Oxford Review of Economic Policy, 2024, **40**, 843–853
https://doi.org/10.1093/oxrep/grae048
Article

OXFORD

Artificial intelligence recommendations: evidence, issues, and policy

Emilio Calvano,*, Giacomo Calzolari, Vincenzo Denicolò,*** and Sergio Pastorello******

*Università LUISS and Einaudi Institute for Economics and Finance, Rome, Italy, Toulouse School of Economics, and CEPR, e-mail: ecalvano@luiss.it
**European University Institute, Bologna University, Italy, and CEPR, e-mail: giacomo.calzolari@eui.ei
***Bologna University, Italy, and CEPR, e-mail: vincenzo.denicolo@unibo.it
****Bologna University, Italy, e-mail: sergio.pastorello@unibo.it

This review paper builds on previous authors' work, including the La Fonte economic research blog of the EUI, available at https://lafonte.eui.eu/2024/03/14/sailing-with-artificial-intelligence-recommendation-systems-and-digital-markets/. Calvano acknowledges funding by the European Union (ERC grant AI-Comp, 101098332) and PRIN 2022 grant 'Algorithms and economic choices', codice Cineca 2022S5RC7R, CUP E53D23006420001. Denicolò and Pastorello acknowledge support from PNRR – M4C2 – (PE00000013) – 'FAIR—Future Artificial Intelligence Research' Spoke 8 'Pervasive AI', funded by the European Commission under the NextGeneration EU programme. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency or the Italian ministry of education. Neither the European Union nor the granting authority can be held responsible for them.

Abstract

Recommender systems (RS) enhance user experiences by providing personalized content and are widely used by popular services like Apple Music, Spotify, Netflix, and YouTube to increase user engagement. However, these systems can also have significant economic implications, including exacerbating market concentration and reducing content diversity. This paper reviews recent economic literature on RS, emphasizing their dual role as both beneficial tools and potential sources of market distortion. The paper underscores the necessity for policies informed by economic research to balance the benefits of RS against their associated risks.

Keywords: artificial intelligence, recommendation systems, search, price competition, platforms, regulation.

JEL codes: L41, L13, D43, D83

I. Introduction

Popular audio streaming services like Apple Music and Spotify offer catalogues with over 100 million tracks and approximately 5 million podcasts. Similarly, Netflix and YouTube provide thousands of hours of entertainment from a vast array of titles and user-generated videos.[1]

While these extensive catalogues ensure there is always something for everyone, finding suitable content can be challenging for users. The overabundance of choices may actually backfire, leading to decreased consumption and subscription cancellations, ultimately reducing both consumer surplus and the platforms' profits.

To address these issues, platforms increasingly use AI-powered algorithms known as recommender systems (RS). These algorithms provide users with personalized recommendations, such as curated playlists and tailored content displays, helping them navigate the vast array of available products and hence boosting consumption.

[1] For detailed information on the number of titles available on Netflix in various countries, visit this https://cordcutting.com/blog/how-many-titles-are-available-on-netflix-in-your-country/ and this https://newsroom.spotify.com/company-info/.

Recommender systems are now ubiquitous across various applications. For instance, Spotify and other music streaming platforms offer personalized playlists such as 'daily mixes'. Netflix uses a combination of tools to select and present specific titles to individual users from its vast catalogue, enhancing user experience and increasing viewership. Major retailers like Amazon and Zalando tailor product displays to individual browsing and purchase histories, as well as the behaviours of similar users. News outlets and aggregators like Google News personalize article recommendations by analysing users' viewing histories, content preferences, and behaviours. App stores like Google Play suggest apps based on download patterns, features, and ratings. Travel apps like Booking.com and Airbnb recommend accommodations and experiences. Financial platforms like Wealthfront use 'robo-advisors' to propose personalized investment portfolios.[2] Finally, social media platforms like Instagram, TikTok, and Twitter/X direct their subscribers' searches across the available content through various systems.

All these recommendations are generated by algorithms rather than humans. The platforms' recommender systems maintain user engagement by ensuring the content, whether personal posts or topics of interest, is relevant and tailored to individual preferences. This personalization is not merely a feature—it is the core product, indispensable to the existence of these services.

Recognizing the significant payoffs from even marginal improvements in the accuracy of the recommendations, firms have increased their investment in these systems. They hire machine learning and data science experts and establish research departments to collaborate with the broader academic community. Over the last few years, many well-known tech companies have emerged as major players in this arena, driving the scientific development and application of recommender systems.

This paper discusses some of the issues that these developments raise for economic analysis. Will markets become more or less concentrated with the diffusion of algorithmic recommendations? Will firms have incentives to increase or decrease their prices? How do algorithmic recommendations affect product choices, entry and exit decisions, and R&D investments? What is the final impact on consumer welfare?

To elaborate, consider the possibility that already popular items may receive even more attention and recommendations, thus becoming even more popular. This phenomenon can occur whenever a recommender system favours items with high initial engagement or positive feedback. It implies that items that become initially popular, whether due to some irrelevant motive or just by chance, may be promoted more often. This creates a cycle where these items continue to gain popularity while less popular items receive less attention, reducing diversity and fairness in recommendations. In many applications, this lack of diversity is problematic. For example, recommender systems that shape users' news consumption may reduce the number of news sources a user is exposed to. Such a process is alarming both at the individual level (as voters become less informed) and for society at large.

The possibility that algorithmic recommendations may carry negative effects has spurred a wide policy debate. Those concerned about the risks argue that the platforms' ability to shape choices carries responsibilities. Regarding specifically the economic consequences of the diffusion of recommender systems, the concern has been raised that algorithmic recommendations may lead to more concentrated markets, with potentially less choice and higher prices.

Economic research is indispensable for assessing the economic impact of recommender systems and designing policies to mitigate the risks associated with the diffusion of this technology. Ideally, we aim to understand how consumer choices change when made with and without the help of this technology.

This article discusses how to conduct this thought experiment using modern economic analysis tools. Section II provides a self-contained introduction to Recommender Systems in general and delves deeper into collaborative filters, a class of algorithms prevalent in many domains of interest. Section III presents empirical evidence showing that RS significantly shape consumption choices in reality. Section IV discusses how economists have conceptualized RS, contrasting various approaches proposed in the literature. In sections V and VI, we present our approach, testing actual state-of-the-art algorithms in controlled and synthetic environments to identify the impact of algorithmic recommendations on product market competition.

II. Recommender systems explained

RS are algorithms that analyse past usage data and user feedback to predict individual preferences for items users have not yet tried. By doing so, these algorithms can help match users with items they might enjoy. They

[2] For more information, see this https://www.nytimes.com/2018/09/07/business/roboadvisers-financial-planning-betterment.html

are sophisticated enough to process millions of personalized user experiences in real time, functioning effectively even in challenging environments where items, such as news articles or social media posts, are perishable and constantly changing and where user preferences are highly variable and may shift rapidly, sometimes within a single day.

RS exemplify a sector where the development of new techniques, innovations, and basic scientific knowledge is predominantly driven by private-sector companies rather than academia. This is because developing these algorithms requires vast amounts of data and resources, enabling private firms that possess these assets to innovate at scale. However, this implies that the architecture of the actual tools deployed in markets is typically kept secret. Scholarly examination of these tools must, therefore, rely on the knowledge disclosed in the scientific literature and the limited information that profit-maximizing firms themselves provide.

(i) Types of recommender systems

RS can be classified into four main types: content-based filtering, collaborative filtering, hybrid methods, and knowledge-based methods.[3]

Content-based filtering recommends items based on the attributes of the items and the user's previous interactions with similar items. This approach relies on understanding the characteristics of items a user has previously enjoyed to suggest similar ones. Key techniques in content-based filtering include *feature extraction*, which identifies and quantifies item attributes, and *similarity measures*, which compute the similarity between items using some measures, such as the Euclidean distance or the cosine similarity. Early systems used simple keyword matching but, over time, more sophisticated techniques such as advanced natural language processing (NLP) models have been employed to improve accuracy and relevance.

A notable feature of personalized recommendations derived from content-based filtering is their independence from other users' data. Consequently, content-based filtering is not reluctant to recommend new or unpopular items, making it less susceptible to the so-called *cold start* problem. However, this approach is limited by the quality and completeness of item attribute data. If the metrics used by the algorithms to classify items do not align with a specific user's preferences, the algorithm may struggle to provide surprising yet interesting recommendations.

Collaborative filtering, further explored in section II(iii), recommends items based on the preferences of similar users, assuming that users who have had similar preferences in the past will continue to do so in the future. It can be further specified into user-based collaborative filtering, which finds similar users and recommends items they liked, and item-based collaborative filtering, which finds similar items to those the user has liked and recommends them. Key techniques include *matrix factorization* methods like 'singular value decomposition' and neighbourhood-based methods like k-nearest neighbours.

The advantage of collaborative filtering is that it provides recommendations without requiring item attribute data. However, it suffers from the cold start problem and requires large user interaction data. Initially, simple nearest-neighbour methods were used, but introducing matrix factorization techniques, especially singular value decomposition, significantly improved performance. Modern systems often use advanced methods like deep learning to capture complex user-item interactions. In subsection (iii) we discuss these systems in more detail, as these systems have been studied in the experiments of Calvano *et al*. (2024*b*), which are described later.

Hybrid recommender systems combine content-based and collaborative filtering methods to leverage the strengths of both approaches while mitigating their weaknesses. Key techniques in hybrid methods include *weighted hybrid*, which combines the scores from different recommendation algorithms, *switching hybrid*, which switches between recommendation methods based on the context, and *feature combination*, which uses features from one method as inputs to another.

Hybrid methods can handle cold start problems better and provide more accurate and diverse recommendations. However, they come with increased complexity in system design and implementation and require more computational resources. Initially, simple combinations of content-based and collaborative filtering were used. Today, complex models such as neural networks, ensemble learning, and reinforcement learning are employed to create more robust hybrid systems.

Knowledge-based recommender systems use domain knowledge and user-specific information to make recommendations. They rely on explicit knowledge about how certain item characteristics meet user preferences rather than historical user behaviour. These systems often ask users specific questions to gather detailed preferences and requirements.

[3] See Aggarwal *et al*. (2016) for a detailed technical analysis of recommender systems.

This approach is particularly useful for complex products or services where user preferences are critical and can be explicitly defined, such as in real estate, finance, or travel planning. These systems provide tailored recommendations by understanding the relationship between user requirements and item attributes and by directly querying users.

(ii) Recent developments

Each of the aforementioned systems can be implemented in *reinforcement learning* mode, where the algorithm learns by optimizing long-term user engagement, continuously adapting based on user feedback and interaction data. The current frontier in recommender systems indeed involves leveraging deep learning techniques to improve recommendation quality.

Specifically, neural networks can capture complex patterns in user behaviour and item characteristics, with techniques like *convolutional neural networks* (CNNs), *recurrent neural networks* (RNNs), and *attention mechanisms* commonly used. Context-aware recommender systems incorporate contextual information, such as time, location, and user's device. These systems often use multi-modal data, combining text, images, and user context. Reinforcement learning models learn to make recommendations by optimizing long-term user engagement, continuously adapting based on user feedback and interaction data.

Recommender systems are also increasingly exploring the potential of leveraging large language models (LLMs) to enhance recommendation quality and user experience. LLMs, such as GPT-3 and GPT-4, can process and generate human-like text, making them well- suited for understanding and predicting user preferences based on natural language inputs. To identify nuanced patterns and trends, these models can analyse vast amounts of unstructured data, including user reviews, comments, and social media interactions. Integrating LLMs into recommender systems opens up opportunities to significantly advance the capabilities and effectiveness of modern RS.

There is also a growing focus on making RS more transparent and interpretable through explainable AI techniques, which help users understand why certain recommendations are made, increasing trust and satisfaction. Additionally, addressing biases in recommender systems is crucial for ensuring fairness, with techniques being developed to detect and mitigate biases, ensuring that recommendations are equitable and do not discriminate against any user group.

(iii) Collaborative filtering

We now provide more detail on a specific class of recommender systems, which will be referred to later on. As explained, RS are statistical tools meant to address the problem of matching items to users. To do so, it must first solve a prediction problem—how much a user may like an item not tried yet. For any user, the algorithm can then rank all the items and provide *personalized* recommendations, that is, show the highest-ranked items to each user.

Therefore, the first building block of collaborative-filtering systems is the dataset (rating matrix) that contains observed *ratings* of items effectively consumed by some users, but just for a few user/item combinations. Let the rating matrix be R, which is a $I \times J$ matrix where I is the number of users and J of items. The matrix contains some non-empty cells with observed ratings r_{ij}, i.e. the rating user i has left in the past for item j. In reality, R is very large (with millions of users and items) and very *sparse*, typically containing just 1–5 per cent of non-blank cells. To visualize the idea, consider the following simple example.

		Items			
		A	B	C	D
Users	1		4.5	2.0	
	2	4.0		3.5	
	3		5.0		2.0
	4		3.5	4.1	1.0

The task of the RS is to predict the missing ratings in the previous table (i.e. a matrix completion exercise) and then make personalized recommendations. A model-based collaborative filtering proceeds to the task using matrix factorization. In other words, the matrix of true ratings is seen as the product of a matrix of item characteristics and a matrix of user characteristics. Both matrices are unknown at the outset and must be estimated.

The underlying model of preferences has several *latent factors*, which do not need to have any specific semantic interpretation. For example, in the case of two factors, the previous matrix can be decomposed as follows:

$$
\begin{array}{cc|c|c|c|c|}
 & & \multicolumn{4}{c}{\text{Items}} \\
 & & \multicolumn{1}{c}{A} & \multicolumn{1}{c}{B} & \multicolumn{1}{c}{C} & \multicolumn{1}{c}{D} \\ \cline{3-6}
 & 1 & & 4.5 & 2.0 & \\ \cline{3-6}
\text{Users} & 2 & 4.0 & & 3.5 & \\ \cline{3-6}
 & 3 & & 5.0 & & 2.0 \\ \cline{3-6}
 & 4 & & 3.5 & 4.1 & 1.0 \\ \cline{3-6}
\end{array}
=
\begin{array}{c|c|c|}
\multicolumn{3}{c}{\text{User matrix}} \\ \cline{2-3}
1 & 1.2 & 0.8 \\ \cline{2-3}
2 & 1.4 & 0.9 \\ \cline{2-3}
3 & 1.2 & 1.0 \\ \cline{2-3}
4 & 1.4 & 0.8 \\ \cline{2-3}
\end{array}
\times
\begin{array}{|c|c|c|c|}
\multicolumn{4}{c}{\text{Item matrix}} \\
\multicolumn{1}{c}{A} & \multicolumn{1}{c}{B} & \multicolumn{1}{c}{C} & \multicolumn{1}{c}{D} \\ \hline
1.5 & 1.2 & 1.0 & 0.8 \\ \hline
1.7 & 0.6 & 1.6 & 0.4 \\ \hline
\end{array}
$$

The entries in the user matrix and item matrix are determined by the algorithm minimizing some function that represents a cost for the distance between the predicted rating and the observed rating (clearly for the user-item pairs where a past rating exists, i.e. the non-empty cells).

As the process clarifies, exploiting user and item correlations, the RS embeds a collaborative element among different users. This observation implies that RS is not just reducing users' search costs, which would instead rely only on idiosyncratic personal characteristics.

III. To what extent do recommender systems determine our decisions?

Anecdotal evidence underscores the significant influence that information from recommender systems has on consumption choices. For instance, MacKenzie *et al.* (2013) report that 75 per cent of the content viewed on Netflix and 35 per cent of the pages visited on Amazon are recommended to users by the platforms' algorithms. Although these figures may overstate the steering power of these recommendations—since some consumption choices would likely have been made without them—they nonetheless highlight their considerable impact.

Identifying the causal impact of recommendations using observational data is a challenging task. For instance, if we observe that items at the top of search result pages on a popular e-commerce platform are frequently purchased, it is difficult to determine how much of this demand is due to their top placement and how much would have occurred regardless because these items are inherently popular. To ascertain this, we need to observe a counterfactual scenario where the recommendation does not influence the outcome. One possible counterfactual could be a scenario where the same product is not featured at all. Other counterfactuals might include items being ordered alphabetically, by popularity, or curated by humans instead of algorithms.

Adopting the counterfactual of human recommendations, Peukert *et al.* (2023) assess algorithmic recommendation's economic benefits and information externalities. They ran a large-scale field experiment conducted with a major news outlet in Germany. They document a sizeable and significant increase in user engagement due to individually targeted recommendations relative to non-targeted professional human curation.

Aguiar and Waldfogel (2021) quantify Spotify's power to steer music consumption by using discontinuity and instrumental variable identification methods. Specifically, they show that including a song in a popular playlist curated by the service has a significant causal effect on consumption. They also show that recommender systems can increase the speed at which good songs become successful and can even create success out of thin air. This latter statement should, however, be taken with a grain of salt. The study focuses by design on songs that eventually make it into the playlist (or barely miss it). So, the songs included in the study were already, by design, selected among the most successful ones.

Empirical evidence on how product rankings influence Amazon shoppers' choices is provided by Farronato *et al.* (2023). They recruited a large number of Amazon customers and used a browser extension to record micro-level data on search results pages and consumer choices. The findings document that consumers seldom browse past the first couple of results, implying that being in one of the top spots is essential to be able to sell at all. Farronato *et al.* (2024) go one step further and manipulate the search results by replacing some products with products down the list. They show that promoted products sell at least as much as the removed ones would have done. Combining these pieces of evidence, one can conclude that manipulating search results may greatly impact consumption. The study specifically refers to Amazon's practice of placing its own brands at the top of those results pages (a conduct usually referred to as self-preferencing).

Finally, Lee and Musolff (2023) use high-frequency proprietary panel data on sales, prices, and recommendations on 200,000 products on Amazon.com to estimate a 'tractable structural model of intermediation power, i.e. a platform's ability to influence market outcomes by steering consumers'.

IV. Alternative views of recommender systems

The evidence reviewed above strongly suggests that RS may significantly impact market outcomes and, as such, should be carefully examined by regulators and antitrust authorities. In recent years, practitioners have developed various tentative theories of harm that identify potential negative impacts of algorithmic recommendations on competition and consumer welfare, such as reinforcing market dominance and distorting consumer choices.

However, more systematic economic analysis is needed to address the challenges posed by RS. Economics can provide a coherent conceptual framework for understanding the market impact of algorithmic recommendations, verifying the consistency of proposed theories of harm, and clarifying the type of empirical evidence needed to assess potential effects. Additionally, economic analysis can demystify the algorithm 'black box' by elucidating the underlying mechanisms and decision-making processes, revealing how they influence consumer behaviour and market outcomes.

A preliminary step for any formal economic analysis is to determine how to conceptualize this new technology. In this section, we discuss the different possibilities that have been considered in the existing literature.

(i) Lower search costs

The most popular approach conceptualizes recommender systems as tools that reduce search costs and information frictions. This perspective tends to view RS as pro-competitive, as the theoretical literature analysing the impact of search frictions on consumer choices and firms' incentives to price and position their goods typically finds that lower search costs are associated with more intense competition, lower prices, and higher consumer welfare. For instance, see the classic contributions of Wolinsky (1983, 1986), and Anderson and Renault (1999).

(ii) Information on product characteristics

A few papers, such as Aridor and Gonçalves (2022) and Calvano and Jullien (2018), model recommender systems as tools that provide consumers with information about product quality. The latter paper, in particular, views recommender systems as instruments used by intermediaries to strategically disclose information relevant to consumers' choices. It examines the incentives for providing such information in a context where the platform seeks to build a reputation for offering reliable recommendations.

(iii) Personalized prominence

While both of the aforementioned approaches capture significant changes associated with the development of information technologies, the most distinctive function of recommender systems appears to be their ability to present items to individual users in a specific order.

This order influences the sequence in which users inspect the available products. Adapting the terminology of Armstrong *et al*. (2009), who define prominence as the position items held in the users' search order, one can say that algorithmic recommendations create *personalized* prominence.

The provision of tailored and accurate recommendations significantly reduces the effort involved in the search process, leading to an increased surplus. This benefit arises from both faster and better choices, as consumers become more selective and compromise less. However, it is important to distinguish between the personalized-prominence view and the search-cost view of recommender systems. Both perspectives suggest that recommendations reduce total search costs. However, while the search-cost view emphasizes a reduction in unit search cost, the personalized-prominence view highlights increased efficiency in the search process. In the latter view, searches are no longer random but follow a pattern based on the algorithm's estimation of an item's value to individual users. As a result, consumers do not need to conduct as many searches as they would under a random search model, leading to a decrease in total search costs even if the unit search cost remains unchanged.

These views also have different implications for the impact of recommender systems on the intensity of competition. The search-cost view suggests that competition becomes more intense because consumers can search more and compare a greater number of products. In contrast, the personalized-prominence view implies that consumers will search less, but their searches will be more efficiently directed. The impact on the intensity of competition is uncertain and may even be negative, as we explain later.

The personalized-prominence approach has been advocated by Lee and Wright (2021), Castellini *et al*. (2023), and our own work Calvano *et al*. (2024*a*). These papers, however, differ in how they model consumer search and the benchmarks they use for comparison. Specifically, Lee and Wright (2021) evaluate the information value of recommendation algorithms by comparing them to purely random choices, whereas Castellini *et al*. (2023) use complete information as their benchmark. Our own work (2024) instead uses the search framework of Wolinsky (1986), as we describe in greater detail later.

(iv) Self-preferencing

While the approaches mentioned so far assume that platforms recommend items that, according to algorithmic predictions, best suit users' tastes, the power to steer choices comes with strings attached. Recommender systems are usually deployed and controlled by intermediaries with their own agendas and interests, which may conflict with those of the users. Hagiu and Jullien (2011) conducted a pioneering study examining the incentives for intermediaries to divert search using a stylized model in which consumers face search frictions. Follow-up studies Lee and Wright (2021) and De Corniere and Taylor (2019), while acknowledging the benefits of recommendations due to costly consumer search, explore the implications of allowing intermediaries to receive commissions from sellers.

(v) Estimation biases

Algorithmic recommendations can be misleading not only when manipulated by platforms but also when they are genuine. At their core, recommender systems are statistical tools that may deliver biased predictions, leading to poor outcomes. Biases can arise due to issues with the algorithms themselves (referred to as 'bias in the algorithm') or the data used to train them ('bias in the data').

But how can RS be inadequate for accurately estimating preferences? Aren't these 'intelligent' algorithms state-of-the-art? Indeed, they are. However, these tools must not only estimate preferences accurately but also be designed to operate at scale, personalizing feeds for millions or billions of users in real time. Consequently, they are 'simple' by design and, as such, are subject to prediction biases.

Bias in the data refers to various issues with the training data that could lead to suboptimal recommendations. For instance, data originates from our interactions and is therefore influenced by previous recommendations, creating a feedback loop. This loop can reinforce existing preferences, narrowing the selection of choices and exacerbating market dynamics. For example, a popular product on Amazon may become even more popular due to frequent recommendations above and beyond its merits, leading to a 'rich-get-richer' dynamic. Additionally, in a world where the success of creators' content or a firm's product is heavily influenced by algorithms, these algorithms shape economic incentives. Product characteristics, such as design and pricing (where applicable), are tailored to align with how algorithms assess and promote them, raising concerns that these systems might lead to unintended consequences such as market monopolization and the amplification of pre-existing biases.

(vi) Policies for recommender systems

Given the problems highlighted in the previous section, a heated debate has emerged regarding the need to regulate these tools, prompting policy-makers to take action. The European Union has taken a pioneering role in this area by enacting three major legislative frameworks: the Digital Services Act (DSA), the Digital Markets Act (DMA), and the Artificial Intelligence Act (AIA). These pieces of legislation collectively aim to enhance the fairness and competitiveness of digital markets while protecting and safeguarding consumer rights through a combination of legal obligations, prohibitions, oversight, and accountability measures. The DSA directly addresses the functioning of recommender systems, defining these systems as:

> fully or partially automated system(s) used by an online platform to suggest in its online interface specific information to recipients of the service or prioritise that information, including as a result of a search initiated by the recipient of the service or otherwise determining the relative order or prominence of information displayed.

The primary concern is that, without regulation, recommender systems could reinforce the market dominance of already powerful players, distort competition, and potentially harm consumer welfare by limiting choice. To tackle these issues, the DSA imposes several obligations on platforms, including the requirement that consumers be informed about how content is ranked and which of their data and behaviours influence this ranking. Furthermore, platforms are mandated 'to mitigate any negative effects arising from personalized recommendations' and adjust their criteria accordingly.

Another significant requirement is that platforms must provide access to their algorithms and training datasets to regulatory authorities, ensuring transparency and accountability. Efforts to minimize algorithmic bias and discrimination are also mandated, emphasizing the need for platforms to design their recommender systems with fairness in mind. Additionally, the DSA stipulates that users must have the ability to alter how content is ranked and must be offered at least one option for a recommender system that is not based on profiling. The design, logic, functioning, and testing of these systems must be thoroughly explained and documented.

The Artificial Intelligence Act (AIA), particularly in Article 5, goes further by prohibiting the use of AI systems that employ manipulative techniques or exploit vulnerabilities. This provision has significant implications for recommender systems, which are often designed to influence user choices. The act raises critical questions about the legality of systems that steer consumer behaviour to extract economic surplus.

Enforcing these broad principles requires the combined effort of economists, computer scientists, law scholars, and practitioners. In particular, economic research should contribute to policy solutions by proposing regulatory frameworks and market interventions that promote fair competition and protect consumer interests. For instance, policy solutions to tackle algorithmic bias should incentivize firms to limit the potential damage by investing in extensive pre-release testing and sandboxing and improving the transparency of recommendation algorithms. Issues due to biased data may be mitigated in various ways, such as by adding more data and imposing data-sharing provisions. In the following, we illustrate an economic research approach that suits these purposes and an example of policy design that follows from it.

V. Gauging benefits and pitfalls: an experiment

In recent research, Calvano *et al.* (2024*a*), we propose an approach to analyse the complex interactions between AI algorithms and markets, aiming to identify the drivers of undesirable outcomes. This is challenging to achieve through empirical research due to data limitations, and theoretical studies struggle with the complexity of modelling AI algorithms.

We conducted a numerical analysis of a model involving a large number of consumers choosing among substitute products. Consumers engage in costly searches and may or may not receive assistance from algorithmic recommendations. When available, these recommendations influence the order in which consumers conduct their searches.

We use collaborative filtering, latent-factors algorithms, and model consumer preferences so that the algorithms' estimates are based on an accurate model of the underlying preferences. However, the estimates remain imperfect due to the limited data available to the algorithms.

Our modelling strategy is based on the following sequence of logical steps:

(i) First, we construct an economy populated by many buyers with preferences over a large number of products. These preferences are flexible enough to generate various scenarios of interest, including horizontal and vertical product differentiation.
(ii) Second, we build a sparse rating matrix R based on buyers' *true* preferences.
(iii) Third, we train a collaborative-filtering algorithm specifically coded for the purposes of the research to predict preferences, and we use it to make personalized recommendations.
(iv) Finally, we numerically simulate consumer choices with and without the assistance of algorithmic recommendations.

To illustrate, consider many consumers who need to purchase a 'California travel guide' from a given set of offers. Consumers are endowed with diverse preferences that reflect their idiosyncratic needs (i.e. adventurous, family-oriented, foodies...) and reflect the different idiosyncratic qualities of those guides (top quality, first-hand guides crafted by someone present in the destination, or lower quality guides crafted using material available in the public domain). In the benchmark, consumers rely only on search: sequentially inspect guides drawn at random bearing a constant cost per product inspected. In the treatment, consumers are assisted by an RS trained to predict their tastes. The RS pre-selects one item for each consumer and presents it to him or her. Consumers then decide whether to buy the recommended item or search for a better match.

When consumers search, they adopt an optimal stopping rule, as in Wolinsky (1986) and Anderson and Renault (1999). The unit search cost is calibrated so that a prespecified fraction of users follows the recommendation.

The experiment is designed to present a problem of statistical preference estimation that mimics the key dimension of real-world problems. For example, the ratio of users over items, as well as the number of parameters to estimate relative to the number of observations and the density of the rating matrix, are similar to real-world problems and data sets. Our data set is generated either using *exogenous data* that is, the matrix R is built with ratings of randomly chosen items per-user (rating reported into R with noise), or with *endogenous data*, to assess the feedback-loop hypothesis.[4]

(i) Results

We find that recommender systems are more effective at making decisions for users than users themselves. Specifically, the expected utility a user can obtain by participating in a market with the help of an RS can increase

[4] In the case of endogenous data, we initialize R with few random ratings per-user, make recommendations, that users follow, they report rating (with noise), thus populating R over time. This process progresses till we reach the target density as in the case of exogenous data.

by as much as 6 per cent compared to a benchmark where the consumer must autonomously perform a costly search for preferred items. Algorithmic recommendations save time and search costs. Additionally, they enhance users' experiences by directing them to items they might otherwise never have discovered.

On the other hand, we also examine the impact of RS on markets. We find that these systems can significantly increase market concentration, turning certain products into 'superstars'. Additionally, sellers tend to increase their prices when consumer behaviour is mediated by an RS, with algorithmic recommendations leading to price increases of up to 16 per cent.

Our approach enables us to delve further into the underlying causes of these significant effects of RS. Specifically, we can identify that 'algorithmic demand'—meaning the demand expressed by consumers assisted by RS—differs markedly from 'human demand'. This difference is illustrated in Figure 1.

By comparing the case of exogenous data with that of endogenous data generated by RS, we find that data endogeneity does influence market outcomes, for example by further increasing market concentration, but this effect is secondary.

In summary, there are both positive and negative effects on consumer welfare. Even if the positive effects tend to prevail, these findings raise legitimate concerns. Specifically, we uncover an inverted U-shaped relationship between the amount of information RS use and consumer welfare. As RS access more data, consumer welfare initially improves due to better matching of products to consumer preferences. However, beyond a certain point, the negative effect of price increases outweighs these benefits. This suggests the need for a balanced approach to data utilization by RS and highlights potential areas for regulatory intervention.

Finally, the approach we have described allows us to investigate the potential for platforms to manipulate recommendations for profit. This is particularly concerning when the platform deploying the RS also sells products that consumers might choose, creating a conflict of interest. While such manipulation could have detrimental effects, our research indicates that it also tends to lower prices for favoured and over-recommended products, thereby mitigating some of the adverse impacts on consumer welfare. This finding opens new routes for understanding the strategic behaviours of platforms and their implications for market competition.

VI. Tampering recommendations with information

As discussed, recommender systems are currently at the forefront of a critical policy debate, which includes discussions on restricting the scope of data they utilize and disseminate. Limiting the amount of information RS

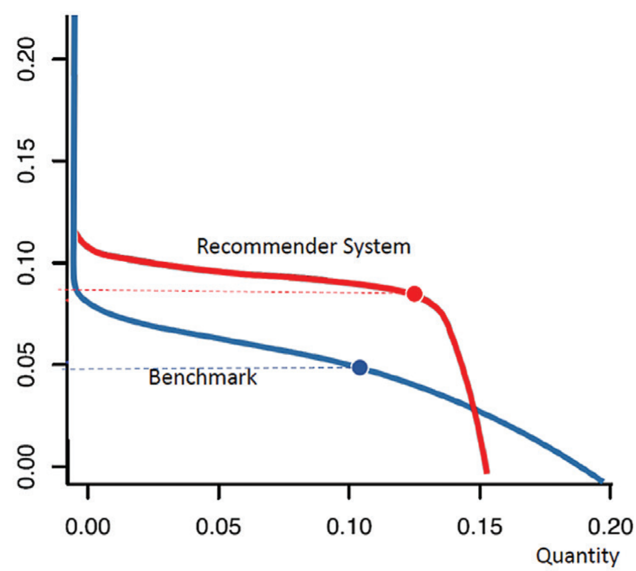

Figure 1: Algorithmic and human (inverse) demand

Notes: The red line, labelled 'recommender system', is the 'algorithmic demand' mediated by the recommender system, the blue line, labelled 'benchmark', is the 'human demand', dots indicate the equilibrium prices for this item.

Source: Authors' calculations.

have access to may increase consumer surplus because such limitations may lower prices more than they degrade the matching. This possibility raises the issue of whether consumers may benefit from specific regulations that somehow restrain algorithms. This is a matter related to the broader policy debate on the desirability of regulating AI and constraining its capabilities.

In a related paper, Calvano *et al*. (2024*b*), we examine a specific form of AI regulation that can be employed in this setting, namely a policy that mandates the platform to make the same recommendations to all consumers instead of offering personalized recommendations. In other words, while the RS would still be capable of identifying the products that could be the best for each consumer, the platform would be constrained to recommend the same products to all of them. This would transform *personalized* prominence into *uniform* prominence.

For example, Article 38 of the 2022 European Digital Service Act (DSA) states that large platforms' recommender systems 'shall provide at least one option for each of their recommender systems which is not based on profiling'. While this provision does not rule out the possibility that the platforms also make personalized recommendations, it clearly aligns with the direction considered in Calvano *et al*. (2024*b*).

Following the implications of this policy, we find that this regulation reduces prices, but it worsens the matching between consumers and products to such an extent that consumer surplus decreases. Although a uniform RS may appear beneficial due to reduced prices, this advantage is offset by a decrease in match value, culminating in an overall negative impact. Quantifying these effects in a realistic environment, the paper provides insight crucial for policy-makers contemplating specific policies. Prohibiting platforms from personalizing their recommendations does not seem to improve consumer welfare.

VII. Concluding remarks

Recommender systems are akin to a double-edged sword. They are powerful tools that enhance our digital experiences, but their broader implications on markets and consumer welfare necessitate careful scrutiny. Moulding the role of RS in our lives becomes a technological quest and a societal duty. This paper provides a first account of the recent economic literature on this topic, a first step in a longer expedition to decode the intricate influence of AI on our existence.

Future research should focus on examining the long-term consequences of recommendation systems. Investigating how such systems influence market entry and exit strategies, product innovation, and overall market competitiveness would provide deeper insights. Additionally, there is a pressing need for evidence-based policies to ensure that deploying these systems benefits consumers and the market as a whole. This involves rigorously assessing the potential for market concentration, the impact on sellers, and the overall diversity of available products. Policymakers must rely on solid evidence to craft regulations that balance the advantages of personalized recommendations with the necessity of maintaining fair competition and consumer choice. Integrating findings from economic research and real-world data will be crucial in developing these policies, ensuring that they are both effective and adaptable to evolving markets and technologies.

References

Aggarwal, C. C., *et al*. (2016), *Recommender Systems,* vol. **1**, Springer.

Aguiar, L., and Waldfogel, J. (2021), 'Platforms, Power, and Promotion: Evidence from Spotify Playlists', *Journal of Industrial Economics*, **69**(3), 653–91.

Anderson, S. P., and Renault, R. (1999), 'Pricing, Product Diversity, and Search Costs: A Bertrand–Chamberlin–Diamond Model', *RAND Journal Economics*, **30**(4), 719–735.

Aridor, G., and Gonçalves, D. (2022), 'Recommenders' Originals: The Welfare Effects of the Dual Role of Platforms as Producers and Recommender Systems', *International Journal of Industrial Organization*, **83**.

Armstrong, M., Vickers, J., and Zhou, J. (2009), 'Prominence and Consumer Search', *The RAND Journal of Economics*, **40**(2), 209–33.

Calvano, E., and Jullien, B. (2018), 'Can We Trust the Algorithms that Recommend Products Online? A Theory of Biased Advice', mimeo.

— Calzolari, G., Denicolo, V., and Pastorello, S. (2024*a*), 'Artificial Intelligence, Algorithmic Recommendations and Competition', *Algorithmic Recommendations and Competition*, 14 May.

— — — — (2024*b*), 'Uniform Algorithmic Recommendations', mimeo.

Castellini, J., Fletcher, A., Ormosi, P. L., and Savani, R. (2023), 'Recommender Systems and Competition on Subscription-based Platforms', in SSRN Wo. [Sl]: SSRN.

De Corniere, A., and Taylor, G. (2019), 'A Model of Biased Intermediation', *The RAND Journal of Economics*, **50**(4), 854–82.

Farronato, C., Fradkin, A., and MacKay, A. (2023), 'Self-Preferencing at Amazon: Evidence from Search Rankings', https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4331880.
— — — (2024), 'Platform Vertical Integration and Consumer Choice: Evidence from a Field Experiment', Working Paper.
Hagiu, A., and Jullien, B. (2011), 'Why Do Intermediaries Divert Search?', *The RAND Journal of Economics*, **42**(2), 337–62.
Lee, G., and Wright, J. (2021), 'Recommender Systems and the Value of User Data', *National University of Singapore Working Paper*, https://836d83c2-c629-40a4-9ca7-b1bf324d720d.filesusr
Lee, K. H., and Musolff, L. (2023), 'Entry into Two-sided Markets Shaped by Platform- guided Search', https://economics.yale.edu/sites/default/files/jmp_entry_into_two-sided_markets_shaped_by_platform-guided.pdf, accessed 23 May 2023.
MacKenzie, I., Meyer, C., and Noble, S. (2013), 'How Retailers Can Keep Up with Consumers', *McKinsey Retail Insights*.
Peukert, C., Sen, A., and Claussen, J. (2023), 'The Editor and the Algorithm: Recommendation Technology in Online News', *Management Science*.
Wolinsky, A. (1983), 'Retail Trade Concentration Due to Consumers' Imperfect Information', *The Bell Journal of Economics*, 275–82.
— (1986), 'True Monopolistic Competition as a Result of Imperfect Information', *The Quarterly Journal of Economics*, **101**(3), 493–511.