Dual Learning: How and How Much Can Platforms Learn from Searching Consumers?*

Maarten C.W. Janssen† Eeva Mauring‡

June 25, 2025

Abstract

Consumers search on a platform to find a product they like. The platform observes which products consumers inspect and buy. Based on these observations it ranks products to maximally learn, in the long term, which product consumers like. We find that a monopoly platform first experiments with rankings and later only ranks products that early consumers bought. This guarantees that later consumers are pickier, helping the platform to learn what consumers really like. The more dissimilar consumer tastes, the more consumers search themselves and the platform learns about products. Competition restricts what platforms learn.

JEL Classification: D40, D83, L10
Keywords: Online platforms; learning; consumer search; product rankings; steering.

*We have benefited considerably from comments by Mark Armstrong, Heski Bar-Isaac, Christoph Carnehl, Jiawei Chen, Michael Choi, Marc Goni, Joao Montez, Andrew Rhodes and Cole Williams and by participants to the virtual CEPR IO seminar, the 2024 consumer search workshop (Istanbul), the 2025 Industrial Committee workshop of the Verein für Socialpolitik, the Durham workshop on Theory and Policy in the Digital Economy, Micro Theory Day at NHH, BECCLE Conference 2025 and seminars in Budapest, Carlos III Madrid, Lausanne, Toulouse, UC Irvine, and UCLA Anderson.

†University of Vienna and CEPR. E-mail: maarten.janssen@univie.ac.at.
‡University of Bergen and CEPR. E-mail: eeva.mauring@uib.no.

1 Introduction

Online platforms possess large amounts of data on items people have searched for. This can be data about products, services, news or other items. Platforms also observe whether or not consumers buy a product or service or how long a user spends on a news item (as an indication of whether the user appreciates the item). Platforms may employ this information to help users find items they like, but there are concerns that platforms use this information in ways that do not (directly) benefit users. Regulators have put rules in place to restrict the ways platforms can use the information they collect; see, e.g. the EU's Digital Markets Act (DMA) and the Digital Services Act (DSA).

What platforms can infer from the information they collect does not only depend on the search behaviour of their users, but also on how they themselves rank the items. On the one hand, if they always present the same ranking of items, users will make similar choices, restricting the type of inferences platforms make as it is unclear whether choices reflect "true preferences" or (mostly) the platform's own ranking. By experimenting with different rankings, and observing behaviour given different rankings, platforms potentially may learn more which products users really like. On the other hand, if platforms only provide random rankings, users will not expect to find high-value items on their next searches and may be easily satisfied with what they are presented with. Thus, how much platforms learn depends on users' behaviour. However, optimal user behaviour in turn depends on what users expect to observe on their next searches given their current observations and their expectations regarding the platform's ranking. This paper studies this process of dual learning.

In particular, we analyze in this paper how and how much a platform can learn by optimally choosing its ranking. We consider that a platform wants to learn as much as possible about which item consumers like best in the long run. Such an objective may be implied, for example, when a platform's main revenue source stems from selling information to interested parties, or from advertising. This objective is also relevant if the platform is interested in user well-being to enlarge the market or to increase its market share. We compare how much can be learned both in case a platform is a monopolist and when it faces competition.

We focus on online markets where consumers search for and buy products, but the ideas of the paper can equally be applied to news aggregators or recommendation platforms. Consumers' value for products has a common and an idiosyncratic component. Platforms observe which products consumers inspect and buy. The platforms choose an algorithm that maps the observed history into a ranking of products for the next consumer. Consumers inspect products sequentially and their behaviour is characterized by an optimal stopping rule that, because of consumer learning, may well be different from a classic reservation value strategy (as in Weitzman, 1979). Consumers can follow

the ranking if they like, but are free to search on the platform in any order they prefer. Consumers sequentially arrive at the platform. Platforms do not observe the actual value a consumer attaches to a product. They only know whether a consumer inspects a product and whether it is bought. As each consumer only uses the platform once, platform learning is about the common value component of consumers' preferences.

Our main result for a *monopoly platform* is a characterization of the maximal amount of learning and a construction of an algorithm that reaches this maximum. The result shows that by experimenting with rankings on early users the platform can learn much more about the common value of the best products than by not experimenting. We say that a platform experiments if it consciously chooses not to provide the best possible ranking for the current consumer using the current information.

To see what experimentation brings, consider first that the idiosyncratic component is absent so that all consumers have the same value for products. *Without experimentation*, the platform would suggest to all future consumers the item that was bought first, which is a product that has a (common) value that is at least equal to the consumers' ex ante reservation value, as in Wolinsky (1986). Subsequent consumers would always buy the same product as they expect to inspect a random product on their next search. *With experimentation*, the platform can hide for a first set of consumers the products that previous consumers have bought by randomizing their ranking. It can gradually shift to showing consumers a ranking where products that previous consumers bought are placed first. Doing so increases later consumers' continuation value of search as they believe that if they continue to search they are more likely to see products with a high value compared to the situation where the platform does not experiment. Thus, later consumers expect that they are not presented with a random set of products if they continue to search, become more picky, and will reject some products that would have been acceptable if they had expected random products on their next search.

We define *a maximum threshold value under experimentation* that is such that the consumer is indifferent between immediately accepting the threshold value and continuing to search expecting that the next item to be searched has a value that is at least as high as the threshold value. We show that by experimenting the platform will ultimately find items with a value larger than this threshold and all consumers immediately buy the first item that is prominently ranked. Moreover, this threshold value is considerably larger than consumers' ex ante reservation value. However, for any positive inspection cost, there is a limit to what the platform learns and not only the very best products are ranked first. In particular, the platform learns less if the inspection cost is larger.

When there is an idiosyncratic component to consumers' value, the larger its weight the better the platform is able to learn which products have the highest common values. With an idiosyncratic component, consumers have a larger incentive to search. In particular, some consumers may be unhappy with products the platform has ranked high,

simply because they draw a low idiosyncratic component of their match value. The larger the weight of the idiosyncratic component, the more likely consumers continue to search themselves. For weights below a critical value, we show that it remains true that the platform ultimately learns that a prominently placed item has a high enough common value such that all consumers immediately buy it.

For weights above this critical value, the platform's problem changes its character. While at lower weights of the idiosyncratic component the platform's main issue is to experiment with rankings early on to be able to make later consumers more picky, at higher weights the issue is when to keep on presenting certain items on more prominent spots (that consumers inspect first) and when to switch to alternative items. Using ideas from statistical hypothesis testing, we show that the platform is able to learn that a product has a common value component that is arbitrarily close to the highest possible value. We also show that the monopoly results are robust to consumer heterogeneity in terms of search costs and the weight they attach to their idiosyncratic component, and to a fraction of consumers or all of them having the option to buy blindly (i.e., without inspecting the product first).

Our second set of results pertains to competing platforms. We show that if platforms want to maximize the extent of learning on their own platform, competition potentially limits the possibility for platforms to experiment, especially *when consumers do not know platforms' market shares* (or how many previous consumers have used which platform). The reason is best understood by realizing that by experimenting, the monopoly platform consciously trades off the benefit of present consumers in favour of future consumers. Competition, however, implies that when they have a choice consumers prefer, *ceteris paribus,* to use a platform that is not experimenting, leaving platforms that do experiment without consumers to experiment on. Proposition 9 formalises this intuition and shows that under competition platforms do not experiment and learn less than under monopoly. *When consumers know market shares*, they know that the platform with the largest market share is able to provide better recommendations, and prefer to use this platform even if it experiments a bit more than others. We show that by experimenting just so much that all consumers prefer to use the platform with the largest market share, the platform can learn in the long run as much as under monopoly, supporting the argument that there is a data-driven tendency for monopoly power in recommendation markets.

Competition may also come in the form of a platform entering the market after an incumbent already has served consumers. We show that a more efficient entrant that has created a more convenient website design where consumers have lower search costs may not be able to gain market share, because the incumbent platform already has been able to experiment with consumers until entry took place. Consumers may therefore prefer to use the incumbent platform as they are more likely to get better recommendations even if that implies that they incur higher search costs.

Related Literature. The environment we study has important similarities and differences with the social learning literature (Banerjee, 1992; Bikhchandani et al., 1992; Smith and Sørensen, 2000, among others). In the classic social learning papers agents have the same preferences and decide whether or not to make a particular choice. Apart from their prior information, in most of these papers they also learn by directly observing the actions of all past agents. In Kremer et al. (2014) and Glazer et al. (2021), agents do not observe the choices of preceding agents, but a principal does and recommends to future agents choices based on this information. In contrast to these papers, in line with online markets, consumers in our paper can choose one out of many products and a platform *communicates* its information through a *ranking* of these products. This type of communication (and choice) is very different from advising a consumer whether or not to adopt one product. Another important difference is that in our paper consumers have to pay an inspection cost if they want to learn their value for a product and they are not restricted to search only a limited set of products. Consumers may decide to stop searching and buy a product simply because they believe it is "good enough" in the sense that they do not want to inspect another unknown item at an inspection cost. If there is information acquisition in the social learning literature, as in Glazer et al. (2021), it is limited to one signal about one option. The fact that there are many products that are costly to inspect impose natural limits on what a platform may learn. Finally, consumers' (agents') preferences include an idiosyncratic component, so that what is a good choice for one consumer may not be a good choice for another consumer in the same situation. Maglaras et al. (2021) studies an environment where consumers learn from reviews of previous consumers and pay a larger "search cost" for an item that is ranked lower by the platform. There is no real search in their model, however, as consumers simply buy the product that maximizes their expected utility before inspecting. Goeree et al. (2006) study an (otherwise classic social learning) environment where the agents' utility has both an idiosyncratic and a common part. They find that in contrast to the herding result of the social learning literature, with an idiosyncratic component to their utility function, agents will eventually always learn the true state of the world. In contrast, our paper focuses on platforms trying to influence the outcome of the social learning process and on consumers searching different alternatives. Our paper also relates to other papers on platforms that recommend products to consumers and learn about product quality.[1]

The interaction between a platform's rankings and consumer search behaviour is also studied from an empirical angle (see, e.g., De los Santos and Koulayev, 2017; Ursu, 2018). The main issue in that literature is how to estimate the impact of a platform's ranking on what consumers click on or what they buy as the ranking itself is also influenced

[1]A literature also exists, such as Teh and Wright (2022), Janssen and Williams (2024), Nocke and Rey (2024), Bar Isaac and Shelegia (2025), and Janssen et al. (2023), where a firm, platform or social influencer with some knowledge steer consumers to products, but there is no learning.

by consumers' search behaviour. The question in this paper, namely how a platform may choose rankings to learn consumer preferences and how that depends on the market structure, is not addressed in the empirical literature, however.

There are two parts of the consumer search literature this paper relates to. First, the literature on directed and ordered search, and the effects of prominence (see, e.g., Weitzman, 1979; Arbatskaya, 2007; Armstrong et al., 2009; Wilson, 2010; Haan and Moraga González, 2011; Rhodes, 2011; Zhou, 2011; Armstrong, 2017; Choi et al., 2018, for early papers on these topics) focusses on the effect of prominence or search order on prices, whereas our focus is on how platforms can use rankings to learn about consumer preferences. Second, the literature on search and learning, initiated by Rothschild (1974) and further developed by Benabou and Gertner (1993) and Janssen et al. (2017), among others, characterizes optimal search behaviour when it is not characterized by a simple reservation value rule, as in Weitzman (1979), because consumers learn when searching about the underlying distribution of product characteristics. In our setting, the platform's ranking algorithm allows consumers to make inferences about the goodness of fit of uninspected products on the basis of the value of the products they already observed.

There is also a rapidly growing literature on how platforms can use their knowledge about consumers searching for keywords to create rankings that maximize their profits (see, e.g., Anderson and Renault, 2025; Ke et al., 2022; Janssen et al., 2023). To the best of our knowledge, our paper is the first to develop a framework analyzing how a platform should adjust its rankings to best learn from consumer choices.

More generally, our paper relates to the literature on dynamic information design (see, for example, Ely, 2017; Renault et al., 2017; Ely and Szydlowski, 2020; Orlov et al., 2020; Smolin, 2021) and the strategic experimentation literature (see, e.g., Bolton and Harris, 1999; Keller et al., 2005). In Hagiu and Jullien (2011), a profit-maximizing platform sometimes makes consumers search more than they would like (in order to influence sellers' behaviour), but there is no learning.

The rest of the paper is organized as follows. The next Section describes the monopoly model, while Sections 3 and 4 state the main monopoly results and their extensions. Section 3 also contains an example of a ranking that induces an optimal search rule that is not characterized by a reservation value. Section 5 discusses the results on competition. Section 6 concludes with a discussion, while proofs can be found in the Appendix.

2 The Monopoly Model

We consider a population of infinitely many consumers using a platform to search for a product. Products are horizontally and vertically differentiated, i.e., they have features that all consumers value in the same way (a common component) and other features that consumers value differently (an idiosyncratic component). Thus, a consumer t's utility

for product j is denoted by

$$u_{tj} = (1-\delta)v_j + \delta m_{tj},$$

where v_j is the common ("value") component, m_{tj} is the idiosyncratic ("match-specific") component and $\delta \in [0,1]$ is the relative weight on the idiosyncratic component. Before inspecting a product, the consumer is uncertain about both components of the product. The common component v_j is a random draw from the CDF $G(\cdot)$ with support $[\underline{v}, \bar{v}]$, while the idiosyncratic component m_{tj} is a random draw from the CDF $F(\cdot)$ with support $[\underline{m}, \bar{m}]$; v_j and m_{tj} are independently distributed. We assume that $f(\cdot)$ and $g(\cdot)$ are strictly positive and finite over the whole support and that $1-F$ and $1-G$ are logconcave (as is usual in the consumer search literature; see, e.g., Anderson and Renault, 1999). For a given δ denote the (derived) distribution of utilities by $H_\delta(u)$, defined over $[\underline{u}, \overline{u}]$ with $\underline{u} = (1-\delta)\underline{v} + \delta\underline{m}$ and $\overline{u} = (1-\delta)\overline{v} + \delta\overline{m}$.

The consumer has to pay an inspection (or search) cost s to observe the utility u_{tj} that product j gives. All consumers have the same search cost s and have to inspect the product before buying.[2] The consumers choose an optimal sequential search strategy and stop searching when the utility exceeds the value of continuing to search. To avoid an uninteresting analysis where consumers never inspect products if the products are randomly ordered, we take s to be small enough, i.e., $s < \int_{\underline{m}}^{\bar{m}}(1-F(m))dm = Em - \underline{m}$ and $s < \int_{\underline{v}}^{\bar{v}}(1-G(v))dv = Ev - \underline{v}$. We sometimes illustrate our results for m and v being uniformly distributed on $[0,1]$, and for this case this implies that $s < \frac{1}{2}$.

Consumers arrive at the platform sequentially and we index consumers by $t = 1, 2, ...$ Consumers know their place in the queue. For each consumer, the platform observes which products a consumer inspects and which product they eventually buy, but it does not know a consumer's utility score. Based on these observations, the platform can make an inference about the consumer's utility of products that are inspected and (not) bought. For each consumer, it can create a ranking of products and this ranking may influence the order in which consumers search. For example, if a product is bought by a previous consumer, then the platform may (or may not) rank that product first, and a first-ranked product may be interpreted by consumers as a product that the platform thinks they may like. Platforms may rank the same item multiple times. Consumers know[3] the ranking algorithm the platform uses and recognizes if multiple copies of the same item have been inserted in the ranking without inspecting them.[4] They also know that other consumers

[2] In extensions, we consider that consumers have different search costs and have the option to buy "blindly", i.e., without inspecting the product first.

[3] If none of the platform's users are informed about the ranking algorithm, the outcome fully depends on consumer expectations. For example, if all consumers expect the platform to rank the products randomly, then the platform cannot learn much. The DSA requires that platforms are transparent about the ranking algorithm they use so that users are, in principle, able to inform themselves of the algorithm used (see the Digital Services Act, paragraph (70), European Parliament, 2022b).

[4] Copies are used by the platform to restrict the number of ways consumers may learn during their search process (see the proof of Proposition 5). Consumers may immediately recognize copies of items

are looking for products on the platform, and that therefore the ranking of products may be informative of what others have inspected and bought in the past. The platform aims to find an item with as high a common component as possible in the long run (and rank it first).[5] That is, it may experiment with rankings for some finite number of consumers without impacting the value of its objective function.

To be more precise, let products have a number $n \in \mathbb{N}$ and denote by N the set of all subsets of $\mathbb{N}$. Let $S(t) \in N$ be the sequence of products that a consumer t has inspected and by $b(t) \in \mathbb{N} \cup \varnothing$ the purchase decision of that consumer. For every consumer t the platform chooses a ranking r_t, which is a function $r_t : \{S(\tau), b(\tau)\}_{\tau=1}^{\tau=t-1} \times Z \to X$, where Z is the realization of a random device and X denotes the set of product permutations of all lengths. The overall ranking algorithm r of the platform is a sequence of those rankings: $r = \{r_t\}_{t=1}^{\infty}$. Let $v_k(t)$ denote the common component of the kth-ranked item in the ranking for the tth consumer. The monopoly platform's objective is to maximise

$$lim_{t\to\infty} Ev_1(t).$$

The timing is as follows.[6] At $t = 0$, the platform chooses and announces its ranking algorithm r. At $t = 1, 2, ...$, the algorithm generates ranking r_t, consumer t arrives, searches through ranking r_t in an optimal order, follows the optimal stopping rule to buy the highest-utility product searched or exits. Then, the next period arrives.

Define for every ranking algorithm r a dynamic process where all consumers t search optimally, believe that all consumers $1, ..., t-1$ before them searched optimally, and consumers' beliefs are consistent with ranking algorithm r. Thus, one can calculate the value of the platform's objective function along this dynamic process. An optimal algorithm r^* is an algorithm that maximizes this function. Alternatively, one can define a Perfect Bayesian Equilibrium (PBE) where the platform's strategy is the choice of its algorithm and consumers' strategies are their search strategies. The above procedure selects the platform's optimal PBE.

2.1 Preliminary analysis: No experimentation

We say that a *platform experiments in period t* if it consciously does not provide the best possible ranking for consumer t using all information available to the platform in period t. We say that a platform *does not experiment* if it does not experiment in any period t.

as copies use the same link that has to be clicked.

[5]For convenience, we call the most prominent spot "the first spot" or "the top spot".

[6]The modelling does not explicitly include prices, but one may think that price is one of the elements that makes up the common component in consumers' utility function. Thus, we implicitly allow prices to differ across products. The reason we do not explicitly introduce prices is that we want to focus on what platforms may learn from consumer choices, while if we included price-setting behaviour by firms, the platform could potentially also learn about product characteristics from firms' price choices. This would complicate matters beyond what we address in this paper.

In this subsection, we will consider the following simple algorithm.

Definition 1. *The NE ranking algorithm puts the item the last consumer has bought prominently on the top spot and places all non-inspected products randomly on the following spots.*

We will show that for small enough values of δ the NE ranking algorithm is the "unique" non-experimentation algorithm. If a platform does not experiment, all items can be divided into three categories: (i) at most one item that has been inspected by previous consumers and never rejected, (ii) a set of items that have been inspected and at least once rejected, and (iii) a set of uninspected items. The critical issue turns out to be which item consumers choose to inspect if they have the choice between inspecting a random item and an item that previously has been rejected. We show below that if δ is small enough, they prefer to inspect a random item. Thus, the best possible ranking for consumer t using all information available to the platform in period t is to put on the top spot the item that consumer $t-1$ bought and clearly mark rejected items so that consumer t can skip them and inspect other items. As the platform has no information about non-inspected items, it has to place them randomly. Thus, the NE algorithm is unique up to permutations of the elements of different categories of items.

The NE algorithm will also be an important reference point for future analysis as (i) it provides a benchmark for how much more the platform can learn by experimenting, and (ii) under a strong form of competition, platforms will find it optimal to choose this algorithm (as we will show in Section 5).

Given the common component in the consumers' utility function, the NE algorithm contains relevant information to consumers and consumers find it optimal to start searching the product on the top spot if the platform chooses the algorithm. Define u_R^* as the reservation utility under random search, or the *ex ante* reservation utility, i.e.,

$$u_R^* = E[max\{(1-\delta)v + \delta m, u_R^*\}] - s.$$

If consumers decide to continue to search beyond the top spot and follow the "recommendation" of the NE algorithm by inspecting items in the order they are presented (and in Lemma 2 below we show that this is the case for small enough δ), then they effectively search among randomly ordered items. Thus, it is clear that any consumer facing the NE algorithm buys an item if the utility it generates is larger than or equal to u_R^*. For any $s > 0, u_R^* < \bar{u}$.

Using the definition of u_R^* we can define v_R^* by $v_R^* \equiv (u_R^* - \delta\underline{m})/(1-\delta)$ if $u_R^* < (1-\delta)\bar{v} + \delta\underline{m}$ and $v_R^* \equiv \bar{v}$ if $u_R^* \geq (1-\delta)\bar{v} + \delta\underline{m}$. For an arbitrary reservation utility $\hat{u}$ (if it exists) we can depict the situation as in Figure 1. The reservation utility is a downward sloping line in the (v, m) space. It crosses the horizontal axis at $\hat{v} = \frac{\hat{u}-\delta\underline{m}}{1-\delta}$ as in Figure 1

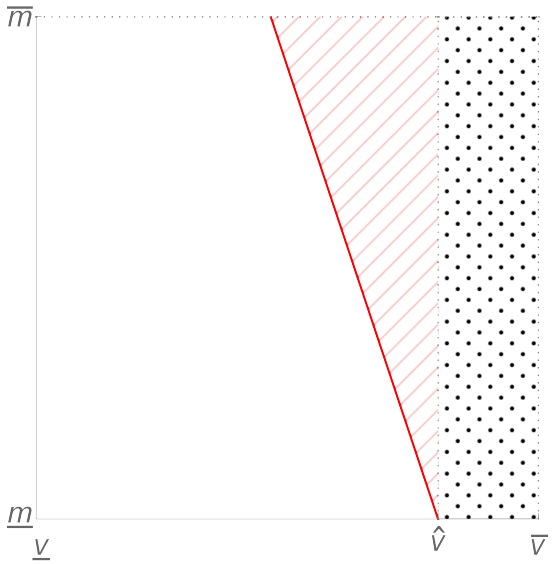

Figure 1: *Consumers accept items north-east from the thick (red) line: in the triangle (hatched) and rectangle (dotted);* $s = 0.2$, $\delta = 0.25$.

if $\hat{u} < (1-\delta)\bar{v} + \delta\underline{m}$. If $\hat{u}$ is a reservation utility and consumers draw a utility level above this line, they accept. The acceptance area consists of a rectangle where the utility level has a common component $v > \hat{v}$ (where all consumers immediately buy such products) and a triangle where the utility level has a common component $v < \hat{v}$ (where only those consumers immediately buy for whom the idiosyncratic component is large enough).

If $v_R^* < \bar{v}$, an implication of the NE algorithm is that the platform eventually puts always the same product on the top spot. The first consumer buys a product A with $u_A \geq u_R^*$ and this product will be recommended to the second consumer. The second (and all future consumers) will certainly buy this product after inspection if it has a common component $v_A \geq v_R^*$. If on the other hand, $v_A < v_R^*$, then eventually there will be a consumer that has a low idiosyncratic component and that consumer will buy a different product B. This product may also be eventually rejected if $v_B < v_R^*$, but if $v_R^* < \bar{v}$ at some point a product with a $v \geq v_R^*$ will be bought, all consumers thereafter will be recommended that product, and they buy it immediately after inspection. Thus, using the NE ranking algorithm, the platform will eventually learn that the recommended product has a $v \geq v_R^*$ if $v_R^* < \bar{v}$ and not more than that, where v_R^* can also be implicitly defined by

$$(1-\delta)v_R^* + \delta\underline{m} = \Pr(u \geq u_R^*)E(u|u \geq u_R^*) + \Pr(u < u_R^*)\left((1-\delta)v_R^* + \delta\underline{m}\right) - s, \qquad (1)$$

or

$$v_R^* = \frac{E(u|u \geq u_R^*) - \delta\underline{m}}{1-\delta} - \frac{s}{(1-\delta)\Pr(u \geq u_R^*)}.$$

Figure 2 illustrates the cutoffs v_R^* for different values of s in the case of the uniform

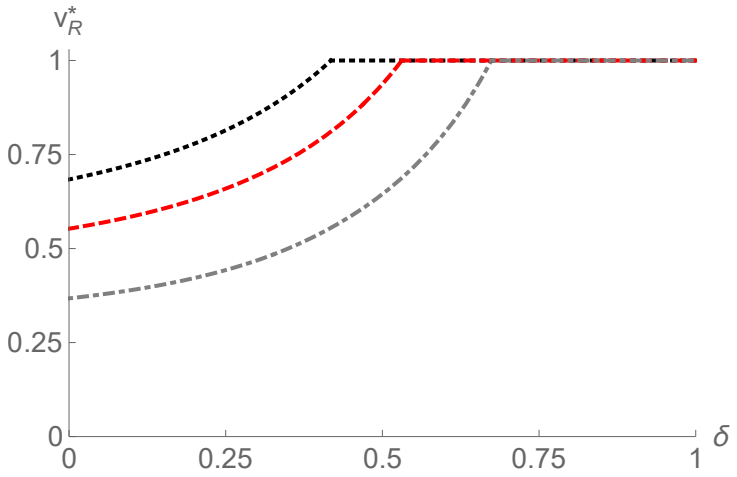

Figure 2: *The cutoff v_R^* as a function of δ for $s = 0.05$ (black dotted), $s = 0.1$ (red dashed), and $s = 0.2$ (grey dot-dashed).*

distribution. For all δ below some critical value $\hat{\delta}_R$, where $\hat{\delta}_R$ is the unique solution to (1) with $v_R^* = \bar{v}$, [7] v_R^* is smaller than the maximum common value of 1.

We will now argue that if δ is small enough a consumer never inspects a position k in the ranking if they know it contains with strictly positive probability a previously rejected item (and an uninspected item with the rest of the probability). Using Figure 1, this means that a consumer does not find it optimal to inspect items in the (red) hatched "triangle region" as these are items that have been *inspected but not bought* by some previous consumer and they would prefer to inspect a random, uninspected item first (and there always will be those items).

To gain intuition, consider a hypothetical situation where a consumer's reservation utility of picking a random item from a certain subset of products equals $\tilde{u} \geq u_R^*$, and define $\tilde{v}$ by $\tilde{v} = (\tilde{u} - \delta \underline{m}) / (1-\delta)$. Suppose also that the consumer knows that the common value of an item equals $\tilde{v}$ and the only uncertainty they face in inspecting this item is the value of the idiosyncratic component. A consumer will not want to inspect this item (with the intention of buying it) above inspecting a random item with reservation utility $\tilde{u}$ if $(1-\delta)\tilde{v} + \delta E(m) - s \leq \tilde{u}$, which is equivalent to $\delta \left(E(m) - \underline{m}\right) \leq s$, or defining

$$\hat{\delta} := \frac{s}{E(m) - \underline{m}},$$

to $\delta \leq \hat{\delta}$.

A specific case in point arises when considering $\tilde{u} = u_R^*$. In that case the consumer has always the option of inspecting a random product giving a reservation utility of u_R^* and if $\delta \leq \hat{\delta}$ they prefer that to inspecting and buying a product with a common component smaller than or equal to v_R^*, i.e., a product in the "triangle region" that under

[7] Note that, as u_R^* is uniquely defined, $\hat{\delta}_R = (\bar{v} - u_R^*)/(\bar{v} - \underline{m})$ is also uniquely defined if $\bar{v} \neq \underline{m}$.

the algorithm considered above is eventually rejected by a consumer. We will say that the platform's ranking weakly improves over time if the platform's knowledge embedded in the ranking for consumer t is such that (if they had a choice) consumers weakly prefer to search in period t to searching in any period before t. Using this notion, we can state the following Lemma.

Lemma 2. *If $\delta < \hat{\delta}$ and the platform's ranking (weakly) improves over time, then consumers prefer to inspect random products to products that have been rejected by previous consumers.*

A first implication of Lemma 2 is that a non-experimenting platform always ranks on the top spot the only item that has been bought but not rejected in the past. A second implication of Lemma 2 is that if $\delta < \hat{\delta}$ a non-experimenting platform will never be able to induce a consumer to inspect an item that with positive probability is rejected by a previous consumer. As there are infinitely many products, a consumer can always inspect a random product on which the platform does not have information and would prefer to do so because of the upward potential such a product has: in contrast to a rejected item, a random item could also have a value in the dotted rectangle area in Figure 1, whereas if the value turns out to be such that $u < u_R^*$ the consumer would want to continue to search anyway. The upper bound $\hat{\delta}$ is such that the expected value of inspecting a rejected item cannot be larger than u_R^*. Thus, we have the following:

Corollary 3. *If $\delta < \hat{\delta}$ a non-experimenting platform will choose the NE algorithm and cannot learn more than that $lim_{t\to\infty}\ v_1(t) \geq v_R^*$.*

By Lemma 2, we know that if $\delta < \hat{\delta}$ consumers prefer inspecting random items ahead of previously rejected items. As a result, the platform serves consumer t's interests best at any t by using the NE algorithm. We argued above that, if the platform uses this algorithm, it finds an item with $v \geq v_R^*$ and does not learn more, and that the algorithm is unique up to inconsequential permutations of items.

If $\delta > \hat{\delta}$, it is much more difficult to characterize the algorithm that a non-experimenting platform will use. The reason is not only that consumers may prefer to inspect previously rejected items ahead of random items (depending on how many times they have been rejected), but also that the platform does not observe the consumers' utilty levels. The latter becomes especially cumbersome if consumers' optimal search strategy is not characterized by a reservation value as in that case an item may be rejected by a consumer even though items with lower utility levels would have been accepted; see the next section for an example. In that case it is difficult to characterize for any period t the platform's posterior value of an item and therefore the best possible ranking for consumer t that uses all the information available to the platform in period t.

3 Results under Monopoly: $\delta = 0$

To show how experimentation may help the platform to learn more and to demonstrate the complexities that may arise due to consumer learning, this section first focusses on the simplest case where $\delta = 0$ (which implies $u = v$).

Consumer learning may potentially complicate the analysis as optimal search strategies may not be characterized by a reservation value as the following example illustrates. The key to the example is the following. Very early consumers must have a reservation value strategy as they search among random items if they do not accept the first item. However, if later consumers do not know whether a platform is randomizing its ranking or not, observing a v that is smaller than the reservation value of a previous consumer makes it more likely that the platform is experimenting, while observing a v that is above the previous reservation makes it more likely that the platform is *not* experimenting. As the continuation value of search is not independent of the current observation, a reservation value strategy may fail to exist.

Example (No reservation value strategy)**.** *Consider the following algorithm: (i) for consumers $t = 1, 2, 3, 4$ rank all items randomly; (ii) for consumer $t = 5$ (6) rank the items consumers $t = 1, 2$ ($t = 3, 4$) bought in a random order on the two top spots, followed by a random ranking of all other items, and (iii) for consumer $t = 7$ with probability ρ rank items randomly, while with probability $1 - \rho$ rank the items consumers $t = 5, 6$ bought in a random order on the two top spots, followed by a random ranking of all other items.*

It is clear that consumers $t = 1, 2, 3, 4$ use the cutoff v_R^* and that consumers $t = 5, 6$ also use a cutoff strategy for their first and second search, but with $\hat{v} > v_R^*$ for their first search (and v_R^* for their second), where

$$\hat{v} = \frac{[1 - F(\hat{v})]\, E(v|v \geq \hat{v})}{1 - F(v_R^*)} + \frac{[F(\hat{v}) - F(v_R^*)]\, \hat{v}}{1 - F(v_R^*)} - s, \tag{2}$$

and the RHS expresses the continuation value of search. We will argue that one can choose ρ such that consumer 7 stops searching if they first inspect an item with $v = \hat{v} - \varepsilon$ (i.e., just below the first cutoff of consumers 5 and 6) and that they continue searching if they inspect an item with $v = \hat{v} + \varepsilon$.

The main step in the argument is to realize that for all $\rho \in (0, 1)$ the continuation value of search after observing $v = \hat{v} + \varepsilon$ is strictly larger than the continuation value of search after observing $v = \hat{v} - \varepsilon$.[8] *The reason is that, conditional on the platform not randomizing its ranking, the only possibility for consumer 7 to observe a $v < \hat{v}$ on their first search is that at least one of consumers 5 and 6 was presented with two items in the top positions that have $v \in [v_R^*, \hat{v})$. On the other hand, and again conditional on the platform not randomizing its ranking, consumer 7 can observe a $v > \hat{v}$ on their first search*

[8] The formal proof of these statements is in the Appendix.

either because one of consumers 5 and 6 encountered this item on their first search and immediately stopped searching, or that they first inspected another item with $v \in [v_R^*, \hat{v})$. *Thus, after observing* $v = \hat{v} + \varepsilon$ *consumer 7 believes it is more likely the platform is not randomizing its ranking than after observing* $v = \hat{v} - \varepsilon$ *and becomes more optimistic to observe a high* v *on their next search. Together with the facts that (i) if* ρ *is close to 0 consumer 7 will always want to continue searching independent of observing* $v = \hat{v} - \varepsilon$ *or* $v = \hat{v} + \varepsilon$, *while (ii) if* ρ *is close to 1 consumer 7 will always stop searching after observing* $v = \hat{v} - \varepsilon$ *and* $v = \hat{v} + \varepsilon$, *this implies that the optimal strategy of consumer 7 is not a reservation value strategy.*

To get around these intricacies related to consumer learning, define for any $s > 0$ a value v^* such that

$$v^*(0) = E(v|v \geq v^*(0)) - s.$$

In words, for $\delta = 0$ the value $v^*(0)$ is such that the consumer is indifferent between (i) accepting an item with $v^*(0)$ and (ii) continuing to search and paying the search cost, knowing that the next item is at least as good as $v^*(0)$. Note the important difference with the definition of v_R^* when setting $\delta = 0$ and realizing that then $u = v$ in (1). In (1), the next-searched item is randomly selected, so there is a positive probability that the next-searched item has a low v; here, the consumer expects that the next-searched item has a $v \geq v^*(0)$. The next proposition (i) states that the platform can never learn more than $v^*(0)$, and (ii) delivers an algorithm that in the long-run ensures that the platform finds at least one item with $v \geq v^*(0)$.

Proposition 4. *If* $\delta = 0$, *there exists an algorithm that achieves that* $lim_{t\to\infty} v_1(t) \geq v^*(0)$ *and this algorithm maximizes the platform's objective function.*

That the platform cannot learn more than that the v_1 in the top spot is larger than or equal to $v^*(0)$ follows from the fact that all consumers $t = 1, 2, ...$ will always accept and buy any $v \geq v^*(0)$. That there exists an algorithm that achieves this amount of learning is proved by constructing an algorithm that simplifies consumer learning. The algorithm implies that the consumers' optimal strategies use reservation values and that these reservation values (weakly) increase most of the time. In particular, in the proof we propose an algorithm that divides consumers in cohorts of size $K > 2$.[9] The first cohort of consumers will get a ranking where all (infinitely many) products are ranked randomly. The platform collects all the items that have been bought and the second cohort gets to see a ranking where these K products are placed on the top spots (and each item gets the top spot once). The platform treats all subsequent cohorts in a similar way as long as at least two different items are never rejected. If less than two never-rejected items remain, then the platform restarts the process. The platform indicates when the process

[9]The proof of Proposition 4 is a special case of the proof of Proposition 5 and therefore omitted.

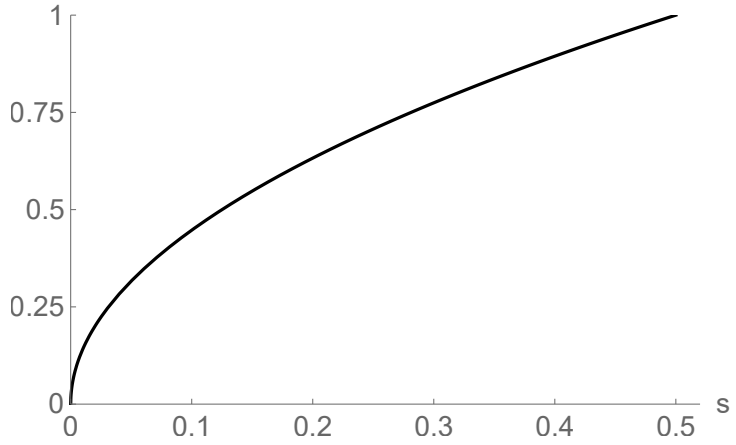

Figure 3: *The value to consumers of having (relative to not having) a platform as a function of* s*;* $\delta = 0$*.*

was restarted last by communicating the cohort j when the process was last restarted: the never-rejected items in a consumer's ranking are followed by j positions filled by copies of a randomly drawn uninspected item (that are recognized by consumers and not inspected).[10]

The proof shows that until the process is restarted consumers in subsequent cohorts become more picky over time and, because $1-G(v^*(0)) > 0$ for any $s > 0$, eventually two items with a $v \geq v^*(0)$ are found. Interestingly, the platform is not able to learn that only one item has a $v \geq v^*(0)$: the platform can only learn so much if consumers (rationally) expect that if they continue to search another item, it also has a value $v \geq v^*(0)$.

To see how much consumers benefit in the search market from a platform that aims to learn as much as possible, we compare the value $v^*(0)$ with the standard consumer search problem in the absence of a platform where the products to be inspected next are randomly drawn. In the latter case the threshold value v_R^* that is acceptable for consumers is given by $s = \int_{v_R^*}^{\bar{v}} (1-G(v))dv$. For the uniform distribution, $v_R^* = 1-\sqrt{2s}$ and $v^*(0) = 1-2s$. Figure 3 depicts $\frac{(1-2s)-(1-\sqrt{2s})}{1-\sqrt{2s}}$, which is the difference in expected utility due to the platform's ranking algorithm relative to the expected utility the consumer gets without the platform. From the Figure it is easy to see that the larger the search cost, the larger the benefit to the consumer, but that even at relatively low inspection costs of say 0.05 consumers benefit around 31.6%.

4 General Monopoly Results: $\delta > 0$

For $\delta > 0$ we have to redefine and replace $v^*(0)$ by v^* by acknowledging that there is a difference between u and v, while keeping the idea that when continuing to search, the

[10] If K is large and $v^*(0)$ not too close to $\bar{v}$, then the number of times that the process is restarted is likely to be zero or a very small number.

consumer pays the search cost and expects that the next item is at least as good as v^*. The minimum possible utility level a consumer gets by accepting a product with a common value of v^* is given by $(1-\delta)v^* + \delta\underline{m}$. If a consumer drawing this utility level expects that on their next search they observe a product that has a common component that is at least as large as v^*, the consumer is indifferent between buying now and continuing to search if

$$(1-\delta)v^* + \delta\underline{m} = (1-\delta)E(v|v \geq v^*) + \delta Em - s. \tag{3}$$

This is the generalized definition of v^* we will use in this section. If $\delta > 0$ there are two cases to be considered, however: (i) if $\delta < \hat{\delta}$, then there exists an interior $v^* < \bar{v}$ that solves (3), while (ii) if $\delta \geq \hat{\delta}$, then we define $v^* = \bar{v}$. As the platform faces very different issues in these two cases, we treat them separately in the next two subsections.

4.1 An Interior Solution: $0 < \delta < \hat{\delta}$

When δ is small enough, the results of the previous section essentially generalize, with v^* replacing $v^*(0)$. The main difference with the previous section is that now different consumers would make different choices whether to buy or continue to search in the same environment as the idiosyncratic component of their value for a product may differ. Nevertheless, the same algorithm will bring the platform maximal learning:

Proposition 5. *If $\delta < \hat{\delta}$, there exists an algorithm that achieves that $lim_{t\to\infty} v_1(t) \geq v^*$ and this algorithm maximizes the platform's objective function, where the unique cutoff v^* is implicitly defined by*

$$\int_{v^*}^{\bar{v}} \frac{1-G(v)}{1-G(v^*)} dv = \frac{s - \delta(Em - \underline{m})}{1-\delta}. \tag{4}$$

The cutoff v^* strictly increases in δ, decreases in s, and exceeds v_R^*.

The comparative statics with respect to s and δ are interesting, but intuitive. When s is larger, consumers are less inclined to inspect products beyond the top spot. This limits the possibility for the platform to learn from consumer choices and, as in the previous section, v^* decreases in s. On the other hand, when δ is larger, consumers tend to have more dissimilar tastes and therefore the probability that consumers are satisfied with the product that is placed on the top spot is smaller. This implies that consumers are more inclined to search, making it easier for the platform to discern products with really high values of v. If the weight on the idiosyncratic component of the utility function δ increases, or s decreases, then δ becomes equal to $\hat{\delta}$ and the platform's incentive to experiment vanishes, as we will see in the next subsection.

Note that by finding products with a common value that exceeds v^* and placing these prominently in the ranking, the platform not only maximizes the amount it learns, but

implicitly also the long-run consumer utility of using the platform. Consumers not only get the highest common value they can ever expect to see on the platform, they also economize on search cost. If consumer utility is positively correlated with the platform's other revenue streams, this re-interpretation of the platform's objective provides another rationale for using it.

It is also interesting to observe that the platform would actually not be able to learn more than what is characterized in the Proposition even if its communication to consumers is not restricted to what can be communicated though an algorithm. In particular, the platform knows which items have been searched by which consumers and in which sequence, but this information cannot be fully communicated by a ranking. However, as items with a $v \geq v^*$ will always be accepted by any consumer, the platform will not be able to learn more even if it could communicate more.

To illustrate how much more the platform can learn by experimenting, it is useful to consider that v and m are uniformly distributed on $[0, 1]$. In this case (3) can be explicitly solved to give

$$v^* = 1 + \frac{\delta - 2s}{1 - \delta}.$$

The solution is smaller than 1 for $\frac{\delta}{2} < s < \frac{1}{2}$, while for $s < \frac{\delta}{2}$ there is no interior solution.

Solving (1) for uniform v and m gives

$$v_R^* = \frac{(2 - \delta) - \sqrt{8(1 - \delta)s - \frac{1}{3}\delta^2}}{2(1 - \delta)}.$$

For example, for $s = 0.05$ and $\delta = \hat{\delta} = 2s$, the experimenting platform learns 38.1% more than a nonexperimenting platform. We later illustrate these cutoffs in Figure 4.

Consumer heterogeneity in s and δ. It is easy to see how our analysis generalizes to consumers having different search costs and weights that they attach to the idiosyncratic component. In particular, suppose that the search costs are distributed over an interval $[s_L, s_H]$, with $s_L > 0$, the relative weight δ is distributed over an interval $[\delta_L, \delta_H]$ with $\delta_L > 0$, s and δ are independently distributed from each other, and the platform does not know a consumer's values of these parameters.

If consumers have different search costs and utility weights, they also have different cutoff values v_t^*. For a consumer t with $(s, \delta) = (s_t, \delta_t)$, equation (3) now becomes

$$(1 - \delta_t)v_t^* + \delta_t \underline{m} = (1 - \delta_t)E(v|v \geq v_t^*) + \delta_t Em - s_t, \tag{5}$$

and we write $v_t^*(s_t, \delta_t)$ for the common value that makes a consumer (s_t, δ_t) who observes the lowest value of the idiosyncratic component indifferent between buying and continuing to search if they expect to see an item with a higher common value on their next search.

We have shown above that if $\delta < \hat{\delta}$ the cutoff value v^* is interior, decreases in s and increases in δ. Translating that result to the context of heterogeneous s and δ, it is clear that consumers with lower s and higher δ have larger values of v_t^*, with the largest $v_t^*(s_t, \delta_t)$ being $v_t^*(s_L, \delta_H)$. As the platform mainly learns when items are rejected, and items with a $v < v_t^*(s_L, \delta_H)$ eventually will be rejected, it follows that the platform's maximization problem has an interior solution if $\delta_H \leq \hat{\delta}_L := \frac{s_L}{E(m)-\underline{m}}$:

Corollary 6. *If consumers differ in their search cost s and their relative weight on the idiosyncratic component δ, then if $\delta_H \leq \hat{\delta}_L$, the platform can learn that $lim_{t\to\infty} v_1(t) \geq v_H^*$, where $v_H^* < \bar{v}$ solves*

$$\int_{v_H^*}^{\bar{v}} \frac{1-G(v)}{1-G(v^*)} dv = \frac{s_L - \delta_H(Em - \underline{m})}{1-\delta_H}.$$

4.2 The Boundary Solution: $\delta > \hat{\delta}$

When $\delta \geq \hat{\delta}$, there is no interior solution to (3). We analyze this situation in two steps. First, we consider that (1) also does not have an interior solution for v_R^*, which is the case for $\delta > \hat{\delta}_R$.[11] In this case, there is no item such that all consumers accept it given that they expect to find a random item on their next search. Second, we consider the intermediate case where $\hat{\delta}_R > \delta > \hat{\delta}$.

If $\delta > \hat{\delta}_R$, then the platform does not need to induce consumers to search more by presenting them with good items on their next search. This is because all items are such that some consumers will not buy them as they prefer to continue searching even if they expect to find a random item on their next search. Thus, the platform does not need to experiment to make consumers reject items. Instead, the platform can in principle observe the fraction of consumers that reject an item to make an inference about the common value of the item as, with a random continuation value of search, an item with a common value v will be rejected with a probability $F\left(\frac{u_R^*-(1-\delta)v}{\delta}\right)$, which as $u_R^* > (1-\delta)\bar{v} + \delta\underline{m}$ is positive for all v and decreasing in v.

If $\delta > \hat{\delta}_R$, the platform's problem has features of a standard statistical inference problem, like tossing a coin to see whether it is fair. There are, however, two fundamental differences. First, the platform wants to select an alternative with a v as close as possible to $\bar{v}$ and therefore has not only to decide whether a particular item is likely to be good enough, but also *when to switch* and try out other items if items tried so far are likely not close to $\bar{v}$. Second, it is not optimal for the platform to decide once and for all based on a finite sample of T consumers. In a standard statistical inference problem there is a given size of the sample selected, but here the platform will always learn more as more consumers use it. Thus, what is likely a good item based on the choices made by the first

[11] Note that $\hat{\delta}_R > \hat{\delta}$ follows from $u_R^* < u^*$.

T consumers may not be a good item based on the choices of the first $2T$ consumers. The proof of Proposition 7 below shows that in the long run the platform is able to find an item with a common value v close to $\bar{v}$ with a probability close to 1.

If $\hat{\delta}_R > \delta > \hat{\delta}$, then (1) has an interior solution for v_R^*, but (3) does not have an interior solution for v^*. Given that for $\delta > \hat{\delta}_R$ the platform does not learn $\bar{v}$ exactly, the platform's problem in case $\hat{\delta}_R > \delta > \hat{\delta}$ is actually closer in spirit to the one analyzed in the previous subsection (i.e., when there is an interior solution for v^*). In particular, the platform can use the same algorithm as in Proposition 5 for the first cohorts of consumers to make consumers more picky and not just accept to buy items with $v \geq v_R^*$. When a cohort is reached whose critical cutoff value $\hat{u}$ would be larger than $(1-\delta)\bar{v} + \delta\underline{m}$, the platform may dilute the continuation value of search by invoking enough randomly selected, so far uninspected, items, to make consumers in this cohort willing to accept all items with a common value close to $\bar{v}$.

We summarize the above discussion as follows.[12]

Proposition 7. *If $\delta > \hat{\delta}$, there exists an algorithm such that for every $\eta, \tau > 0$ the probability that $lim_{t\to\infty} v_1(t) \geq \bar{v} - \eta$ is at least $1 - \tau$.*

Note that unlike when there is an interior solution (and $\delta < \hat{\delta}$), the platform does not maximize long-run consumer utility of using the platform when $\delta > \hat{\delta}$ by learning that the common value of a product is close to $\bar{v}$ and placing this item on the top spot. The reason is that, as consumers reject all items with positive probability, consumers would benefit if an infinite number of items with a common value of $\bar{v}$ are placed in top positions in case they draw a low value of the idiosyncratic component for many top-ranked items. As explained above, however, the platform will be able to find a few items with a large common value, but even in the long run cannot find an infinite number of such items.

Figure 4 summarizes the discussion in this section so far by illustrating (for v and m being uniformly distributed on $[0, 1]$) that by experimenting with rankings the monopolist can learn considerably more than a non-experimenting monopolist.

4.3 Blind Buying

We now consider how our results are affected if consumers can buy a product without inspecting it (as they may trust the platform's "recommendation") and the platform cannot observe whether or not consumers inspected the product, i.e., the platform can only observe whether consumers buy a product and make inferences about the consumers' optimal inspection behaviour. We refer to buying without inspecting as "blind buying".[13]

[12] If consumers differ in their values for s and δ, and for at least some consumers (5) does not have an interior solution, i.e., in the notation introduced in the previous subsection we have $\delta_H > \hat{\delta}_L$, then we can prove the same proposition using the algorithm outlined in the proof.

[13] Doval (2018) studies blind buying in the environment of Weitzman (1979).

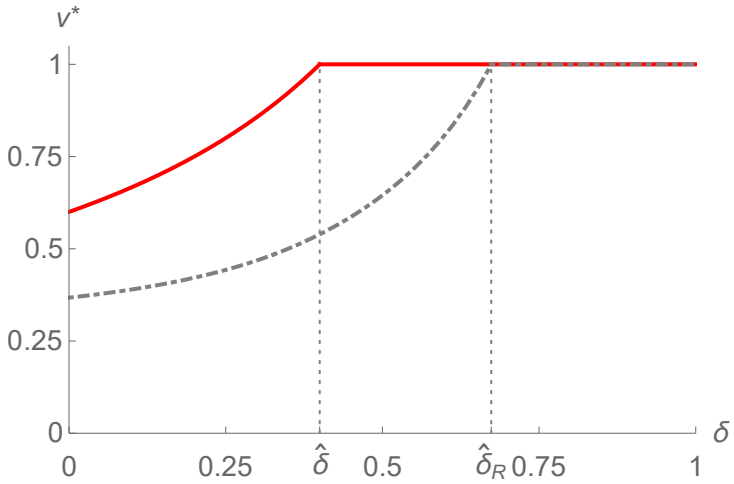

Figure 4: *The cutoffs v^* (red solid) and v_R^* (grey dot-dashed) as a function of δ for $s = 0.2$.*

If the platform only observes that consumers bought a product, but does not know whether they inspected the product (or believes that they have not), then it does not learn from the consumers' choice. Clearly, if consumers have the option of blind buying the platform cannot learn more than what we have considered so far. However, if only a fraction of consumers consider blind buying, while another fraction always inspects before buying (as we previously assumed all consumers would do), then the previous analysis remains valid. What matters is that eventually some consumers inspect products and they continue to search if products yield a utility level that is too low.

In the remainder of this subsection we suppose that all consumers consider buying blindly, but do so only when it is rational to do so. To this end, for every $\delta > 0$ define by $\hat{s}_R(\delta)$ the largest search cost such that $u_R^* = (1-\delta)\bar{v} + \delta\underline{m}$, i.e., this is the largest search cost such that $v_R^* = \bar{v}$. For $s < \hat{s}_R(\delta)$, in Figure 1 only the "triangle" region exists and the expected utility of searching over randomly ordered items is $u_R^* > (1-\delta)\bar{v} + \delta\underline{m}$. For all $\delta > 0$, let $m_R^*(s)$ satisfy $(1-\delta)\bar{v} + \delta m_R^*(s) = u_R^*$. Note that $m_R^*(s=0) = \bar{m}$, $m_R^*(s = \hat{s}_R(\delta)) = \underline{m}$, and $m_R^*(s)$ strictly decreases in s.

The highest expected utility that consumers can potentially get by buying blindly is $(1-\delta)\bar{v} + \delta E(m)$. Conversely, by inspecting consumers can always ensure a utility of u_R^* because they can always ignore the platform's ranking and sample items in a random order. Thus, inspecting is certainly better than buying blindly for all consumers $t \geq 1$ if

$$(1-\delta)\bar{v} + \delta E(m) < (1-\delta)\bar{v} + \delta m_R^*(s) - s,$$

or

$$\delta E(m) + s < \delta m_R^*(s).$$

At $s = 0$, the inequality holds. At $s = \hat{s}_R(\delta)$, the inequality becomes $\delta E(m) + \hat{s}_R(\delta) < \delta\underline{m}$

which does not hold. Since m_R^* decreases in s, a unique $\bar{s}(\delta) < \hat{s}_R(\delta)$ exists such that $\delta E(m) + \bar{s}(\delta) = \delta m_R^*(\bar{s}(\delta))$ and the inequality holds for all $s < \bar{s}(\delta)$. Thus, for any $s < \bar{s}(\delta)$, all consumers inspect and we can apply Proposition 7 to show that the platform learns a product has a v close to $\bar{v}$.

Proposition 8 states this result more formally and also that for intermediate values of s the platform may learn strictly more than the cutoff v_R^*, but never as much as v^*.

Proposition 8. *(i) For any $\delta > 0$ if $s < \bar{s}(\delta)$, then an algorithm exists that for every $\eta, \tau > 0$ achieves that the probability that $lim_{t\to\infty} v_1(t) \geq \bar{v} - \eta$ is at least $1 - \tau$. (ii) If δ is small enough, $\tilde{s}(\delta) > \hat{s}_R(\delta)$ exists such that for all $s \in (\hat{s}_R(\delta), \tilde{s}(\delta))$ an algorithm exists that achieves that $lim_{t\to\infty} v_1(t) \geq \hat{v}$ with $v_R^* < \hat{v} \leq v^*$. If $v^* < \bar{v}$, the platform is never able to learn $lim_{t\to\infty} v_1(t) \geq v^*$.*

The second part of Proposition 8 states that if, as in subsection 4.1, the maximal the platform can learn is $v^* < \bar{v}$ it may still be able to learn more than v_R^* if δ is small and s is not too large. This is best explained by considering that $\delta = 0$ and $s > 0 = \hat{s}_R(0)$, but small, so that $v_R^* < \bar{v}$. As by blind buying the first consumer gets a pay-off of Ev, while by inspecting they are able to get v_R^* (with v_R^* being close to $\bar{v}$ for $s \to 0$) it is clear that for small enough search costs they will choose to inspect. Later consumers would prefer to buy blindly *if they expected the platform to prominently rank the product that the first consumer bought*. But the platform can learn more in the long run by ranking products randomly for a first cohort of consumers and then ranking the products bought by them highly for a next cohort of consumers, etc (as in the algorithm used in Proposition 5). This is easy to see for the uniform distribution. In that case, by inspecting items consumers in the second cohort can use the cutoff $\hat{v}_2$ defined by

$$\hat{v}_2 = \frac{1 + \hat{v}_2}{2} \frac{1 - \hat{v}_2}{1 - v_R^*} + \hat{v}_2 \frac{\hat{v}_2 - v_R^*}{1 - v_R^*} - s.$$

The RHS of this expression is the expected value $\frac{1+\hat{v}_2}{2}$ of the next item searched if it has a value larger than the cutoff $\hat{v}_2$ times the conditional probability the item has a higher value than $\hat{v}_2$ plus the conditional probability the item has a smaller value than $\hat{v}_2$ times the pay-off of continuing to search, which is $\hat{v}_2$ minus the search cost. This can be explicitly solved as $\hat{v}_2 = 1 - \sqrt{2(1 - v_R^*)s}$. If, on the other hand, these consumers buy blindly they settle for $E(u|u \geq u_R^*) = \frac{1+v_R^*}{2}$. As in case of the uniform distribution $v_R^* = 1 - \sqrt{2s}$, inspecting yields a higher pay-off if $1 - \sqrt{2s\sqrt{2s}} > 1 - \frac{1}{2}\sqrt{2s}$, or $s < \left(\frac{1}{2}\right)^8$. Thus, even if they can buy blindly these consumers strictly prefer not to do so for sufficiently small search costs and the platform is able to learn more than that an item has a $v \geq v_R^*$ if $v_R^* < \bar{v}$. By continuity a similar argument continues to hold if δ is a small positive number. The proof shows that this continues to be true for other distribution functions.

Finally, a simple argument shows that $\hat{v} < v^*$ if $v^* < \bar{v}$. Recall that v^* is such that

$u^* = (1-\delta)v^* + \delta\underline{m}$ and all items are accepted in the long run. The cutoff v^* is the highest possible cutoff that can be reached when blind buying is allowed because the platform can learn weakly less. Suppose then that the cutoff v^* is supposed to be reached and consider a T that is so large that the platform knows that certain items have a common value above a number that is slightly smaller than v^*. Then consumers $t > T$ expect to get a utility (almost) equal to $E(u|u > u^*) - s$ when inspecting, whereas they expect to get (almost) $E(u|u > u^*)$ by buying blindly. As a result, all consumers $t > T$, i.e., sufficiently far in the queue, prefer to buy blindly. Thus, the platform learns strictly less than when consumers cannot buy blindly.

5 Competition

Finally, we consider markets where platforms are competing with each other for consumers and ask to what extent the results under monopoly continue to hold. In particular, we ask whether platforms are able to learn to the same extent as under monopoly.

An important part of the analysis under monopoly is that the platform can experiment with different rankings to learn consumer choices for a variety of rankings. Experimentation implies that the utility of early consumers is sacrificed to the benefit of later consumers. This is easiest seen by considering the second consumer using the platform. If the first consumer bought a particular product (whether or not it is the first-inspected product), then the platform gets positive information about the common value of that product and *if* it wants to serve the second consumer best, it will place that product on the top spot. However, doing so limits the scope of learning from the choices that the second consumer makes and a monopoly interested in what it can learn in the long run wants to hide the product bought by the first consumer from the second consumer.

Competition limits, however, the extent to which platforms can experiment with rankings as consumers only care about their own utility and not about the utility of future consumers. Accordingly, early consumers will avoid using a platform that they expect experiments most as they know this is not to their own advantage. Obviously, without further regulation, if consumers are not using the platform, it cannot learn from them either. Note, however, that EU regulation in the DMA attempts to enforce exactly that other platforms are able to use consumer search data from other (dominant) platforms.[14] Below we will discuss the importance of this regulation in more detail.

To study the implications of competition in the absence of regulation, we extend the monopoly model in the following way. First, at stage 0 platforms $i, i = 1, 2$, simultane-

[14]It states "[g]atekeepers should therefore be required to provide access, on fair, reasonable and non-discriminatory terms, to those ranking, query, click and view data in relation to free and paid search generated by consumers on online search engines to other undertakings providing such services, so that those third-party undertakings can optimise their services and contest the relevant core platform services."(European Parliament, 2022a, DMA, paragraph (61)).

ously choose how to rank products in every period, i.e., they choose $r_i = \{r_t^i\}_{t=1}^{\infty}$. Note that as a platform's ranking depends on consumers' past choices and the platform also observes whether a consumer has used its platform or not, r_t^i may depend on which of the consumers $1, ..., t-1$ have used the platform. Second, consumers enter sequentially and choose which platform to use for their search, which products to inspect, and which of the inspected products to buy. Consumers know the platforms' algorithms and single home (potentially because of switching costs).[15] Thus, an individual consumer's full search process takes place on the platform of their initial choice. As under monopoly, consumers do not observe previous consumers' behaviour on any of the platforms. Third, each platform aims to offer consumers that buy on it a first-ranked item with as high a common component as possible in the long run. More formally, platform i's objective is to maximise

$$lim_{t\to\infty}\pi_t^i v_1^i(t),$$

where π_t^i denotes the probability consumer t buys on platform i and $v_1^i(t)$ is the common component of the first-ranked item in platform i's ranking for the tth consumer.

The competition result below depends on whether consumers know which platform previous consumers have used (or some usage statistic, such as market shares). To ease discussion we distinguish between consumers not knowing and knowing the market shares.

If *consumers do not know market shares*, then competition severely limits the extent to which platforms can learn about the common value of items. In particular, platforms cannot do better than under the *no experimentation algorithm* introduced in subsection 2.1. The reason is that the only information consumers have before deciding which platform to use is their knowledge of the algorithms. They will rationally decide to use the platform that is experimenting less. Thus, competing platforms choose the *NE algorithm.* Interestingly, this is true even though the platforms know market shares (or previous consumers' choices): even though a platform's algorithm may, in principle, experiment more when the platform has a larger market share than its competitor, the platform will be "undercut" by the competitor if it actually does so. We prove this result for v^* being interior. For larger values of δ, the analysis is more complicated as it is unclear whether competition allows platforms to keep items long enough on the top spot to learn that their v is close to $\overline{v}$.

The situation is different if *consumers know market shares.* Now consumers have access to two sources of information when deciding which platform to use: the platforms' algorithms *and* market shares. If platforms use the same algorithms, but one has a larger

[15]This captures in the most simple way one of the implications of the EU's DSA (cf., footnote 4). In the unlikely event that none of the consumers knows anything about the platforms' choice of r, there is no competition between platforms and the results of Section 3 apply. Thus, to the extent that the DSA informs consumers of platforms' ranking algorithms, this section may be viewed as an evaluation of the possible effects of the DSA in markets with competing platforms.

market share, then consumers are better off using the platform with the largest market share. For example, if both platforms use the *NE algorithm* and one platform has had T users and the other $T' < T$ users in the past, then it is more likely that the platform with T users "recommends" a product with $v \geq v_R^*$ on the top spot (simply because it has had more consumer inspecting products). The platform with a higher market share has a competitive advantage and consumers will continue choosing it even if it experiments more than its competitor. Once a platform has more users than the other, it can build up its advantage, and eventually become a data-driven monopolist, learning the same amount as the monopoly platform.

Thus, we have the following proposition.[16]

Proposition 9. *(i) If $\delta < \widehat{\delta}$ and consumers do not know market shares, then under competition $lim_{t\to\infty} v_1^i(t) \geq v_R^*$, which is strictly smaller than under monopoly. (ii) If, on the other hand, consumers know market shares, then $lim_{t\to\infty} v_1^i(t) \geq v^*$ and a platform learns as much under competition as under monopoly.*

The EU's DMA intends to prevent the second part of the proposition from happening by allowing firms with smaller market shares to learn from the data that consumers generate on a dominant competing platform. Our analysis suggests, however, that the devil is in the details. The consumer search data is only useful if the ranking of items is known for each particular search query. Aggregate data is much less useful for other platforms to learn to the same extent. Moreover, if consumers believe that the larger platform is still at least marginally better, they will continue using that platform.

5.1 Asymmetric Competition: Incumbent and Entrant

We finally explore a different form of competition, namely, between an incumbent and a more efficient entrant and ask how much more efficient than the incumbent should the entrant be to attract consumers to its platform. We measure efficiency by how easily consumers can navigate a platform's website and find out information about products. Thus, the incumbent's efficiency is measured by the consumers' search cost on its platform, s_I, and the entrant's efficiency by s_E. As the incumbent has already learned about consumer preferences, while the entrant has not, the only interesting case is $s_E < s_I$.

Consider that the incumbent has been in the market long enough so that it ranks items with $v \geq v^*(s_I)$ first. Then a consumer's expected utility of searching on the incumbent platform is $u^*(s_I) = (1-\delta)E(v|v \geq v^*(s_I)) + \delta Em - s_I = (1-\delta)v^*(s_I) + \delta\underline{m}$. The first consumer on the entrant's platform has an expected utility of $u_R^*(s_E) = (1-\delta)v_R^*(s_E) + \delta\underline{m}$ because the entrant has not observed any choices to learn from. Thus, the consumer

[16] Results showing that our competition result is robust to similar extenions (on hetereogeneous consumers and blind buying) as we considered in the previous section are available on request.

prefers to use the entrant's platform iff $u_R^*(s_E) > u^*(s_I)$. The formal result below then follows directly from Proposition 5 and the facts that (i) for the same s an interior monopoly cutoff v^* satisfies $v^*(s) > v_R^*(s)$ and (ii) both cutoffs decrease in s.

Corollary 10. *For any $s_I > s_E > 0$ there exists $\Delta(s_I, s_E) > 0$ such that the entrant can enter the market only if $\frac{s_E}{s_I} < \Delta(s_I, s_E)$.*

For general distribution functions F and G, it is difficult to determine exactly how large the efficiency gains of the entrant should be for it to be able to enter the market. However, for specific distributions, it is easy to see that the efficiency gains may need to be very large for entry to take place. To see this, define $\frac{s_E}{s_I}$ as the measure of inefficiency of the entrant relative to the incumbent: if the entrant is maximally efficient (i.e., $s_E = 0$), the measure is zero and if the entrant has no efficiency advantage, the measure is one. Then, if v and m are uniformly distributed on $[0, 1]$, and considering $v_R^*(s_E)$ and $v^*(s_I)$ to be interior, the expected utility is higher on the entrant platform if

$$v_R^*(s_E) = \frac{(2-\delta) - \sqrt{8(1-\delta)s_E - \frac{1}{3}\delta^2}}{2(1-\delta)} > 1 + \frac{\delta - 2s_I}{1-\delta} = v^*(s_I),$$

or

$$\frac{s_E}{s_I} < \frac{2s_I + \frac{\delta^2}{6s_I} - \delta}{1-\delta}.$$

For $\delta = 0$, this inequality reduces to $\frac{s_E}{s_I} < 2s_I$. For a relatively inefficient incumbent with an s_I being equal to 0.1 this requires the entrant to be 5 times more efficient. For a more efficient incumbent, the entrant needs to be increasingly more efficient.

6 Discussion and Conclusion

This paper asks to what extent online platforms may learn how consumers value products on the basis of observing the consumers' search and purchase behavior. Importantly, this potentially is an environment of dual learning as, depending on the platforms' ranking algorithms, consumers may also learn on the basis of the utility of inspected items what the value of not-yet-inspected items is. A monopoly platform that cares about learning in the long run will experiment with product rankings and eventually prominently rank only products that early consumers have bought. This benefits consumers in the search market as they will only accept to buy products that are strictly better than if their next search was a random item. In this paper, we have not emphasized the welfare properties of experimentation as the platform may use their knowledge regarding consumer preferences in other markets where consumers (or advertisers) are harmed. These other markets are not explicitly modelled in this paper.

Competition between platforms may severely limit the ability of platforms to learn if consumers do not know which platform previous consumers have used. If consumers do have access to this information, for example in the form of platforms' market shares, then consumers prefer to use the platform with the largest market share, allowing it to experiment to some extent without losing consumers to its competitors. This will lead to a user (data) driven monopoly with the same long-run ability to learn.

In order to focus on learning from consumer search behaviour, we have treated firm behaviour as exogenous throughout the paper. An interesting next step in the analysis is to endogenize firm behaviour, for example, by endogenizing pricing decisions or allowing firms to have some information about consumer values that they could signal in their price-setting behavior. Another interesting avenue for future research is to allow products to differ vertically and consumers to differ in their value for quality. Will this give rise to echo chambers where some platforms cater to consumers with a high willingness to pay for quality and other platforms that do not? Finally, where the current paper focuses on what platforms may learn from observing the search and purchase behaviour of different consumers, platforms may also be able to learn about individual characteristics on the basis of consumers repeatedly using the same platform for different search queries. We see our paper as making a start with investigating what platforms may learn by creating rankings. There are many directions that can be taken from here.

7 Appendix: proofs

Proof of Lemma 2.

The proof is by induction. As with a non-experimenting platform consumers 1 and 2 face a fully random ranking and a fully random ranking after the top spot respectively, we start by considering consumer 3. Clearly the best item to place on the top spot for consumer 3 is the item bought by consumer 2. There is a positive probability that consumer 2 rejected the item that was placed on the top spot for them and bought another item. Thus, there is a positive probability that for consumer 3 the platform has a choice which item to place after the top spot. Suppose the ranking is such that the platform puts the item that was bought by consumer 1 and not bought by consumer 2 on the second spot. After having rejected the item of the top spot, will consumer 3 inspect that item or will the consumer prefer to skip that item and continue to inspect random items?

Inspecting random items yields an expected utility of u_R^* as the consumer can always continue to inspect other random items if the first items turn out to generate a utility $u < u_R^*$. Inspecting a rejected item instead (after having rejected the top-spot item, and

not coming back to the item in the top spot) yields an expected utility of

$$\Pr\left(u \geq u_R^* \,|v \in [\dot{v}, v_R^*]\right) E(u|u \geq u_R^*, v \in [\dot{v}, v_R^*]) + \Pr\left(u < u_R^* \,|v \in [\dot{v}, v_R^*]\right) u_R^* - s \quad (6)$$

$$< \Pr\left(u \geq u_R^* \,|v \in [\dot{v}, v_R^*]\right) [(1-\delta)v_R^* + \delta Em] + \Pr\left(u < u_R^* \,|v \in [\dot{v}, v_R^*]\right) u_R^* - s,$$

where $\dot{v} := \frac{u_R^* - \delta \bar{m}}{1-\delta}$ is the lowest common value that is accepted if the cutoff utility is u_R^*. The first term in the first line of (6) stands for the possibility that this item generates a higher utility than the consumer's subsequent continuation value u_R^*, in which case it is bought. The second term accounts for the possibility that the rejected item's utility is below u_R^* in which case the consumer continues. The inequality follows because $E(u|u \geq u_R^*, v = \dot{v}) = u_R^*$ and $E(u|u \geq u_R^*, v = v_R^*) = (1-\delta)v_R^* + \delta Em$ so that for all $v' \in [\dot{v}, v_R^*)$, $E(u|u \geq u_R^*, v = v') < (1-\delta)v_R^* + \delta Em$.

The second line in (6) is smaller than $u_R^* = (1-\delta)v_R^* + \delta \underline{m}$, if, and only if,

$$\delta(Em - \underline{m}) < \frac{s}{\Pr\left(u \geq u_R^* | v \in [\dot{v}, v_R^*]\right)},$$

which is certainly the case if $\delta < \hat{\delta}$.

Thus, on a non-experimenting platform consumer 3 optimally first explores the item in the top spot (which is the item bought by consumer 2). If that item has a utility $u \geq u_R^*$ they will buy it and stop searching, while if it has a utility $u < u_R^*$ they will continue searching random items, will never return to the item on the top spot and will also not inspect the item on the second spot that was rejected by consumer 2. The same conclusion would hold if the platform put the rejected item with positive probability on some spots: all these spots will be skipped in favour of inspecting random items and as there are infinitely many of them, there will always be such items to inspect.

Let us then suppose that all consumers $t = 1, ..., T-1$ first explore the item in the top spot (which is the item bought by the consumer $t-1$). If that item has a utility $u \geq u_R^*$ consumer t will buy it and stop searching, while if it has a utility $u < u_R^*$ they will continue searching random items, never return to the item on the top spot, and also not inspect any spot where an item that has been rejected by one or more of consumers $1, ..., T-1$ is placed with positive probability. It is then clear that it is optimal for consumer T to follow the same search rule as all rejected items have an expected value that is smaller than $(1-\delta)v_R^* + \delta Em$ and therefore for each next search it would be better to choose a random item than a previously rejected item.

Detailed calculations for the example in Section 3.

Assume, without loss of generality, that the product bought by consumer 5 happens to be ranked first for consumer 7 if the platform doesn't randomise the ranking for consumer

7. Consider first that consumer 7 observes $v = \hat{v} - \varepsilon$ at their first search. After observing this v the consumer updates the probability the platform is randomizing as

$$\frac{\rho f(v)}{\rho f(v) + (1-\rho)\frac{f(v)\left(F(\hat{v})-F(v_R^*)\right)}{\left[1-F(v_R^*)\right]^2}} = \frac{\rho}{\rho + (1-\rho)\frac{F(\hat{v})-F(v_R^*)}{\left[1-F(v_R^*)\right]^2}} =: \mu_L.$$

This is because the only way consumer 7 may observe $v = \hat{v} - \varepsilon$ when the platform is not randomizing is if consumer 5 buys such an item. For this to be the case it should be that one of the two top-ranked items for consumer 5 had a value $v' \leq \hat{v} - \varepsilon$ so that the better of the two top-ranked items of consumer 5 has $v = \hat{v} - \varepsilon$. We can approximate the probability that one of the items consumer 5 observed had a value $v' \leq \hat{v} - \varepsilon$ by $(F(\hat{v}) - F(v_R^*))/(1 - F(v_R^*))$. Thus, the continuation value of making another search is (for ε small) approximately

$$\begin{aligned} &\mu_L \left([1 - F(\hat{v})]\, E(v|v \geq \hat{v}) + F(\hat{v})\hat{v}\right) + \\ &(1-\mu_L)\left[E(v|v \geq \hat{v}) - \frac{\left(F(\hat{v}) - F(v_R^*)\right)^2}{\left(1 - F(v_R^*)\right)^2}\left(E(v|v \geq \hat{v}) - \hat{v}\right)\right] - s \end{aligned} \tag{7}$$

because, if the platform does not randomize, consumer 7 only gets an item with $v < \hat{v}$ on their next search with probability $\frac{\left(F(\hat{v})-F(v_R^*)\right)^2}{\left(1-F(v_R^*)\right)^2}$.

Consider next that consumer 7 observes $v = \hat{v} + \varepsilon$ at their first search. The posterior probability that the platform is randomizing after observing such a v is

$$\frac{\rho f(v)}{\rho f(v) + (1-\rho)\frac{f(v)}{1-F(v_R^*)}\left[\frac{1}{2} + \frac{1}{2}\frac{F(\hat{v})-F(v_R^*)}{1-F(v_R^*)}\right]} = \frac{\rho}{\rho + \frac{1-\rho}{1-F(v_R^*)}\left[\frac{1}{2} + \frac{1}{2}\frac{F(\hat{v})-F(v_R^*)}{1-F(v_R^*)}\right]} =: \mu_H.$$

This can happen if consumer 5 encounters an item with $v = \hat{v} + \varepsilon$ on their first search and immediately buys it, or if consumer 5 first searches the other item with a value $v \in [v_R^*, \hat{v})$. Thus, after observing a $v = \hat{v} + \varepsilon$, the continuation value of making another search for consumer 7 is (for ε small) approximately

$$\begin{aligned} &\mu_H \left([1 - F(\hat{v})]\, E(v|v \geq \hat{v}) + F(\hat{v})\hat{v}\right) + \\ &(1-\mu_H)\left[E(v|v \geq \hat{v}) - \frac{\left(F(\hat{v}) - F(v_R^*)\right)^2}{\left(1 - F(v_R^*)\right)^2}\left(E(v|v \geq \hat{v}) - \hat{v}\right)\right] - s \end{aligned} \tag{8}$$

We can now make the following observations. First, if ρ is close to 1, consumer 7 wants to stop searching at their first search both when $v = \hat{v} - \varepsilon$ and when $v = \hat{v} + \varepsilon$. The reason is that the continuation value of search is very close to the continuation value of search of consumers 1-4 and they have a reservation value strategy to stop searching if $v > v_R^*$. Second, if ρ is close to 0, consumer 7 always wants to continue searching both

when $v = \hat{v} - \varepsilon$ and when $v = \hat{v} + \varepsilon$. The reason is that the continuation value of search is very close to

$$E(v|v \geq \hat{v}) - \frac{(F(\hat{v}) - F(v_R^*))^2}{(1 - F(v_R^*))^2} (E(v|v \geq \hat{v}) - \hat{v}) - s$$

and this is strictly larger than the continuation value of search of consumers 5 and 6 as represented by the RHS of (2) as $\frac{F(\hat{v})-F(v_R^*)}{1-F(v_R^*)} < 1$.

Third, the continuation value of search after observing a $v = \hat{v} + \varepsilon$ is strictly larger than the continuation value of search after observing a $v = \hat{v} - \varepsilon$, i.e., (8) is strictly larger than (7), if the weight on the first term is larger in (7), i.e., $\mu_L > \mu_H$. This is the case if

$$\frac{1}{2} + \frac{1}{2} \frac{F(\hat{v}) - F(v_R^*)}{1 - F(v_R^*)} > \frac{F(\hat{v}) - F(v_R^*)}{1 - F(v_R^*)},$$

which is always true as for any $s > 0$ we have $F(\hat{v}) < 1$.

These three observations imply that there are values of ρ such that consumer 7 wants to continue to search after observing a $v = \hat{v}+\varepsilon$ and wants to abort search after observing a $v = \hat{v} - \varepsilon$.

Proof of Proposition 5.

We divide the proof into 4 steps for clarity.

(1) An interior v^* exists and is unique if $\delta < \hat{\delta}$. A consumer drawing the minimum possible utility level given $v \geq v^*$, $(1 - \delta)v^* + \delta\underline{m}$, buys if

$$(1 - \delta)v^* + \delta\underline{m} \geq (1 - \delta)E(v|v \geq v^*) + \delta Em - s,$$

which gives

$$v^* \geq E(v|v \geq v^*) + \frac{\delta(Em - \underline{m}) - s}{1 - \delta}. \tag{9}$$

Clearly inequality (9) can hold only if $\delta(Em - \underline{m}) < s$ or $\delta < \hat{\delta}$. We can write (9) as

$$v^* \geq \frac{\int_{v^*}^{\bar{v}} vg(v)dv}{1 - G(v^*)} + \frac{\delta(Em - \underline{m}) - s}{1 - \delta} = \frac{\bar{v} - v^*G(v^*) - \int_{v^*}^{\bar{v}} G(v)dv}{1 - G(v^*)} + \frac{\delta(Em - \underline{m}) - s}{1 - \delta}$$

$$= v^* + \frac{\bar{v} - v^* - \int_{v^*}^{\bar{v}} G(v)dv}{1 - G(v^*)} + \frac{\delta(Em - \underline{m}) - s}{1 - \delta}.$$

Thus, we have that an interior v^* is implicitly defined by

$$v^* = \bar{v} - \int_{v^*}^{\bar{v}} G(v)dv + (1 - G(v^*)) \frac{\delta(Em - \underline{m}) - s}{1 - \delta}. \tag{10}$$

Equation (10) can be rewritten as (4). Since the RHS of equation (4) is positive, we also need the LHS to be positive for an interior v^* to exist. Now

$$\lim_{v^* \to \underline{v}} \int_{v^*}^{\bar{v}} \frac{1-G(v)}{1-G(v^*)} dv = \int_{\underline{v}}^{\bar{v}} 1 - G(v) dv = Ev - \underline{v} > 0,$$

and

$$\lim_{v^* \to \bar{v}} \int_{v^*}^{\bar{v}} \frac{1-G(v)}{1-G(v^*)} dv = \lim_{v^* \to \bar{v}} \frac{-(1-G(v^*))}{-g(v^*)} = \frac{0}{g(\bar{v})} = 0,$$

as $g(\bar{v}) > 0$. Thus, a sufficient condition for an interior solution for v^* to exist is $Ev - \underline{v} > \frac{s - \delta(Em - \underline{m})}{1-\delta}$ or

$$(1-\delta)(Ev - \underline{v}) + \delta(Em - \underline{m}) > s.$$

This inequality holds because, by assumption, both $Ev - \underline{v} > s$ and $Em - \underline{m} > s$.

A sufficient condition for v^* to be uniquely defined is that the derivative of the LHS of equation (4), or

$$\frac{\partial}{\partial v^*} \int_{v^*}^{\bar{v}} \frac{1-G(v)}{1-G(v^*)} dv = -1 + \int_{v^*}^{\bar{v}} \frac{(1-G(v))g(v^*)}{(1-G(v^*))^2} dv, \tag{11}$$

is negative for all $v^* \in (\underline{v}, \bar{v})$. The derivative can be rearranged so that it is negative if for all v^*

$$g(v^*) \int_{v^*}^{\bar{v}} \frac{1-G(v)}{1-G(v^*)} dv < 1 - G(v^*). \tag{12}$$

We show that inequality (12) holds because the RHS of the inequality decreases faster than the LHS for all $v^* < \bar{v}$, while the two sides equal at $v^* = \bar{v}$. At $v^* = \bar{v}$, using l'Hopital's rule, the LHS equals $g(\bar{v}) \frac{-(1-G(\bar{v}))}{-g(\bar{v})} = g(\bar{v}) \cdot 0 = 0$ and the RHS is also equal to zero. Taking the derivative with respect to v^* on both sides of inequality (12) gives $-g(v^*) \left[1 - \int_{v^*}^{\bar{v}} \frac{(1-G(v))g(v^*)}{(1-G(v^*))^2} dv\right] + g'(v^*) \int_{v^*}^{\bar{v}} \frac{(1-G(v))}{1-G(v^*)} dv$ for the LHS and $-g(v^*)$ for the RHS. Thus, the derivative of the LHS of (12) is larger than the derivative of the RHS if

$$\left[\frac{g^2(v^*)}{1-G(v^*)} + g'(v^*)\right] \int_{v^*}^{\bar{v}} \frac{(1-G(v))}{1-G(v^*)} dv > 0.$$

This is indeed the case if $\frac{g^2(v^*)}{1-G(v^*)} + g'(v^*) > 0$ which follows from the fact that $1-G(v)$ is logconcave: logconcavity implies that for the first and second derivative of $\ln(1-G(v))$ we have $\frac{-g}{1-G}$ and $\frac{-g'(1-G)-g^2}{(1-G)^2} = \frac{-1}{1-G}\left[g' + \frac{g^2}{1-G}\right] \leq 0$. In sum, an interior v^* exists and is unique if $\delta(Em - \underline{m}) < s$.

(2) The platform can never learn more than that $v \geq v^*$. Learning more than that $v \geq v^*$ would require that consumers at some point reject items that have v^* as their common component, but we will show that items with a $v \geq v^*$ will never be rejected by consumers. Consumer $t = 1$ and $t = 2$ use a random cutoff rule u_R^* and all items with a $v \geq v_R^*$ are immediately accepted (that is, bought after being inspected). This implies

that all $v \geq v^*$ are accepted as $v^* > v_R^*$. Suppose then by induction that all consumers up to and including T accept all items with a $v \geq v^*$. We will argue that then also all items with a $v \geq v^*$ are accepted by consumer $T+1$. The only reason why consumer $T+1$ would not want to accept some item with a $v \geq v^*$ is that they become more positive about observing an item with an even higher utility on their next search(es). But they can only become more positive if they think that the platform may be able to distinguish between items with a $v \geq v^*$ based on choices of consumers $t = 1, ..., T$. But consumers know this is not the case (because of the induction assumption).

(3) The platform can achieve $lim_{t\to\infty} v_1(t) \geq v^*$**.** We show that the platform can do so by using the following algorithm:

(i) split consumers into cohorts $j = 1, 2, ...$, each of a fixed finite size $K > 2$,

(ii) show each consumer in cohort $j = 1$ a different ranking of infinite length that contains randomly ordered items,

(iii) for each cohort $j > 1$,

- call items that at least one member of cohort $j-1$ immediately bought after the first inspection "qualifying items" and denote the number of them by k_{j-1};
- If $k_{j-1} \in \{0, 1\}$, restart the process at step (i) by treating cohort j like cohort 1: show cohort j a ranking of infinite length that contains randomly ordered unrepeated items;
- if $k_{j-1} \in \{2, ..., K\}$, (a) rank the qualifying items on the top spots, (b) rotate the first-ranked item so that all qualifying items are shown on rank 1 for at least one consumer in cohort j, and show the other qualifying items randomly on spots $2, ..., k_{j-1}$, (c) fill the spots $k_{j-1}+1$ and $k_{j-1}+2$ with copies of a randomly drawn uninspected item and fill the j' spots after the top $k_{j-1}+2$ spots with copies of another randomly drawn uninspected item, where j' is the cohort with which the process was last restarted.

Consider first cohort $j = 1$. Clearly all consumers in cohort 1 use the random cutoff u_R^* and purchase the first item they inspect with $u \geq u_R^*$. Since each consumer of cohort 1 faces a different ranking, $k_1 = K$.

Consider $j = 2$. All members of cohort 2 know that the top k_1 ranks contain distinct items where a consumer of cohort 1 drew $u \geq u_R^*$. We first make two remarks on the stopping problem of consumers in cohort 2.

First, consumers in cohort 2 can search over randomly ordered items outside the platform at any point so their value of searching beyond rank k_1 (excluding the recall value) is u_R^*. Thus, they search beyond rank k_1 only if for all the k_1 ranked items $u < u_R^*$.

Second, after having seen the ranking and inspected any set of items, the consumer's belief about the distribution of common values of the remaining items is unchanged as the platform ranks items in the qualifying set randomly (beyond the first rank).

As a result of these two observations, we can focus on the consumer's optimal stopping problem when they can make at most k_1 independent draws from a utility distribution truncated from below. Given that the consumer's stopping problem does not involve learning about the distribution of items in the ranking, the consumer's optimal policy can be calculated using the one-step-look-ahead rule.

In particular, for a consumer of cohort 2, after inspecting the first-ranked item and uncovering utility u, the value of inspecting one more item, uncovering u' and then stopping with the better of the two items is

$$V(u) := P(u' \geq u | u'' \geq u_R^*) E(u' | u' \geq u, u'' \geq u_R^*) + P(u' < u | u'' \geq u_R^*) u - s,$$

where u' is the utility draw that the searching consumer gets for the next item and u'' is the utility draw of the previous consumer who bought it. (A next such item to search exists because $k_1 = K > 2$.) The consumer is just indifferent between stopping with u and continuing if $u = u_2$ such that $u_2 = V(u_2)$. The solution for u_2 clearly satisfies $u_2 > u_R^*$. Thus, a consumer in cohort $j = 2$ will stop and purchase the first item among the k_1 top-ranked items if they find one with $u \geq u_2$. If not, they return to the item with the highest u among the k_1 inspected items if it exceeds u_R^* and continue searching over randomly ordered items (on or off the platform) otherwise.

Some consumers in cohort 2 end up buying an item bought by some consumer in cohort 1, with some buying after an initial inspection and some after returning (and other members of cohort 2 may buy an item off the platform). Thus, after consumers in cohort 2 have made their purchase decisions, some number $k_2 \leq k_1$ of items qualify to be shown to cohort 3 in the top k_2 ranks: these were bought by consumers in both cohorts 1 and 2 and the consumers in cohort 2 bough them after first inspection. All of these items had $u \geq u_2$ for consumers in cohort 2.

Consider $j = 3$. With cohort 3, two things can happen. If $k_2 < 2$, then consumers in cohort 3 are treated like cohort $j = 1$, i.e., shown an randomly ordered ranking of unrepeated items, so the process is essentially restarted. Note that consumers in cohort $j'' \geq 4$ (until the next restart) are then shown a ranking where the qualifying items (and the two copies after these) are followed by $j' = 3$ copies of one item so they know that the process restarted last with cohort 3 and that they are treated like cohort $j'' - j' + 1$.

If, instead, $k_2 \geq 2$, then consumers in cohort 3 face a ranking where the top k_2 items satisfied $u \geq u_2$ for cohort 2. The consumers can infer k_2 from the fact that an item is copied on spots $k_2 + 1$ and $k_2 + 2$. In this case, the appropriately modified remarks that we made for cohort 2 hold also for cohort 3. We can, thus, focus on the consumer's

optimal stopping problem when they can make at most k_2 independent draws from a utility distribution truncated from below. The truncation is given by the cutoff value $v_2 > v_R^*$ where v_2 satisfies $u_2 = (1-\delta)v_2 + \delta\underline{m}$.

Thus, the reservation value to decide whether to buy the first-ranked item after initial inspection of a consumer in cohort 3, u_3, is derived analogously to u_2. Explicitly, u_3 solves

$$u_3 = P(u \geq u_3 | u' \geq u_2) E(u | u \geq u_3, u' \geq u_2) + P(u < u_3 | u' \geq u_2) u_3 - s,$$

where u_3 clearly satisfies $u_3 > u_2$.

By a similar argument, for any cohort $j+1$ that is shown a ranking that contains qualifying items, the cutoff utility u_{j+1} satisfies $u_{j+1} > u_j$.

The only thing left to prove is that eventually as a result of this process, there are at least two items with $v \geq v^*$ uncovered by a cohort that faces a ranking that contains no copies. Suppose otherwise: that either no or only a single item out of K have $v \geq v^*$ conditional on having $u \geq u_R^*$. Let the lowest possible v such that $u \geq u_R^*$ be denoted $\dot{v}$ so that $\dot{v} := \max\{v_R^* - \frac{\delta(\bar{m}-\underline{m})}{1-\delta}, 0\}$. Thus, the probability of the event of interest happening for a single cohort that faces a ranking without copies is at most

$$\rho_K := P(v < v^* | v \geq \dot{v})^K + K \times P(v \geq v^* | v \geq \dot{v}) P(v < v^* | v \geq \dot{v})^{K-1}$$

$$= \left(\frac{F(v^*) - F(\dot{v})}{1 - F(\dot{v})} \right)^K + K \frac{1 - F(v^*)}{1 - F(\dot{v})} \left(\frac{F(v^*) - F(\dot{v})}{1 - F(\dot{v})} \right)^{K-1} < 1.$$

If this happens for all cohorts that face a ranking that contains no copies, then the process will always be restarted at some point, in total infinitely many times. But then an upper bound on the total probability that no or only a single item of K have $v \geq v^*$ conditional on having $u \geq u_R^*$ goes to zero: $\lim_{n\to\infty} \rho_K^n = 0$.

(4) The interior cutoff v^* strictly increases in δ, decreases in s, and exceeds v_R^*. It is easy to show that the RHS of (4) decreases in δ and increases in s, but does not depend on v^*. The LHS of (4) depends directly on neither δ nor s, but decreases in v^* (by step (1)). These facts together establish that $\frac{\partial v^*}{\partial \delta} > 0$ and $\frac{\partial v^*}{\partial s} < 0$.

Finally, we show that $v^* > v_R^*$ if $\delta < \hat{\delta}$. Recall that v_R^* is defined in equation (1) for $\delta < \hat{\delta}$. For ease of reading, we let $P(\triangle)$ and $P(\square)$ denote the probabilities with which an item has a utility either in the triangle or in the rectangle as indicated in Figure 1 (for $\hat{v} = v_R^*$):

$$P(\triangle) := \int_{\dot{v}}^{v_R^*} \int_{\underline{m} + \frac{(1-\delta)(v_R^* - v)}{\delta}}^{\bar{m}} f(m) g(v) dm dv,$$

where, as in step (3), $\dot{v} := \max\{v_R^* - \frac{\delta(\bar{m}-\underline{m})}{1-\delta}, 0\}$ denotes the lowest possible v such that $u = u_R^*$, and $P(\square) := 1 - G(v_R^*)$. Then we can rewrite (1) as

$$(1-\delta)v_R^* + \delta\underline{m} = P(\square) \left[(1-\delta) E(v | v > v_R^*) + \delta E(m) \right] \tag{13}$$

$$+P(\triangle)E(u|u\in\triangle)+(1-P(\triangle)-P(\square))\left[(1-\delta)v_R^*+\delta\underline{m}\right]-s,$$

i.e., with probability $P(\square)$ ($P(\triangle)$) the item's u is "in the rectangle" (triangle) and the consumer gets expected utility of u being in the rectangle (triangle). With the remaining probability the product generates utility below u_R^* and the consumer continues with value $(1-\delta)v_R^*+\delta\underline{m}$.

Recall that for the experimenting monopolist, the cutoff v^* is defined by

$$(1-\delta)v^*+\delta\underline{m}=(1-\delta)E(v|v\geq v^*)+\delta E(m)-s. \tag{14}$$

Now the LHS of (14) and (13) as functions of v_R^* and v^* respectively are identical. The RHS of (13) is lower than in (14) for $v_R^*=v^*$ as long as $E(u|u\in\triangle)<E(u|u\in\square)=(1-\delta)E(v|v>v^*)+\delta E(m)$.

To show that $E(u|u\in\triangle)<E(u|u\in\square)$ we argue that $E(u|v=\tilde{v})$ increases in $\tilde{v}$ if $\tilde{v}$ is in the support of the triangle, i.e., if $\tilde{v}\in[\dot{v},v^*]$. Let $\tilde{m}:=\underline{m}+\frac{(1-\delta)(v^*-\tilde{v})}{\delta}$. Then

$$E(u|v=\tilde{v})=(1-\delta)\tilde{v}+\delta E(m|m>\tilde{m})=(1-\delta)\tilde{v}+\frac{\delta}{1-F\left(\tilde{m}\right)}\int_{\tilde{m}}^{\bar{m}}mf(m)dm$$

$$=(1-\delta)\tilde{v}+\delta\tilde{m}+\delta\int_{\tilde{m}}^{\bar{m}}\frac{1-F(m)}{1-F\left(\tilde{m}\right)}dm=(1-\delta)v^*+\delta\underline{m}+\delta\int_{\tilde{m}}^{\bar{m}}\frac{1-F(m)}{1-F\left(\tilde{m}\right)}dm,$$

where the one but last equality follows from integrating by parts. Now

$$\frac{\partial E(u|v=\tilde{v})}{\partial\tilde{v}}=(1-\delta)\left[1-\int_{\tilde{m}}^{\bar{m}}\frac{(1-F(m))f\left(\tilde{m}\right)}{\left(1-F\left(\tilde{m}\right)\right)^2}dm\right],$$

which is positive as long as the term in the squared brackets is positive. But showing that this term is positive is analogous to showing that the expression in (11) is negative, which we did in step (1). Thus, $E(u|u\in\triangle)<E(u|u\in\square)$ and the solution to (13) is smaller than for (14): $v_R^*<v^*$.

Proof of Proposition 7.

The algorithm ranks items below the top spot randomly so that the consumers always have a random continuation value u_R^*. To specify precisely which item the algorithm puts on the top spot and when, if ever, to change it, we introduce some preliminaries. Let the lowest acceptable level of the common component given u_R^* be denoted $\dot{v}$, i.e., items with $v<\dot{v}$ are not accepted even if the consumer gets the highest idiosyncratic draw. Note that any item j with a common component $v_j\in[\dot{v},\bar{v}]$ can be characterised by $\tilde{m}_j\in[\underline{m},\bar{m}]$ so that consumer t buys item j iff $m_{tj}\geq\tilde{m}_j$, with probability $Pr(m_{tj}\geq\tilde{m}_j)=1-F(\tilde{m}_j)=:p_j$, where $p_j\in(0,1]$ due to the fact that $s>0$. Thus, from the platform's perspective a consumer inspecting item j can be thought of as a Bernoulli trial

with success probability p_j. Note that $\tilde{m}_j$ decreases in v_j. For $v_j = \bar{v} - \eta =: v_\eta$, denote $Pr(m_{tj} \geq \tilde{m}_j) =: p_\eta$ and for $v_j = \bar{v}$, denote $Pr(m_{tj} \geq \tilde{m}_j) =: p_R$. The proof focusses on the nontrivial case where η is small so that $\dot{v} < v_\eta$.

Below we show that the platform can achieve any desired combination of precisions $\tau > 0$ and $\eta > 0$. To do so, we will introduce three other errors ζ, α and ϵ, that must (and can) be chosen small enough to achieve the desired τ (and η).

Each item j that has been inspected by n consumers is associated with the random variable X_j that counts the number of consumers of n who bought j. Thus, X_j is a binomially distributed random variable with parameters n and p_j, $X_j \sim B(n, p_j)$. The parameter p_j can be estimated by the realised proportion of purchases $\hat{p}_j = \frac{x_j}{n}$. By the central limit theorem, for n large enough, the sampling distribution of $\hat{p}_j$ becomes arbitrarily close to the normal distribution with mean p_j and variance $\frac{\sigma_j^2}{n}$. More precisely, the Berry-Esseen theorem adjusted for the binomial distribution shows that the maximal approximation error between the normal and the binomial distribution $B(n, p)$ (as measured by the Kolmogorov-Smirnov distance) is

$$\sup_{x \in \mathbf{R}} \left| Pr\left(\frac{X_j - np_j}{\sqrt{np_j(1-p_j)}} \leq p \right) - \Phi(p) \right| \leq \frac{C[p_j^2 + (1-p_j)^2]}{\sqrt{np_j(1-p_j)}},$$

where $\Phi(p)$ is the cdf of the standard normal distribution and $C = \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \approx 0.40973$, a constant (Schulz, 2016). The Berry-Esseen theorem implies that if the maximal tolerated approximation error is some $\zeta > 0$, then the approximation error is smaller than ζ if $n > \bar{n}_\zeta$, where

$$\bar{n}_\zeta = \frac{C^2[p_j^2 + (1-p_j)^2]^2}{p_j(1-p_j)\zeta^2}.$$

Note that $\bar{n}_\zeta$ is finite for any $p_R \in (0, 1]$ because, with probability one, all the items that have been bought at least once have $v \in (\dot{v}, \bar{v})$, thus, $p_j \in (0, 1)$. As a result, any desired $\zeta > 0$ can be achieved with a finite number of observations n by choosing $n \geq \bar{n}_\zeta$. We will show later that a "small enough" ζ must be chosen to achieve a desired $\tau > 0$.

We are now ready to introduce the algorithm:

1. Rank the item bought by consumer $t = 1$, denoted by A, on the top spot for the next $n' - 1$ consumers, where $n' > \bar{n}$ is a finite constant. Calculate the sample mean $\hat{p}_A(n')$. If the null hypothesis $H_0 : p_A \geq p_\eta$ is not rejected with a significance level α, keep item A on the top spot and move to consumer $t = n' + 1$. Otherwise, restart the process, calling the restart period $t = 1$ again.

2. For all $t' \geq n'+1$, calculate the sample mean $\hat{p}_j(t')$ over the entire sample $t = 1, ..., t'$ where j was the item on the top spot. If the null hypothesis $H_0 : p_j \geq p_\eta$ is not rejected with a significance level α, keep item j on the top spot and move to consumer $t' + 1$. Otherwise, restart the process, calling the initial period $t = 1$.

We now argue that this algorithm ensures that the platform finds out if the random variable X_j associated with a top-ranked item j has a population mean p_j close to the target population mean p_R, where "close to" is defined as $p_j \geq p_\eta$. We formulate the platform's problem in the terminology of hypothesis testing: if the top-placed product is j, the null hypothesis is $H_0 : p_j \geq p_\eta$, and the alternative hypothesis is $H_1 : p_j < p_\eta$ (so the platform uses a lower-tail test). Denote the nominal significance level chosen by the platform, i.e., the nominal probability of type I error, by α. Since the platform chooses α, it can choose an $\alpha > 0$ as small as it wants. We will show later that a "small enough" α must be chosen to achieve a desired $\tau > 0$.

The platform is interested in not rejecting H_0 if H_0 is true for the top-spot item (the probability of which is affected by the platform's chosen nominal probability of type I error α and, as explained below, the normal-binomial approximation error ζ), and rejecting H_0 if H_0 is false for the top-spot item (the probability of which is affected by the nominal probability of type II error β and the normal-binomial approximation error ζ).

We claim that, if n consumers (with $n > \bar{n}_\zeta$) are shown item j first and the true common component of j is $v_j \geq v_\eta$, then with probability at least $1 - \alpha - \zeta$ the sample mean $\hat{p}_j$ lies within the acceptance region and the null hypothesis is not rejected. In particular, the approximation error ζ increases the probability of type I error (rejecting H_0 when H_0 is true) the most if the difference in the normal cdf and the approximated binomial cdf happens to be largest at the least favourable acceptance probability for H_0. The total probability of type I error is, thus, at most $\alpha + \zeta$. In other words, conditional on an item with $v_j \geq v_\eta$ being ranked first, j will remain first with probability $1 - \alpha - \zeta$ for all sample sizes $n' \geq n > \bar{n}_\zeta$.

The platform also wants to ensure that an item that is ranked on the top spot and has $v < v_\eta$ is removed from the top spot with a high probability. In particular, we argue that if the true mean of j is $p_j = p_\eta - \lambda$ for some $\lambda > 0$, the nominal probability of type II error, β, goes to zero if n increases for standard levels of $\alpha + \zeta$. To do that, note that the platform must use the t-test statistic in the hypothesis testing because each random variable X_j associated with an item j s.t. $v_j \geq v_\eta$ has a different population variance σ_j and, since the platform does not know exactly which item it is evaluating, it does not know the population variance and, hence, the variance of the sampling distribution of the sample mean of X_j that it faces. The probability of a type II error β (for a lower-tail test) is given by $\beta = Pr(\hat{p}_j > p_\eta - t_{n-1,\alpha}\frac{s}{\sqrt{n}} | p_j < p_\eta)$ where s is the sample standard deviation of X_j and $t_{n-1,\alpha}$ the critical value. Thus, if the true population mean of X_j is $p_j = p_\eta - \lambda$, the nominal probability of type II error is

$$\beta = Pr\left(t > \frac{p_\eta - t_{n-1,\alpha}\frac{s}{\sqrt{n}} - (p_\eta - \lambda)}{s/\sqrt{n}}\right) = Pr\left(t > \frac{\lambda}{s/\sqrt{n}} - t_{n-1,\alpha}\right).$$

For n large enough, the critical value $t_{n-1,\alpha}$ grows at a slower rate than $\sqrt{n}$, whereas $\frac{\lambda}{s/\sqrt{n}}$ grows in n at rate $\sqrt{n}$. Thus, β disappears as n increases.

Similarly to above, the approximation error ζ increases the total probability of a type II error the most if the difference between the binomial and normal cdfs happens to be largest at the most favourable acceptance probability for H_0. Thus, the total probability of type II error is at most $\beta + \zeta$. Because β vanishes as n becomes large enough, a finite number $\bar{n}_\beta$ exists such that for all $n \geq \bar{n}_\beta$, the total probability of type II error $\beta + \zeta$ is at most $\zeta + \epsilon$ for some $\epsilon > 0$. The error $\epsilon > 0$ can again be chosen to be as small as desired by choosing $n \geq \bar{n}_\beta$ for the appropriate (finite) $\bar{n}_\beta$.

Altogether, by setting $n' > \bar{n} := \max\{\bar{n}_\beta, \bar{n}_\zeta\}$ in the algorithm, the platform ensures that the first-ranked item has $v \geq \bar{v} - \eta$ with a probability at least $1 - \tau$ where τ solves

$$1-\tau = Pr(v_A \geq v_\eta | v_A \geq \dot{v})[(1-\alpha-\zeta)+(\alpha+\zeta)(1-\tau)]+P(v_A < v_\eta | v_A \geq \dot{v})(1-\zeta-\epsilon)(1-\tau).$$

If the first-bought item has $v_A \geq v_\eta$, it remains ranked first with a probability at least $1 - \alpha - \zeta$. If it is discarded, the process is started anew. If the first-bought item has $v_A \in [\dot{v}, v_\eta)$, it is discarded with probability at least $1 - \zeta - \epsilon$, in which case the process is started anew again. Solving this equation for τ gives

$$\tau = \frac{(G(v_\eta) - G(\dot{v}))(\zeta + \epsilon)}{(1 - \alpha - \zeta)(1 - G(v_\eta)) + (G(v_\eta) - G(\dot{v}))(\zeta + \epsilon)}.$$

Note that τ is affected by all α, ζ, and ϵ and recall that these can all be chosen as small as the platform wants. In particular, for any desired $\eta > 0$, any $\tau > 0$ can be chosen as small as the platform wants by choosing ζ and ϵ small enough.

Proof of Proposition 8.

(i) The proof that consumers prefer inspecting to buying blindly if $s < \bar{s}(\delta)$ is in the text before the Proposition. What still must show that an algorithm exists that for every $\eta, \tau > 0$ achieves that the probability that $lim_{t\to\infty} v_1(t) \geq \bar{v} - \eta$ is at least $1 - \tau$. In fact, the platform can use the same algorithm that is specified in the proof of Proposition 7: all consumers prefer inspecting to buying blindly and $v_R^* = \bar{v}$ so consumers behave exactly the same way as when blind buying is not allowed if that algorithm is used.
(ii) Consider next that $s > \hat{s}_R(\delta)$. We first focus on $\delta = 0$ (so that $\hat{s}_R(\delta) = 0$) and consider a small $s > 0$. Now the first consumer prefers to inspect (using the cutoff $v_R^* < \bar{v}$). Suppose that the platform's algorithm specifies that it ranks products randomly for the first T_1 consumers and then ranks the products bought by these first T_1 consumers on top spots in a random order for the next T_2 consumers. After inspecting any number of products less than T_1, consumer t with $T_1 < t \leq T_1 + T_2$ has a continuation value $\tilde{v}$ that

satisfies

$$\tilde{v} = \frac{1 - G(\tilde{v})}{1 - G(v_R^*)} E(v|v \geq \tilde{v}) + \frac{G(\tilde{v}) - G(v_R^*)}{1 - G(v_R^*)} \tilde{v} - s. \tag{15}$$

To prove that for sufficiently small search costs consumers t with $T_1 < t \leq T_1 + T_2$ would prefer to continue inspecting highly-ranked items if they so far only found items with a $v < \tilde{v}$ rather than blind buy and settle for $E(v|v \geq v_R^*)$, we need to prove that for sufficiently large T_1 and for small s

$$E(v|v \geq \tilde{v}) - \frac{(1 - G(v_R^*))\, s}{1 - G(\tilde{v})} > E(v|v \geq v_R^*),$$

or

$$(1 - G(\tilde{v}))\, E(v|v \geq \tilde{v}) - (1 - G(v_R^*))\, s > (1 - G(\tilde{v}))\, E(v|v \geq v_R^*).$$

Taking the total differential, we need to prove that in a neighborhood of $s = 0$

$$-g(\tilde{v})\tilde{v}d\tilde{v} - (1 - G(v_R^*))\, ds + g(v_R^*)sdv_R^* > (1 - G(\tilde{v}))\, \frac{g(v_R^*)s}{(1 - G(v_R^*))^2} dv_R^* - g(\tilde{v})d\tilde{v}E(v|v \geq v_R^*),$$

where we use the fact $(1 - G(\tilde{v}))\, E(v|v \geq \tilde{v}) = \int_{\tilde{v}} g(v)vdv$. As $\int_{\tilde{v}} g(v)(v - \tilde{v})dv = (1 - G(v_R^*))\, s$ yields in a first-order approximation that $-(1 - G(\tilde{v}))\, d\tilde{v} = (1 - G(v_R^*))\, ds - g(v_R^*)sdv_R^*$ we have

$$-g(\tilde{v})\tilde{v}d\tilde{v} + (1 - G(\tilde{v}))\, d\tilde{v} > (1 - G(\tilde{v}))\, \frac{g(v_R^*)s}{(1 - G(v_R^*))^2} dv_R^* - g(\tilde{v})d\tilde{v}E(v|v \geq v_R^*)$$

or

$$\left[1 + \frac{g(\tilde{v})}{(1 - G(\tilde{v}))} \left(E(v|v \geq v_R^*) - \tilde{v}\right)\right] d\tilde{v} > \frac{g(v_R^*)s}{(1 - G(v_R^*))^2} dv_R^*.$$

To show that this holds, it is sufficient to show that

$$\lim_{s \to 0} \frac{E(v|v \geq v_R^*) - \tilde{v}}{(1 - G(\tilde{v}))} = -\infty,$$

as $d\tilde{v}/ds$ and dv_R^*/ds are negative and $\frac{g(v_R^*)s}{\left(1 - G(v_R^*)\right)^2}$ is positive. So, the LHS would be positive and the RHS would be negative. As

$$\lim_{s \to 0} \frac{E(v|v \geq v_R^*) - \tilde{v}}{(1 - G(\tilde{v}))} = \lim_{s \to 0} \frac{-\frac{g(v_R^*)s}{(1 - G(v_R^*))^3} - \frac{\partial \tilde{v}}{\partial s}}{-g(\tilde{v})\frac{\partial \tilde{v}}{\partial s}} = \frac{1}{g(\tilde{v})} \left(\lim_{s \to 0} \frac{g(v_R^*)s}{(1 - G(v_R^*))^3 \frac{\partial \tilde{v}}{\partial s}} + 1\right)$$

$$= \frac{1}{g(\tilde{v})} \lim_{s \to 0} \left(1 + \frac{-g(v_R^*)s}{(1 - G(v_R^*))^3 \left(\frac{1 - G(v_R^*)}{1 - G(\tilde{v})} + \frac{g(v_R^*)s}{(1 - G(\tilde{v}))} \frac{1}{1 - G(v_R^*)}\right)}\right)$$

$$= \frac{1}{g(\tilde{v})} \lim_{s\to 0} \left(1 - \frac{1}{\frac{(1-G(v_R^*))^4}{(1-G(\tilde{v}))g(v_R^*)s} + \frac{(1-G(v_R^*))^2}{(1-G(\tilde{v}))}} \right) = \frac{1}{g(\tilde{v})} \lim_{s\to 0} \left(1 - \frac{1}{\frac{(g(\bar{v})(\bar{v}-v_R^*))^4}{g(\bar{v})(\bar{v}-\tilde{v})g(v_R^*)s} + \frac{(g(\bar{v})(\bar{v}-v_R^*))^2}{g(\bar{v})(\bar{v}-\tilde{v})}} \right),$$

where the first step uses l'Hopital and the third step uses the fact that $-(1-G(\tilde{v}))\,d\tilde{v} = (1-G(v_R^*))\,ds - g(v_R^*)s dv_R^*$ and $-(1-G(v_R^*))\,dv_R^* = ds$ yields that in a neighborhood of $s=0$,

$$\frac{\partial \tilde{v}}{\partial s} = -\frac{(1-G(v_R^*))}{(1-G(\tilde{v}))} - \frac{g(v_R^*)s}{(1-G(v_R^*))\,(1-G(\tilde{v}))}.$$

Now, as $s = \int_{v_R^*}^{\bar{v}} (1-G(v))dv =: h(v_R^*)$, and we know that $v_R^*(s) = \bar{v}$ at $s=0$ we can give a second-order Taylor approximation of $h(v_R^*)$ in a neighborhood of $\bar{v}$ as

$$h(v_R^*) = 0 + (1-G(\bar{v}))(\bar{v}-v_R^*) + \frac{1}{2}g(\bar{v})(\bar{v}-v_R^*)^2 + o\left((\bar{v}-v_R^*)^2\right) = \frac{1}{2}g(\bar{v})(\bar{v}-v_R^*)^2 + o\left((\bar{v}-v_R^*)^2\right),$$

where $o(\cdot)$ is Laudau's little-o notation. Thus, we have that in a neighborhood of $s=0$,

$$\bar{v} - v_R^* = \sqrt{\frac{2s}{g(\bar{v})}} + o(\sqrt{s}).$$

Similarly, as $\int_{\tilde{v}}^{\bar{v}} (1-G(v))dv =: h(\tilde{v})$ we have that the second-order Taylor approximation of $h(\tilde{v})$ in a neighborhood of $\bar{v}$ is

$$h(\tilde{v}) = \frac{1}{2}g(\bar{v})(\bar{v}-\tilde{v})^2 + o\left((\bar{v}-\tilde{v})^2\right),$$

so that in a second-order approximation $\int_{\tilde{v}}^{\bar{v}} (1-G(v))dv = (1-G(v_R^*))\,s$ yields

$$\frac{1}{2}g(\bar{v})(\bar{v}-\tilde{v})^2 + o\left((\bar{v}-\tilde{v})^2\right) = g(\bar{v})s\left(\sqrt{\frac{2s}{g(\bar{v})}} + o(\sqrt{s})\right)$$

or

$$\bar{v} - \tilde{v} = \sqrt{2s\sqrt{\frac{2s}{g(\bar{v})}}} + o(s\sqrt{s}).$$

Thus,

$$\begin{aligned}
\lim_{s\to 0} \frac{E(v|v\geq v_R^*) - \tilde{v}}{(1-G(\tilde{v}))} &= \frac{1}{g(\tilde{v})} \lim_{s\to 0} \left(1 - \left[\frac{(g(\bar{v})(\bar{v}-v_R^*))^4}{g(\bar{v})(\bar{v}-\tilde{v})g(v_R^*)s} + \frac{(g(\bar{v})(\bar{v}-v_R^*))^2}{g(\bar{v})(\bar{v}-\tilde{v})} \right]^{-1} \right) \\
&= \frac{1}{g(\tilde{v})} \lim_{s\to 0} \left(1 - \left[\frac{o(s^2)}{o(s^{1\frac{3}{4}})} + \frac{o(s)}{o(s^{\frac{3}{4}})} \right]^{-1} \right) = -\infty.
\end{aligned}$$

We next consider that $\delta > 0$ and small and argue that the platform can use the same algorithm. If a consumer t, $T_1 < t \leq T_1 + T_2$, inspects products they want to continue to

search if they find a $u < \tilde{u}$ on the first search, where $\tilde{u}$ is implicitly defined by

$$\tilde{u} = \frac{1 - H_\delta(\tilde{u})}{1 - H_\delta(u_R^*)} E(u|u \geq \tilde{u}) + \frac{H_\delta(\tilde{u}) - H_\delta(u_R^*)}{1 - H_\delta(u_R^*)} \tilde{u} - s.$$

For sufficiently large T_1 the pay-off of buying blindly equals $(1-\delta)E(v|u \geq u_R^*) + \delta E(m)$. Thus, it is clear that for sufficiently large T_1 and for small s we have that consumers $t > T_1$ would like to inspect products, instead of buying blindly, if

$$E(u|u \geq \tilde{u}) - \frac{(1 - G(u_R^*))\, s}{1 - G(\tilde{u})} > (1-\delta)E(v|u \geq u_R^*) + \delta E(m).$$

Above we proved this to be the case for $\delta = 0$ and s in a right-neighborhood of $\hat{s}_R(0)$. As both the LHS and the RHS are continuous in δ this inequality must also hold for δ close enough to 0 and s in a right-neighborhood of $\hat{s}_R(\delta)$. Thus, using the algorithm, the platform learns more than the random cutoff, $\hat{v} > v_R^*$ for these parameters.

Finally, we show that the platform can no longer achieve the cutoff v^* for any combination of δ and s such that $v^* < \bar{v}$, i.e., if $\delta < \hat{\delta}$. Recall that v^* is such that $u^* = (1-\delta)v^* + \delta \underline{m}$ and without blind buying all items are accepted in the long run (as $t \to \infty$). The cutoff v^* is the highest possible cutoff that can be reached when blind buying is allowed because the platform can learn weakly less when blind buying is allowed versus prohibited. Suppose that the platform can reach the cutoff v^* in the long run and consider a $t = T$ so large that the top-ranked item has almost surely $v \geq v^*$. Then consumers $t > T$ expect to get a utility (almost) equal to $E(u|v > v^*) - s$ by inspecting and $E(u|v > v^*)$ by buying blindly. Thus, all consumers $t > T$, i.e., sufficiently far in the queue, prefer to buy blindly and the platform cannot identify items with $v \geq v^*$ for $v^* < \bar{v}$.

Proof of Proposition 9.

(i) First consider that consumers do not know market shares and $\delta < \hat{\delta}$. We show that a platform always has an incentive to "experiment less" than the competitor. An important part of the proof is to consider what "experimenting less" means. By Lemma 2 a platform never prominently ranks items that have been inspected and not bought. Thus, the only question is how platforms rank uninspected items versus the items that have been bought in the past and subsequently never rejected (call these item Bs). Note that as long as a platform does not experiment, there is only one item B on that platform. We say that platform 1 "experiments first at t" if in period t platform 1 ranks the only item B on the top spot with a probability less than one.

We first show that, all else equal, a consumer prefers a ranking where an item B is on a single position with probability one rather than a ranking where item B is on position k with probability $\pi > 1/2$ and another position k' with probability $1 - \pi$. To see that, let us consider the expected value that the consumer receives from searching through the

latter ranking. Once the consumer draws an item with utility u, their value function is

$$V(u) = \max\{u, E_{u'}[V'(u')]\} - s,$$

where V and V' can differ because the consumer searches through the ranking in a particular order. Note that $V(u)$ is convex. Before drawing u, the expected value from searching through the ranking starting with positions k and k' is

$$E_u[V(u)] = \pi E[\max\{u, V_R(u')\}|u \in \triangle \cup \square] + (1-\pi)E[\max\{u, V''(u')\}] - s,$$

where $V''(u') = \max\{u', E[V_R(u'')|u \in \triangle \cup \square]\} - s$. The expression can be read as follows. With probability π, the consumer inspects item B first (since it is on position k) in which case they know that it is either in the "triangle" or "rectangle" because it has been bought before. The consumer accepts B if it is better than continuing search. If they reject B, they know that the rest of the items in the ranking are randomly drawn uninspected items and V_R is defined as the value of searching over these. With probability $1-\pi$, a random item is on position k and the consumer knows that item B is on the next-searched position k' so their continuation value takes that into account. The fact that $V(u)$ is convex and expectation is a linear operator imply that $E_u[V(u)]$ is convex, too. But then the expected value $E_u[V(u)]$ is higher if outcomes are more spread, i.e., at $\pi = 1$ rather than an interior π. Thus, a consumer prefers a ranking that places item B on a single position with probability one rather than on two positions with some interior probabilities. By an analogous argument, a consumer prefers a ranking that places an item B on a single position with probability one rather than on more than two positions with some interior probabilities.

Now we show that if platform 1 experiments with a positive probability, then platform 2 wants to undercut it. Suppose that neither platform experiments at $t < \tau$ and platform 1 experiments with a positive probability at τ: platform 1's ranking places item B on the top spot with a probability $\pi_\tau < 1$ at $t = \tau$. We claim that platform 2 benefits from deviating to a ranking that at τ experiments with probability $\pi'_\tau < \pi_\tau$.

Since consumer $t = 1$ expects both platforms to make equally good recommendations, the consumer is indifferent between them. Consumer $t = 2$ then expects both platforms to have attracted the same number of consumers at $t = 1$. Thus, consumer 2, too, is indifferent between the platforms because they expect the platforms to make equally good recommendations. This is true for all consumers $t < \tau$. Now consider consumer τ. They know that platform 1 experiments and, as we showed above, strictly prefers searching on platform 2 if it experiments less at τ. Thus, platform 2 has an incentive to post a ranking that at τ experiments with a slightly lower probability than π_τ, for two reasons. First, it attracts consumer τ. Second, platform 2 learns (almost) the maximal amount that

it would have learned if it had experimented with any other probability $\hat{\pi}_\tau$: if $\hat{\pi}_\tau > \pi_\tau$, platform 2 expected to learn nothing at τ as consumer τ would have searched on platform 1; if $\hat{\pi}_\tau < \pi_\tau$, platform 2 would have attracted consumer 2, but learned less.

As a result of this deviation, consumer $\tau + 1$ strictly prefers platform 2 (if it expects the platforms to experiment the same amount at $\tau+1$) because they know that platform 2 has attracted one additional consumer (in expectations) and likely makes them a better recommendation. Platform 2, thus, strictly benefits from undercutting platform 1 in terms of experimentation in all τ where platform 1 experiments with a positive probability.

Thus, the only potential equilibrium is where no platform experiments. To see that this constitutes an equilibrium, consider a deviation by platform 1 to an algorithm that experiments first with some consumer $\tau \geq 2$ (as platforms experiment with consumer $\tau = 1$ by definition). By a similar argument as above, any consumer $\tau \geq 2$ strictly prefers a platform that doesn't experiment to platform 1. Thus, this deviation is not profitable and $lim_{t\to\infty} v_1^i(t) \geq v_R^*$.

(ii) Now suppose that consumers know market shares. To simplify the argument, but without loss of generality, assume that there are two platforms and that a consumer who is indifferent between platforms breaks their indifference by using the platform that has served more consumers in the past. We show that a monopoly emerges and the monopolist can learn as much as in the monopoly model.

The algorithm that any platform can use to achieve $lim_{t\to\infty} v_1(t) \geq v^*$ with probability a half is the following:

(i) If at t an equal number of past consumers have used the two platforms or fewer past consumers have used platform i, rank items according to their posterior vs.

(ii) If at t, more past consumers have used platform i, use the same algorithm as in the monopoly case (specified in the proof of Proposition 5, on p. 31 for $\delta < \hat{\delta}$ and in the proof of Proposition 7, on p. 35 for $\delta \geq \hat{\delta}$).

Without loss of generality, suppose that consumer 1, who is indifferent between the platforms, uses platform 1. Now consider consumer 2. If $\delta < \hat{\delta}$, both platforms have a fully random ranking, so consumer 2 uses platform 1 (because of the tie-breaking rule). If $\delta \geq \hat{\delta}$, platform 2 has a fully random ranking, while platform 1 ranks the item bought by consumer 1 on the top spot so consumer 2 strictly prefers to go to platform 1. By a similar argument, for all consumers $t \geq 3$, either platforms 1 and 2 offer fully random rankings (platform 1 when it has restarted its process and the consumer belongs to cohort one and platform 2 because it has not attracted any consumers) and consumer t uses platform 1 (because of the tie-breaking rule) or platform 1 offers a strictly better ranking (containing at least on the top spot an item that has a higher expected inspection value than u_R^*, while platform 2's ranking is fully random) and consumer t strictly prefers to use platform

1. Thus, the initial consumer determines fully which platform gains the entire market (here, platform 1) and, because this platform can use the same algorithm as the monopoly platform, it learns as much as under monopoly. The platform that does not get the first consumer cannot deviate to a better ranking algorithm simply because it can never offer a better than fully random ranking if it does not attract at least a single consumer.

References

Anderson, S. P. and R. Renault (1999): "Pricing, product diversity, and search costs: A Bertrand-Chamberlin-Diamond model," RAND Journal of Economics, 719–35.

——— (2025): "Search Direction: Position Externalities and Position Auction Bias," Review of Economics Studies, forthcoming.

Arbatskaya, M. (2007): "Ordered Search," RAND Journal of Economics, 38, 119–26.

Armstrong, M. (2017): "Ordered consumer search," Journal of the European Economic Association, 15, 989–1024.

Armstrong, M., J. Vickers, and J. Zhou (2009): "Prominence and Consumer Search," RAND Journal of Economics, 40, 209–33.

Banerjee, A. V. (1992): "A simple model of herd behavior," Quarterly Journal of Economics, 107, 797–817.

Bar Isaac, H. and S. Shelegia (2025): "Monetizing Steering," Joural of Political Economy: Microeconomics, forthcoming.

Benabou, R. and R. Gertner (1993): "Search with learning from prices: does increased inflationary uncertainty lead to higher markups?" Review of Economic Studies, 60, 69–93.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," Journal of Political Economy, 100, 992–1026.

Bolton, P. and C. Harris (1999): "Strategic experimentation," Econometrica, 67, 349–74.

Choi, M., A. Y. Dai, and K. Kim (2018): "Consumer search and price competition," Econometrica, 86, 1257–81.

De los Santos, B. and S. Koulayev (2017): "Optimizing click-through in online rankings with endogenous search refinement," Marketing Science, 36, 542–64.

Doval, L. (2018): "Whether or not to open Pandora's box," Journal of Economic Theory, 175, 127–58.

Ely, J. C. (2017): "Beeps," American Economic Review, 107, 31–53.

Ely, J. C. and M. Szydlowski (2020): "Moving the goalposts," Journal of Political Economy, 128, 468–506.

European Parliament (2022a): "Digital Markets Act," https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2022.265.01.0001.01.ENG, [Online; accessed 01.03.24].

——— (2022b): "Digital Services Act," https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065, [Online; accessed 01.03.24].

Glazer, J., I. Kremer, and M. Perry (2021): "The wisdom of the crowd when acquiring information is costly," Management Science, 67, 6443–56.

Goeree, J. K., T. R. Palfrey, and B. W. Rogers (2006): "Social learning with private and common values," Economic Theory, 28, 245–64.

Haan, M. and J. Moraga González (2011): "Advertising for Attention in a Consumer Search Model," Economic Journal, 121, 552–79.

Hagiu, A. and B. Jullien (2011): "Why do intermediaries divert search?" RAND Journal of Economics, 42, 337–62.

Janssen, M., T. Jungbauer, M. Preuss, and C. Williams (2023): "Search Platforms: Big Data and Sponsored Positions," CEPR Discussion Papers 18639.

Janssen, M. C., A. Parakhonyak, and A. Parakhonyak (2017): "Non-reservation price equilibria and consumer search," Journal of Economic Theory, 172, 120–62.

Janssen, M. C. and C. Williams (2024): "Influencing search," RAND Journal of Economics, 55, 442–62.

Ke, T. T., S. Lin, and M. Y. Lu (2022): "Information design of online platforms," HKUST Business School Research Paper.

Keller, G., S. Rady, and M. Cripps (2005): "Strategic experimentation with exponential bandits," Econometrica, 73, 39–68.

Kremer, I., Y. Mansour, and M. Perry (2014): "Implementing the "wisdom of the crowd"," Journal of Political Economy, 122, 988–1012.

Maglaras, C., M. Scarsini, D. Shin, and S. Vaccari (2021): "Social learning from online reviews with product choice," Columbia Business School Research Paper.

Nocke, V. and P. Rey (2024): "Consumer search, steering, and choice overload," Journal of Political Economy, 132, 1684–739.

Orlov, D., A. Skrzypacz, and P. Zryumov (2020): "Persuading the principal to wait," Journal of Political Economy, 128, 2542–78.

Renault, J., E. Solan, and N. Vieille (2017): "Optimal dynamic information provision," Games and Economic Behavior, 104, 329–49.

Rhodes, A. (2011): "Can prominence matter even in an almost frictionless market?" Economic Journal, 121, F297–F308.

Rothschild, M. (1974): "Searching for the lowest price when the distribution of prices is unknown," Journal of Political Economy, 82, 689–711.

Schulz, J. (2016): "The optimal Berry-Esseen constant in the binomial case," PhD Dissertation, University of Trier.

Smith, L. and P. Sørensen (2000): "Pathological outcomes of observational learning," Econometrica, 68, 371–98.

Smolin, A. (2021): "Dynamic evaluation design," American Economic Journal: Microeconomics, 13, 300–31.

Teh, T.-H. and J. Wright (2022): "Intermediation and steering: Competition in prices and commissions," American Economic Journal: Microeconomics, 14, 281–321.

Ursu, R. M. (2018): "The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions," Marketing Science, 37, 530–52.

Weitzman, M. L. (1979): "Optimal Search for the Best Alternative," Econometrica, 641–54.

Wilson, C. (2010): "Ordered Search and Equilibrium Obfuscation," International Journal of Industrial Organization, 28, 496–506.

Wolinsky, A. (1986): "True monopolistic competition as a result of imperfect information," Quarterly Journal of Economics, 101, 493–511.

Zhou, J. (2011): "Ordered Search in Differentiated Markets," International Journal of Industrial Organization, 29, 253–62.