle*
2024-07	Rami Kacem, Abir Khribich & Damien Bazin *Investigating the Nonlinear Relationship between Social Development and Renewable Energy Consumption: A Nonlinear Autoregressive Distributed Lag (ARDL) Based Method*
2024-08	Abir Khribich, Rami Kacem & Damien Bazin *The Determinants of Renewable Energy Consumption: Which Factors are Most Important?*
2024-09	Abir Khribich, Rami Kacem & Damien Bazin *Assessing Technical Efficiency in Renewable Energy Consumption: A Stochastic Frontier Analysis with Scenario-Based Simulations*
2024-10	Gianluca Pallante, Mattia Guerini, Mauro Napoletano & Andrea Roventini *Robust-less-fragile: Tackling Systemic Risk and Financial Contagion in a Macro Agent-Based Model*
2024-11	Sandye Gloria *Exploring the Foundations of Complexity Economics: Unveiling the Interplay of Ontological, Epistemological, Methodological, and Conceptual Aspects*
2024-12	Frédéric Marty *L'Intelligence Artificielle générative et actifs concurrentiels critiques : discussion de l'essentialité des données*
2024-13	Leonardo Ciambezi *Left for Dead? The Wage Phillips Curve and the Composition of Unemployment*
2024-14	Sophie Pommet, Sylvie Rochhia & Dominique Torre *Short-Term Rental Platforms Contrasted Effects on Neighborhoods: The Case of French Riviera Urban Destinations*
2024-15	Katia Caldari & Muriel Dal Pont Legrand *Economic Expertise at War. A Brief History of the Institutionalization of French Economic Expertise (1936-1946)*
2024-16	Michela Chessa & Benjamin Prissé *The Evaluation of Creativity*

2024-17 Supratim Das Gupta, Marco Baudino & Saikat Sarkar
Does the Environmental Kuznets Curve Hold across Sectors? Evidence from Developing and Emerging Economies

2024-18 Patrice Bougette, Oliver Budzinski & Frédéric Marty
Ex-ante versus Ex-post in Competition Law Enforcement: Blurred Boundaries and Economic Rationale

2024-19 Alexandre Truc & Muriel Dal Pont Legrand
Agent-Based Models: Impact and Interdisciplinary Influences in Economics

2024-20 Sanjit Dhami & Paolo Zeppini
Green Technology Adoption under Uncertainty, Increasing Returns, and Complex Adaptive Dynamics

2024-21 Adelina Zeqiri, Issam Mejri & Adel Ben Youssef
The Metaverse and Virtual Reality in Tourism and Hospitality 5.0: A Bibliometric Study and a Research Agenda

2024-22 Adel Ben Youssef, Mounir Dahmani & Mohamed Wael Ben Khaled
Pathways for Low-Carbon Energy Transition in the MENA Region: A Neo-Institutional Perspective

2024-23 Dorian Jullien & Alexandre Truc
Towards a History of Behavioral and Experimental Economics in France

2024-24 François Claveau, Jacob Hamel-Mottiez, Alexandre Truc & Conrad Heilmann
The Economics of JEM: Evidence for Estrangement

2024-25 Alexandre Truc, François Claveau, Catherine Herfeld & Vincent Larivière
Gender Homogeneity in Philosophy and Methodology of Economics: Evidence from Publication Patterns

2024-26 Flora Bellone, Arnaud Persenda & Paolo Zeppini
The Rise of China in the Global Production Network: What Can Autocatalytic Sets Teach Us?

2024-27 Jeanne Mouton & Benoit Rottembourg
Auditing the Ranking Strategy of a Marketplace 's Algorithm in the Frame of Competition Law Commitments with Surrogate Models: The Amazon 's Buy Box Case