Search Platforms: Big Data and Sponsored Positions[*]

Maarten Janssen[†] Thomas Jungbauer[‡]

Marcel Preuss[§] Cole Williams[¶]

July 2, 2024

We study a search platform ranking firms' products across sponsored and organic positions, accounting for the incentives of both firms and consumers. To characterize an optimal ranking when the number of firms is large, we formulate a Mixing Principle for Consumer Search, adapting tools from the social learning literature. The platform assigns the products it deems best to sponsored positions and obfuscates the content of organic positions subject to consumers' participation constraints. Obfuscation serves to maximize the platform's revenue from both sponsored position auctions and commission fees. Our results allow us to analyze the welfare effects of sponsored positions.

In 2024, the worldwide market for digital advertising is projected to reach \$740bn. Its biggest component—accounting for about 40% or \$307bn—is *search advertising*.[1] Search platforms (such as Google, Tripadvisor or Yelp) assist consumers looking for a product or service.[2] The consumer submits a keyword, the platform provides a list of results, and the consumer inspects these results in whatever order they prefer. In producing this list, the platform may draw on information about the consumer (demographics, past searches, order histories, etc.) as well as on other consumers' behavior. Search advertising refers to the

[*]We are grateful to Nageeb Ali, Simon Anderson, Heski Bar-Isaac, Marco Battaglini, Emiliano Calvano, Federico Echenique, William Fuchs, Kinshuk Jerath, Justin Johnson, Bruno Jullien, Louis Kaplow, Tony Ke, H. Tai Lam, Massimo Motta, Volker Nocke, Antonio Penta, Patrick Rey, Sandro Shelegia, Vasiliki Skreta, Steve Tadelis, Nikhil Vellodi, Rakesh Vohra, Michael Waldman, Mark Whitmeyer, Leeat Yariv, Yang Yu and seminar participants at the Workshop on Digital Platforms (Barcelona, 2022), The Economics of Platforms Workshop (Capri, 2023), SAM Workshop (Bristol, 2023), IIOC (Washington D.C., 2023), Consumer Search and Switching Cost Workshop (Rotterdam, 2023), SAET (Paris, 2023), Strategy and Economics of Digital Markets Conference (Ithaca, 2023), EARIE (Rome, 2023), Tinbergen Institute Digital Platform Conference (Amsterdam, 2023), AMES (Beijing, 2023), Norwegian Business School (Oslo, 2024) and Hal White Antitrust Conference (2024), for their helpful comments and valuable discussions. All remaining errors are our own.

[1]See https://www.statista.com/outlook/dmo/digital-advertising/worldwide. The other major digital advertising categories are audio, banner, classifieds, influencer, in-app, social media and video, none of which generates an amount of spending comparable to search advertising.

[2]We refer to search platforms and platforms interchangeably throughout.

practice of platforms to auction specific positions in the list of search results to advertisers (firms) for their products or services.

The power of these paid search results, or *sponsored positions*, stems from (a) the way they steer consumer search activity and (b) the platform leveraging its information about consumers. Online search pioneers and Google founders Sergey Brin and Larry Page discussed the commercial value of sponsored positions (slots) well before the advent of search advertising, warning of their implications:

> "We expect that advertising funded search engines will be inherently biased towards the advertisers and away from the need of the consumers."[3]

In this paper, we study how search platforms leverage their information to shape online markets by creating a ranking to maximize their revenue from both selling sponsored slots and earning commission fees. We show that when ranking a large number of search results, a platform has an incentive to assign the firms it deems most relevant for the consumer to sponsored positions, while *obfuscating* the remaining results across organic positions (Theorem 1). Thus, the impact of sponsored positions and their welfare effects are more complex than this quote suggests.[4]

The extent to which the platform is able to obfuscate organic positions depends on consumers' outside option. To ensure consumers who do not buy from sponsored positions inspect organic ones rather than leave the platform, the platform may find it optimal to introduce *premium positions* across which it obfuscates more relevant products (Theorem 2).[5]

Obfuscation plays a dual role in maximizing auction revenue from sponsored slots and commission fees. On one hand, by obfuscating organic search results, the platform increases the probability a consumer buys from a sponsored position, raising advertisers' valuation of obtaining such a position. On the other hand, if the consumer does not buy from the sponsored slot, obfuscation causes them to inspect a larger number of organic positions, increasing the platform's expected earnings from commission fees.

---

[3] This quote first appeared in 1998 in the appendix of a paper later in the same year published as Brin and Page (1998) introducing the Google search engine to the wider scientific community.

[4] This result is supportive of the FTC's recent announcement to sue Amazon for products by deliberately worsening search quality in order to increase their profits (see https://www.ftc.gov/news-events/news/press-releases/2023/09/ftc-sues-amazon-illegally-maintaining-monopoly-power).

[5] The relevance of premium positions is supported by empirical findings that consumers overwhelmingly search more prominent positions. In the case of Amazon, for example, more than 80% of clicks happen on the first page (Yu, 2024).

Even though our results indicate that a platform's ranking of search results does not reveal its entire information, we find the platform has an incentive to invest in learning about consumers, as better information increases firm's bids for sponsored positions and the platform's ability to assign firms relevant for the consumer to premium positions.[6] Abandoning sponsored positions may actually decrease consumer welfare.

Building on the seminal consumer search papers of Wolinsky (1986) and Anderson and Renault (1999), we study a search platform providing a ranking of a finite but large number of keyword-relevant firms in response to a representative consumer query. Leveraging its access to rich consumer data, the platform has some understanding of the consumer's preferences, represented by a *match score* for each firm that is informative of the consumer's ultimate *match value*. The platform chooses a ranking algorithm assigning firms across one sponsored and many organic positions based on their bids and match scores. In our baseline model—which we extend in several directions—ex ante symmetric firms submit bids for the sponsored slot and charge a uniform price. The consumer then sequentially inspects the platform's search results in their preferred order, each time incurring a search/inspection cost, until they acquire a product/service or abort search altogether.

As the ranking of search results is strategically chosen by the platform and match scores and match values are correlated, inspecting one slot allows consumers to learn about the value of products not yet inspected. It is well-known that learning in consumer search complicates the characterization of the optimal search behavior; see, e.g., Rothschild (1974), Janssen et al. (2017) and Garcia and Shelegia (2018). By means of examples we show that due to learning, with a small number of search results it may neither be optimal to obfuscate organic slots nor to assign the sponsored slot to the firm with the best match score.

We overcome these issues by adapting the so-called *mixing property* of stochastic processes described within the context of social learning (Mossel et al., 2020), introducing a *Mixing Principle for Consumer Search*. It states that—due to the independence of match values across firms—the match value of the firm in the sponsored slot can be strongly correlated with the match values of at most a few firms in organic positions (Lemma 1). As such, when the number of search results increases, additional information about the continuation value of search conveyed by a realized match value vanishes. We believe that our Mixing Principle

[6] The extent to which user data collection can be excessive plays an important role in the current antitrust case in front of the United States District Court for the District of Columbia vs. Google LLC.

is fruitful for other search and inspection environments in which learning is important.

In order to understand why full obfuscation—i.e., minimizing information provision—of organic slots is optimal with a larger number of firms, consider the consumer's problem. Under full obfuscation, they essentially pick a random firm once they search beyond the sponsored slot.[7] It follows that full obfuscation minimizes the consumer's utility from inspecting organic positions, thereby minimizing the likelihood of the consumer to search beyond the sponsored slot. As a result, advertisers' willingness to bid for the sponsored slot is maximized, even though obfuscation minimizes the probability that a consumer returns to the sponsored slot (return demand). The Mixing Principle, imposing bounds on return demand, is essential in establishing this obfuscation result.

The logic of why the platform designs the auction such that the firm with the highest match score wins the the sponsored slot is as follows. Note that in our baseline model all firms submit identical bids. As a consequence, the sponsored position contains informational value for both consumers and firms when the platform assigns the highest match score firm to the sponsored slot in case of a tie. Therefore, consumers examine the sponsored slot first. Firms, on their part, learn about their favorable match score through winning the auction, raising their willingness to bid. Thus, it is not benevolence that drives platforms to take match scores into account when allocating the sponsored slot (instead of selling the slot to the highest bidder). In fact, doing so increases firms' bids to acquire the sponsored position.

We also show that the better a search platform is able to predict consumers' match values (in the sense of Lehmann (1988)) the higher platform profits and consumer surplus (Proposition 1). Consumers are more likely to buy from the sponsored slot if a firm with a more predictive match score wins the sponsored position. As a result, firms are willing to up their bids for the sponsored position, while consumers expect both a better match and lower search costs.

Obfuscating organic positions does not only maximize the advertising firms' willingness to bid for the sponsored position, but also the number of organic slots a consumer is expected to inspect. As such, the probability of a sale intermediated by the platform increases, independent of both whether the platform offers sponsored positions or not and the consumer's outside option. The consumer's alternatives do however affect the set of firms over which the platform randomizes. The better the outside option, the higher the match

[7] This remains true with premium positions when the consumer has a viable outside option.

scores of the firms assigned to premium positions. When prices are uniform, maximizing sales probability maximizes sales commissions. Therefore, full obfuscation optimizes both the auction revenue from the sponsored slot and sales commissions.

Our main obfuscation result persists when adding other real-world features of search advertising. If firms also hold consumer-relevant information, the platform can extract this information through the auction for the sponsored slot. Thus, our main intuition about an informative sponsored slot and obfuscation of organic slots persists (Proposition 2). Alternatively, when firms quote different prices, the platform continues to obfuscate organic positions and sells the sponsored slot to the firm with the highest match value (Proposition 3). However, to earn high commission fees, the platform now will put firms with higher match score and *the highest prices* in premium positions (Corollary 1). Thus, heterogeneous prices generate another reason to obfuscate over premium positions.

Finally, we address welfare effects of introducing sponsored positions by investigating how the platform ranks products in the absence of sponsored positions. Note that the platform's asymptotically optimal ranking creates several inefficiencies, harming consumers. First, the expected match value of consumers who buy at the sponsored position decreases (as, with obfuscation consumers are more willing to stop searching than without). Second, consumers who continue to search beyond the first position expect to spend more time inspecting other products to achieve any given match value. Finally, with heterogeneous prices, consumers tend to pay higher prices in organic slots as these firms generate higher commissions from a sale for the platform. If in the absence of sponsored positions, the platform maximizes only sales commissions revenue, then (perhaps surprisingly) consumers are even worse off (Corollary 2). Independent of whether the platform offers sponsored positions, it fully obfuscates organic slots. With a sponsored position, however, the platform allocates the top spot to a firm with the best match score, providing the consumer with more information.

Starting with Athey and Ellison (2011), Chen and He (2011) and Eliaz and Spiegler (2011), a growing literature on position auctions explicitly takes into account that the value of a position depends on consumers' search patterns.[8] In contrast to these papers, we focus on a platform that has information about consumer preferences, allowing us to address the important policy question of how online search platforms leverage their information about

[8]Other important papers on information gatekeepers include Baye and Morgan (2001), Armstrong and Zhou (2011) and De Corniere and Taylor (2019).

consumers to steer consumer search.

More recent papers on search platforms have addressed questions that are different from ours such as which rankings maximize consumer and industry surplus (Anderson and Renault, 2021), and the effects of brand advertising (Motta and Penta, 2022), different sales mechanisms of sponsored positions (Bar-Isaac and Shelegia, 2022) or a social influencer recommending products Janssen and Williams (2022).[9] None of these papers consider, however, the issues we address related to the effects of sponsored positions.[10]

In the context of a multi-product firm, Nocke and Rey (2023) find that "garbling" of information may be optimal for the seller as it induces a buyer to inspect a larger number of items before terminating search. This is related to the second role of obfuscation in our paper. Similar to Chen and He (2011) and Anderson and Renault (2021), however, in their paper a firm-consumer pair either constitutes a match or not, and the value of this match is constant across firms. As a consequence, consumers do not learn about firms not yet inspected. Also, neither of these papers studies the interactions of sponsored and organic positions.

The effects of sponsored slots and different rankings of organic slots on choices have also been studied empirically. Ghose et al. (2014), Ursu (2018), and Donnelly et al. (2022), among others, show that personalized rankings affect consumer choices and induce positive welfare effects. Using a field experiment of mobile search platforms in 13 Asian cities where consumers search for restaurants, Sahni and Nair (2020) find that users significantly more often call at restaurants if these have acquired a sponsored position. In line with our main result, this suggests sponsored slots provide valuable information to consumers.

The remainder of this paper is organized as follows. Section 1 introduces the main model. Section 2 presents our main results, which Section 3 generalizes to scenarios where (i) both the platform and the firms hold consumer-relevant information and to (ii) firms with heterogeneous prices. In Section 4 we discuss the welfare effects of sponsored positions and how they depend on the platform's objectives. Finally, Section 5 concludes with a discussion.

[9] Armstrong and Zhou (2022) consider how information provision affects consumers in the absence of consumer search.

[10] Long et al. (2022) and Ke et al. (2022) also analyze how a platform uses its information about consumer preferences to allocate firms to sponsored and organic positions, but their modeling is very different and consumer learning plays a very limited role.

1 The Model

The market comprises a platform, n firms selling horizontally differentiated products, and a representative consumer. The consumer demands one unit of the product and has an unobserved *match value* v_i with firm i. Match values are independently and identically distributed across firms according to a continuous distribution G with bounded support $[\underline{v}, \bar{v}] \subset \mathbb{R}$.[11] The platform's information regarding the consumer's match value with firm i is summarized by a *match score* θ_i, following a continuous distribution $F(\theta_i|v_i)$ with compact support $[\underline{\theta}, \bar{\theta}] \subset \mathbb{R}$. Higher scores indicate "good news" in the sense that $F(\cdot|v_i')$ has likelihood ratio dominance over $F(\cdot|v_i)$ if $v_i' \geq v_i$, i.e., a higher match score is suggestive of a higher match value. Let z denote an independent non-atomic random variable with support Z that the platform can use as a randomization device to play a mixed strategy. We then denote the probability measure on $\Omega = [\underline{\theta}, \bar{\theta}]^n \times [\underline{v}, \bar{v}]^n \times Z$ by μ and denote probabilities and expectations w.r.t. μ by $\mathbb{P}[\cdot]$ and $\mathbb{E}[\cdot]$, respectively.

The platform displays a *ranking* $x \in X$ of the firms to the consumer, where X denotes the set of firm permutations.[12] The platform forms the ranking by running an auction in which firms submit bids, the winner of the auction is placed at the top in the "sponsored" position (i.e. $x(i) = 1$ implies i is sponsored), and all other firms are allocated to the remaining "organic" positions. Denote firm i's bid by $b_i \geq 0$. Given the bids and scores, the platform's algorithm determines which firm wins the sponsored position and how the remaining firms are allocated to the organic positions. The platform's *algorithm* is a function mapping bids, scores, and realizations of the randomization device to rankings, $a : \mathbb{R}_+^n \times [\underline{\theta}, \bar{\theta}]^n \times Z \rightarrow X$. Let $\mathcal{A}_n$ denote the set of (measurable) algorithms in the game with n firms. Denoting the vector of bids by $\mathbf{b} = (b_1, b_2, \ldots, b_n)$, the platform also specifies a *payment rule* $\rho : \mathbb{R}_+^n \rightarrow \mathbb{R}_+^n$ with $\rho(\mathbf{b}) = (\rho^1(\mathbf{b}), \ldots, \rho^n(\mathbf{b}))$ and whereby $\rho^i(\mathbf{b}) \leq b_i$ is the amount firm i pays the platform conditional on it winning the sponsored position. Let $\mathcal{P}_n$ denote the space of (measurable) payment rules in the game with n firms.

Note that the algorithm and payment rule jointly define the structure of the sponsored search auction. For example, an algorithm that always places a firm with the highest bid in the sponsored position and a payment rule satisfying $\rho^i(\mathbf{b}) = b_i$ correspond to a standard

[11] The results can be extended to the case with unbounded match values.

[12] That is, X is the set of bijections from $\{1, \ldots, n\}$ to itself.

first-price auction. More generally, the platform can use the information contained in the scores to determine the winner of the auction.

Consumers are initially uninformed of their match values with firms and can only uncover them through costly sequential search. At each point along the search path, the consumer can select any position in the ranking and incur the inspection cost $s > 0$ to learn the match value with the firm located in that position, buy the good from a firm whose match value the consumer has already inspected, or exit the market and take an outside option $\eta \in \mathbb{R}$ at the same price p.[13] Defining $r(\phi)$ to be the reservation value for a firm with match value above ϕ (i.e. $\mathbb{E}\left[\max\{v_i - r(\phi), 0\} | \theta_i \geq \phi\right] = s$), we focus on the interesting case in which $\eta < r(\bar{\theta})$—otherwise, no algorithm designed by the platform attracts consumers.

When searching they only observe their match values with different firms but not the platform's match score. In principle, the consumer search problem is reminiscent of Pandora's box problem as studied in Weitzman (1979). However, knowledge that the platform utilizes in a ranking algorithm creates interdependence between match values so that inspecting the goods of one firm provides the consumer with information about other firms. Consumers have perfect recall when searching.

The timing of interaction is as follows. First, the platform commits to an algorithm and payment rule observed by the firms and consumer. Second, firms privately submit their bids to the platform. Third, Nature determines the match scores and values. Fourth, the platform receives the firms' scores and bids, and the algorithm determines the position each firm takes in the list. The consumer receives the list and then proceeds with their search. At each point along the search path, the consumer's information consists of the the algorithm, payment rule, as well as the realized match values at all positions previously inspected.

Consumers receive the match value minus the price of a good they purchase net the search costs. A firm's profit equals the revenue minus production cost and any fee paid for the sponsored position. Unless explicitly stated otherwise (as in Sections 3.2 and 4) we assume that all firms charge price p and normalize their production costs to 0. The platform's expected profit corresponds to the sum of the expected revenue from the sponsored search auction and the expected revenue from sales commission fees, i.e., a fee firms pay in the case of a transaction that is a fixed percentage $q \geq 0$ of the transaction sum. We focus on

[13] Assuming the consumer's outside option η to come at price p simplifies notation but is innocuous for our results. In essence, at every point in time, the consumer buys from the platform or continues to search only if their expected payoff from doing so exceeds $\eta - p$.

symmetric Perfect Bayesian Equilibria.

We end this section with a few comments on the model. First, we assume that the platform commits to its ranking algorithm. We see this as a reasonable approximation of the real world situation where platforms submit a ranking of alternatives within a split second after the consumer has typed its key words. Consumers typically use the same platform repeatedly learning about how the resulting rankings satisfy their needs. Platforms may, of course, work on different algorithms to improve their functioning, but will implement new algorithms only once in a while. Without commitment, other outcomes than the ones we focus on in this paper may be supported. For example, if firms and consumers believe that the platform's ranking, including the sponsored position, is completely random, then the platform may not be able to do better than indeed randomly allocating firms to positions.

Second, the questions we address are akin to the ones studied in the literature on information design, as initiated by Kamenica and Gentzkow (2011), in the sense that the platform chooses which information to release to firms and consumers. However, unlike most of that literature, the platform's choice set is limited in that it can only choose a ranking of alternatives and gives the same information to both consumers and firms.

Third, the platform may decide not to rank all firms. However, it is clear that as the platform also generates revenue from commissions, it will never want to do so. In the worst case, the platform can always rank items in such a way that consumers will decide themselves not to search among a subset of the items.

Fourth, the model treats firms' prices as exogenous. In particular, prices do not depend on whether or not a firm is recommended. We think that this is realistic in many cases in which the revenue a firm makes is only to a limited extent dependent on the sales via the search platform. Implicitly, we also assume that all firms charge identical prices, but that turns out to be inessential as we will explain in 3.2.

2 Obfuscation to Maximize Revenue

In this section we state and explain our main result: as the number of firms grows large, it is optimal for the platform to obfuscate the organic slots to the maximal extent possible—subject to consumer participation constraints—and to allocate the sponsored position to the firm with the highest match score. To focus on how consumers learn along their search path and how our Mixing Principle deals with that, we initially simplify and consider in subsection 2.1 that the consumers' outside option is not binding so that even if the platform fully obfuscates the organic slots consumers continue to use the platform if the sponsored position did not result in a good enough match. As such, revenues from sales commissions are automatically generated when consumers find a good match eventually. By means of two examples, we show in this part of the analysis that for small n the optimal algorithm may *neither* obfuscate organic slots *nor* to put the firm with the highest match score in the sponsored position.

Then in subsection 2.2, we consider a binding outside option, i.e., if the platform uniformly obfuscates all organic positions, consumers will not choose to search any organic slots. This outside option may reflect competition among platforms, alternative sales channels and the availability of substitute products. In order to generate maximal revenue from both the sponsored slot and commission fees, the platform will then determine a subset of firms with the best match scores and randomly assign these firms to limited "premium positions" such that the consumer finds it optimal to continue searching these positions until they find a product with a good enough match.

Even though a platform does not have the incentive to provide a full ranking according to match scores, it still has an incentive to invest in acquiring information on consumers' match values. To further emphasize this point, we show in subsection 2.3 that the platform generates more revenue if its match scores are better predictors of consumers' match values.

Our analysis focuses on an algorithm that yields approximately maximal profits when the number of firms to be ranked gets large.[14] To this end, it is important to note that after choosing the algorithm and payment rule, a proper subgame is played by the firms, the consumer, and nature. Let $\mathcal{X}_n \subset \mathcal{A}_n \times \mathcal{P}$ denote the nonempty subset of algorithms and payment rules for which an equilibrium of the subgame exists. We focus on the "platform's preferred equilibria" of the subgame by defining $\Pi : \bigcup_{n=1}^{\infty} \mathcal{X}_n \to \mathbb{R}$ to be the supremum of the platform's expected profit taken over the equilibria of the subgame. The results can be extended to apply to any arbitrary selection of equilibria through the use of more complicated algorithms. Thus, in this paper we define asymptotic optimality as follows.

Definition 1. *A sequence $\{a_n, \rho_n\}_{n\in\mathbb{N}}$ is **asymptotically optimal** if for every sequence $\{a'_n, \rho'_n\}_{n\in\mathbb{N}}$ and $\epsilon > 0$ there exists an n^* such that $n \geq n^*$ implies $\Pi(a_n, \rho_n) + \epsilon > \Pi(a'_n, \rho'_n)$.*

[14] For finite n there is not one algorithm that is approximately optimal which follows from our examples below.

2.1 Non-Binding Outside Option

To specify a non-binding outside option, we define the function $d_i : \mathbb{R} \times \Delta\Omega \to \mathbb{R}$ by

$$d_i(v, \lambda) = \int_{\Omega} \max\{v_i - v, 0\} d\lambda(\omega) - s, \tag{1}$$

which reflects the option value of inspecting firm i and paying the search cost s when the consumer can alternatively select a product with value v and the distribution over the state is λ. Observe that $d_i(v, \lambda)$ is strictly decreasing in v whenever $\max \operatorname{supp} \lambda > v$ and takes the value of s otherwise. The *ex ante reservation price* is the unique value $\bar{r}$ satisfying $d_i(\bar{r}, \mu) = 0$. Given only the prior, the consumer is indifferent between taking a good with value $\bar{r}$ and first inspecting firm i and then taking the larger value between the two. We will say that a consumer has a non-binding outside option if $\eta < \bar{r}$.

Next we define a class of algorithms, where the match value found at an organic position does not convey any information about the match value at other organic positions.

Definition 2. *An algorithm $a \in \mathcal{A}_n$ exhibits* ***uniform obfuscation*** *if the firms that lose the auction are assigned to each of the organic positions with uniform probability.*

Using this definition, we can now state our main result when the outside option is non-binding.

Theorem 1. *If $\eta < \bar{r}$, then there is a sequence of uniformly obfuscating algorithms that is asymptotically optimal. Along this sequence, the firm with the highest match score wins the sponsored position.*

It is important to note that uniform obfuscation serves the dual role of maximizing revenue from the sponsored slot (by increasing the probability consumers buy from that slot and thereby increasing the bids of firms) as well as the revenue from sales commissions (by increasing the probability the consumer buys from an organic slot, given it does not buy from the sponsored slot). This second, but not the first, role of obfuscation extends Nocke and Rey (2023) who show that a multi-product firm may want to use obfuscation of product orderings to maximize profits. Note, however, that the settings are very different in that, in their setting, consumer learning is restricted in the sense that as all products have the same value, and thus, a consumer never continues searching if a match occurs.

In our setting, consumer learning is important because a ranking algorithm introduces interdependence between the consumer's conditional match values across firms: consumers may use observed match values to make inferences about the match values with firms they have not yet inspected. The ability to influence the consumer's learning over the course of search introduces strategic tension in the platform's objective of designing the algorithm. On one hand, the sponsored firm's initial demand can be increased by providing a less informative ranking of the organic positions because this reduces the consumer's expected payoff from continuing search beyond the sponsored firm. On the other hand, supplying some information in the organic ranking can bolster the sponsored firm's return demand as observing a low match value at an organic firm makes the consumer pessimistic about the remaining organic firms, making returning back to the sponsored firm more attractive.

In addition, for certain realizations of match scores, the platform may prefer not to put the firm with the highest match score in the sponsored position as the consumer may become more optimistic about finding an even better match value (and thus continues to search) if the platform puts the highest match score first. Hiding information from the consumer may then increase the probability that the consumer immediately buys from the sponsored slot.

Theorem 1 provides a clear picture of two properties of optimal algorithms by focusing on the many real-world applications where the number of potentially relevant firms for a search query is large. This allows us to adapt the the property of *mixing* (Mossel, Mueller-Frank, Sly, and Tamuz, 2020, Lemma 1) to establish our main results. In order to develop the intuition for why the result holds for large n, we start by presenting two examples showing that, due to learning effects, our main result may fail to hold with a small number of firms. HERE DISCUSSION (and maybe footnote) OF $q = 0$ but can also be small in examples The first example shows that uniform obfuscation may fail to be optimal, while the second example shows that the platform may not want to put the firm with the best match score in the sponsored position. The second example also illustrates that, with a small number of firms, if the platform always puts the firm with the highest match score in the sponsored position, consumers may have non-monotonic reservation values: they continue to search for intermediate match values, but stop searching for high or low match values.

Example 1 Suppose there are three firms. The consumer's match value is either low ℓ, medium m, or high h and a product is only worth purchasing if it provides at least a medium value. A firm's match score is L when the value is low and H otherwise, i.e., the platform can distinguish firms with low match values from other firms, but cannot distinguish firms with medium and high match values.[15] Suppose the platform employs the following algorithm. The firm with the highest bid is placed in the sponsored position, ties are broken in favor of the firm with the highest match score and further ties are broken with equal probability. For the two non-sponsored firms, if only one of them has a high score it is placed in the second position with probability $\alpha \geq \frac{1}{2}$, otherwise they are arranged in the organic positions with equal probability. Uniform obfuscation is a special case where $\alpha = \frac{1}{2}$.

The consumer's optimal search proceeds in the following manner. If the sponsored firm's value is high h, the consumer buys it immediately since there is no advantage from continuing. If instead the sponsored firm's value is low ℓ, then given the algorithm, the consumer learns that all remaining firms must likewise have low match values and so the consumer exits the market. If, however, the consumer observes m in the sponsored slot, then it might still be prudent to continue searching as some remaining firm might deliver a higher match value.

In Online Appendix S.1, we show that there are values such that (*i*) if $\alpha = \frac{1}{2}$, the consumer continues searching when observing m in the first and second search, but halts otherwise, and that (*ii*) for some values of $\alpha > \frac{1}{2}$ the consumer inspects the second slot if the sponsored position provides a medium match value, but does not find it optimal to search further. Note that by providing some information in the organic slots, the non-uniformly obfuscating algorithm makes inspecting the second firm more desirable, but increases a sponsored firm's return demand as consumers do not inspect the third firm. In addition, it also affects the expected profits of a non-sponsored slot. Online Appendix S.1 shows that firms are willing to bid more for the sponsored slot with an algorithm that is not uniformly obfuscating, i.e., $\alpha > \frac{1}{2}$, because the effect on return demand dominates the other effects.

Example 2 As in the example above, suppose that there are three distinct match values l, m, h and that the platform cannot distinguish medium and high match values, but perfectly recognizes a low value match. Contrary to example 1, every product is, in principle, worth buying (that is, better than any outside option). Also, assume there are four firms.

Consider the following two algorithms. The first one always assigns the firm with the highest match score to the sponsored position and uniformly obfuscates the organic positions. The second algorithm does almost always the same except in the event that two match scores are

[15] This example departs from the assumptions of our model in that the distribution of match values conditional on the match scores do not share the same support. This is insignificant to the particular example since we could modify the distributions to $\mathbb{P}(\{\ell\}|L) = \mathbb{P}(\{m, h\}|H) = 1 - \varepsilon$ so that the conclusion continues to hold for $\varepsilon > 0$ sufficiently small.

L and two are H. In that case the platform puts a firm with the lowest match score in the sponsored position and uniformly obfuscates the organic positions.

In Online Appendix S.2 we show that there are parameters under which the second algorithm gives the winning firm a higher probability of selling. In particular, with the first algorithm the consumer buys immediately from the sponsored position whenever it contains an h or l value, but they continue to search if they see an m value. On the other hand, with the second algorithm the consumer buys immediately from the sponsored position whatever its match value. Since the second algorithm ensures that the consumer always buys from the sponsored slot and never buys from an organic position, it is clear that the platform gets its highest possible profit as a result of firms bidding maximally for the sponsored slot.

The main idea exploited in Online Appendix S.2 is that under the first algorithm when consumers find an l in the sponsored slot, they know that all remaining slots must contain l and therefore they buy immediately. The second algorithm exploits this pessimism: even though the consumer knows that after observing an l in the sponsored slot the realized match scores may now be either $\{L, L, L, L\}$ or $\{L, L, H, H\}$, the consumer may still buy immediately if the ex ante probability of L is sufficiently high.[16] The effect of this is that when the consumer observes an m on the first search, under the second algorithm they are more pessimistic about finding an h when continuing to search than under the first algorithm. Thus, they may stop searching under the second algorithm, but not under the first.

Proof Outline We are now ready to convey the main elements of the proof of Theorem 1. A key idea in the argument is that when there are many firms, then for *any* possible algorithm, when the consumer finds that their match value with the sponsored firm lies below $\bar{r}$, they almost certainly have a better option for how to proceed with their search than to buy the sponsored firm's product. The tool we use to formalize this idea is the fact that independently and identically distributed (IID) random variables have the property of *mixing* (Mossel, Mueller-Frank, Sly, and Tamuz, 2020, Lemma 1). Intuitively, mixing means that any event defined on the same probability space as a sequence of IID random variables can only be strongly related to a finite number of them.

Lemma 1 (Mixing Principle for Consumer Search)**.** *Consider a collection of events* $\{E_n\}_{n\in\mathbb{N}}$ *in* Ω *such that* $\mathbb{P}(E_n) > \alpha$ *for some* $\alpha > 0$*. For every* $v \in \mathbb{R}$ *and* $\delta > 0$*, there exists an* $n^* \in \mathbb{N}$

[16] To some extent, this is reminiscent of Bayesian Persuasion as the platform pools good and bad events subject to the constraint that the consumer stops searching immediately.

such that, if $n \geq n^*$ then $|d_i(v, \mu) - d_i(v, \mu(\cdot|E_n))| < \delta$ for some firm $i \leq n$.

The lemma implies that if the number of firms is large, then for a given event, there always exists a firm whose expected option value conditional on the search histories induced by the event is arbitrarily close to its unconditional option value. The Mixing Principle for Consumer Search follows from the fact that the difference between the probability of an event H conditional on another event E and the unconditional probability of H cannot be too large if the probability of E is bounded away from zero. To understand the logic underlying this claim, consider a game in which a player flips 1000 coins and wins if and only if at least $x \in \{500, ..., 999, 1000\}$ coins turn up heads. Let $E(x)$ be the event that the player wins, and let $\mathbb{P}(H)$ denote the probability that a randomly chosen coin has turned up heads. Then, for $\mathbb{P}(H|E(x)) - \mathbb{P}(H)$ to become large, we must choose a high value of x. But selecting a higher x makes $E(x)$ very unlikely, demonstrating a fundamental trade-off between the likelihood of an event and the effect it can have on the prior. Consequently, an event that occurs with a positive probability bounded away from zero cannot be arbitrarily informative and, thus, cannot shift the conditional probability too much away from the prior.

Using the Mixing Principle for Consumer Search, we derive two important implications. First, we develop an upper bound on platform profits attainable from any algorithm (Lemma B.1). Second, we show how uniform obfuscation, if the firm with the highest match score wins the sponsored position, achieves this upper bound (Lemma B.2). The examples above illustrate that the platform can leverage a consumer's learning to its own advantage when n is small. The Mixing Principle for Consumer Search implies this is not true when n is large.

For the intuition of the first claim, we argue that the probability the consumer buys from the sponsored firm when it offers a match value less than $\bar{r}$ vanishes as the number n of firms grows large. To see why this is true, take any $\epsilon > 0$ and let E_n be the event in which the consumer buys from the sponsored firm and it delivers a match value that is less than $\bar{r} - \epsilon$ when there are n firms. If the probability of this event fails to vanish as the number of firms grows large, then by the Mixing Principle there will eventually be some position with an arbitrarily high probability of containing a firm $i \leq n$ for whom $d_i(\bar{r} - \epsilon, \mu(\cdot|E_n)) > 0$. But this implies a contradiction as there will eventually be a position with an option value that is, in expectation, positive over the event E_n. Thus, the limiting demand facing a sponsored firm with match score θ_i cannot be larger than the probability that its match value exceeds $\bar{r}$, which is equal to $1 - G(\bar{r}|\theta_i)$, which is increasing in θ_i. Consequently, because the platform's profit cannot exceed the amount accrued by the sponsored firm and the match scores are bounded by $\bar{\theta}$, an upper bound on platform profit is $p(1 - G(\bar{r}|\bar{\theta}))$.

For the second claim, we next argue that when n grows large, an algorithm for which the firm with the highest match score wins the sponsored position and uniformly obfuscates the organic slots gives the firm in the sponsored slot a demand that is arbitrary close to this upper bound. The algorithm has two features that are important. First, by allocating the sponsored slot to the firm with the highest match score, the platform makes it attractive for consumers to start their search at the sponsored slot. When n grows large, there is almost surely a firm that has a match score that is arbitrarily close to $\bar{\theta}$. Second, even for large n, the platform could in principle choose a ranking that allows consumers to learn, but by uniformly obfuscating the organic slots, it effectuates that consumers do not learn anything from observing their match value at the sponsored slot and they expect a randomly selected firm to have a match value close to $\bar{r}$. This makes it least attractive for consumers to continue searching beyond the sponsored slot and from the above we know that the return demand is arbitrarily small for large n. Thus, it is *not* the case that when n grows large consumers cannot learn from the platform's ranking, but the platform cannot gain from rankings that do allow consumers to infer information about match values at firms that are not inspected yet. Together, the above two features ensure that a firm's demand in the sponsored slot approaches $1 - G(\bar{r}|\bar{\theta})$.

The last step of the proof argues that for large n there exists an equilibrium where firms find it optimal to bid an amount in the bidding stage that is close to $p(1 - q)(1 - G(\bar{r}|\bar{\theta}))$ realizing an expected profit for the platform close to the maximal attainable profits. This step has two parts. We intuitively discuss here the argument where firms do not have private information on which to condition their bid so that in a symmetric equilibrium the platform receives identical bids from all firms. Section 3.1 shows that the result actually holds for an arbitrary distribution of private information about the consumer between firms and the platform. So, let $\hat{a}_n$ be the algorithm which: (1) assigns the sponsored position to the firm with the highest bid, breaking ties in favor of a firm with the highest score and remaining ties with uniform probability; (2) assign those firms that do not win the auction to organic positions with uniform probability. Specify the payment rule $\hat{\rho}_n^i(\mathbf{b})$ to be equivalent to a second-price auction, equal to the highest bid among all firms $j \leq n$ excluding i.

Let us provide the necessary conditions for there to be a symmetric equilibrium in which all firms submit the same bid $\beta_n \in \mathbb{R}$. Given the above algorithm, payment rule, and strategy profile, let $\pi(m,n)$ be a firm's expected profit (excluding the bid paid to the platform) given that it is placed in the mth position in the game with n firms. A firm's expected profit is then

$$\frac{1}{n}(\pi(1,n)-\beta_n)+\sum_{m=2}^{n}\frac{1}{n}\pi(m,n).$$

Supposing all other firms continue to submit β_n, let $\tilde{\pi}(m,n)$ for $m \geq 2$ denote a firm's expected profit from submitting a bid strictly below β_n given that it is placed in the mth position in the game with n firms. A firm's expected payoff from such a "downward deviation" is thus

$$\sum_{m=2}^{n}\frac{1}{n-1}\tilde{\pi}(m,n).$$

Finally, supposing all other firms submit β_n, let $\hat{\pi}(1,n)$ denote a firm's expected profit (excluding payment to the platform) from submitting a bid strictly higher than β_n and winning the sponsored position. The expected payoff from such an "upward deviation" is

$$\hat{\pi}(1,n)-\beta_n.$$

Combining these conditions, placing an identical bid of β_n is a best reply for all firms if and only if

$$\frac{n}{n-1}\hat{\pi}(1,n)-\frac{1}{n-1}\sum_{m=1}^{n}\pi(m,n)\leq\beta_n\leq\sum_{m=1}^{n}\pi(m,n)-\frac{n}{n-1}\sum_{m=2}^{n}\tilde{\pi}(m,n). \tag{2}$$

The remainder of the argument shows that for large n the LHS of this expression approaches $p(1-q)(1-G(\bar{r}))$, while the RHS approaches $p(1-q)(1-G(\bar{r}|\bar{\theta}))$. Thus, for large n there is a continuum of equilibrium bids. By choosing a reserve price that is close to the highest equilibrium bid, the platform can easily resolve this equilibrium multiplicity to its own advantage achieving a profit close to $p(1-q)(1-G(\bar{r}|\bar{\theta}))$.

2.2 Binding Outside Option

We now consider the case in which the consumer's outside option is binding, i.e., the outside option has a value $\eta > \bar{r}$. If the platform continues to assign the firm with the highest match score to the sponsored position but otherwise uniformly obfuscates firms across organic positions, the consumer inspects the sponsored position, but never inspects an organic one. Therefore, even though such an obfuscating algorithm is asymptotically optimal in maximizing the sponsored position auction revenue as n grows large, the platform only earns commission when the consumer buys from the firm in the sponsored slot. Below we show that, in this case, the platform can do better than obfuscating over the entire set of firms.

Below we allow the platform to divide its $n-1$ organic positions into two sets, so-called *premium positions* and the remaining ones.[17] Consider premium positions for example as those located on the first page or otherwise endorsed by the platform.[18] This allows us to formally define an algorithm that obfuscates firms with the best match scores across premium positions:

Definition 3. *An algorithm $a \in \mathcal{A}_n$ uniformly obfuscates firms with higher match scores across premium positions if it assigns each of the firms that lose the auction—and whose match scores exceeds a value ϕ—to each of the premium positions with uniform probability.*

A search platform, leveraging big data techniques, has a fairly accurate idea of the match value distribution of products and the consumer's outside option. Based on this information, the platform commits to an algorithm that assigns the sponsored position to the firm with the highest match score and randomly assigns all firms that lose the auction and whose match score exceeds a value ϕ to premium positions.

When the consumer then types in a keyword, the platform learns the consumer's match scores and executes the algorithm. Theorem 2 below establishes that there is a sequence of cutoff values $(\phi_n)_{n\in\mathbb{N}}$ such that an algorithm that assigns the likely best matches to premium positions is approximately optimal for the platform when the number of firms is large.

Theorem 2. *If $\eta > \bar{r}$, there is a sequence of algorithms that uniformly obfuscate firms with higher match scores across premium positions that is asymptotically optimal. Along this sequence, the firm with the highest match score wins the sponsored position.*

The main idea behind this result is as follows. Let $r(\phi)$ be the function satisfying

$$\begin{aligned} s &= \mathbb{E}\left[\max\{v_i - r(\phi), 0\} | \theta_i \geq \phi\right] \\ \Leftrightarrow \quad s &= \int_{r(\phi)}^{\bar{v}} (1 - G(v|\theta \geq \phi))dv. \end{aligned} \tag{3}$$

[17] That is, $a : \mathbb{R}_+^n \times [\underline{\theta}, \bar{\theta}]^n \times Z \to X \times \{0, 1, \ldots, n-1\}$ where $a(\mathbf{b}, \boldsymbol{\theta}, z) = (x, \tilde{n})$ implies that the algorithm selects ranking x and demarcates the first $\tilde{n}$ organic positions as premium positions.

[18] Note that in Section 2.1—as will become clear below—the platform asymptotically weakly prefers not to define premium positions.

In words, absent additional information, the consumer is indifferent between searching a premium position and a product with value $r(\phi)$ if the cutoff match score for premium positions is ϕ.[19] Note that from an ex ante perspective, the consumer inspects a premium position only if $\phi > \zeta$ with $\zeta = r^{-1}(\eta)$. Moreover, observe that—conditional on $\theta > \phi$—draws from premium positions are independent for the consumer.

With an algorithm that assigns the sponsored position to a firm with the highest match score, however, the consumer inspects the sponsored position first, thereby learning about the match scores of firms in premium positions. With a small number of firms, a particularly low match value at the sponsored position may induce a consumer to become so pessimistic that they prefer their outside option η over inspecting any premium position, even if $\phi > \zeta$.

If the number of firms to be ranked is large, however, the Mixing Principle of Consumer Search not only allows us to define an upper bound for platform profits,

$$(1 - q)p(1 - G(\eta|\bar{\theta})) + pq,$$

but also establishes that the effect of learning from the sponsored slot vanishes. No matter how low the match value of the firm in the sponsored position the consumer observes, it will not sway them enough to choose the outside option η over inspecting another premium position. In the proof of Theorem 2 we show that one sequence of premium cutoffs that establishes asymptotic optimality is given by $\phi_n = \zeta + \frac{1}{n}$.

As the number of firms grows large, i.e, $n \to \infty$, $r(\phi_n)$ converges from above to $r(\zeta) = \eta$. Thus, on the one hand—invoking the logic derived in the proof of Theorem 1—the revenue accrued from the auction for the sponsored position is maximized. On the other hand, learning about firms in premium positions vanishes, and there is a premium position with a positive option value relative to the outside option η with arbitrarily high probability. Therefore, the consumer prefers inspecting any premium position over their outside option. It follows that an algorithm uniformly obfuscating a set of firms across premium positions is approximately optimal when the number of firms is large.

Intuitively, we have shown that our main result generalizes when the platform faces competition or when substitute products are readily available. The platform exhibits an incentive to assign the sponsored position to the firm with the highest match score, while it

[19] We assume throughout that $\eta < r(\bar{\theta})$.

wants to provide the consumer with as little information as possible through organic search results, i.e., obfuscate them to the maximal extent possible subject to the constraint that the consumer prefers inspecting premium positions over leaving the platform altogether and taking their business elsewhere. As such, premium positions are a viable instrument for platforms to address threats such as competition and substitute products.

2.3 Improved Platform Information

Now that we better understand for which purposes the platform may (not) use the information it possesses, we can also answer the question how the market is affected by the platform having more accurate information. Consider an improvement in the quality of the platform's information in the sense of Lehmann (1988). That is, if $F(\theta_i|v_i)$ is the platform's initial score distribution, then the platform has better information if its new score distribution $\tilde{F}(\theta_i|v_i)$ is such that

$$\tilde{F}^{-1}\left(F(\theta_i|v_i)|v_i\right)$$

is nondecreasing in v_i for all θ_i.[20]

The next proposition argues that both the platform and consumers are better off if the platform has better information.

Proposition 1. *When there are many firms, improving the quality of the platform's information leads to higher platform profit and consumer surplus under the algorithms specified in Theorems 1 and 2.*

When the platform has the ability to design its algorithm with more accurate information about consumers' match values, it can create a sponsored position that is more valuable for firms, and extract more revenue from the sponsored search auction. Moreover, better information can also lead to more revenue through commission fees as the platform has more latitude to keep consumers from exiting the platform by designing more attractive premium positions.

Proposition 1 is consistent with the interpretation of search platforms being critical gatekeepers in online markets. The platform sells "preferred access" to consumers to firms, which is

[20] Dewatripont et al. (1999) and Persico (2000) discuss economic applications of Lehmann information.

more valuable to a firm if the likelihood that the consumer likes the firm's product is higher. This is why more accurate information increases the platform's profits.

Interestingly, consumers are also better off as they are more likely to find a product they will like enough so that they will not continue searching. They benefit from improved information both through a higher expected match value conditional on buying and through a reduced expected total search cost. Firms may be slightly worse off as the ones with an organic slots are less likely to sell (even though for large n, their profits were anyway already small to begin with) and the firm in the sponsored slot has paid so much more that it is indifferent between winning and losing.

3 Generalizations

In this Section we show that our main result continues to hold when we take into account that (i) firms may also have some information regarding consumers' match values, or (ii) firms charge different prices.

3.1 Privately Informed Firms

We have so far assumed that only the platform has access to relevant information regarding consumer preferences, and firms do not possess such information. We think this is relevant in many instances where firms do not have the relevant technology in place to digest large amounts of information. However, there are other instances where firms also do have relevant information in addition to the platform, perhaps from other sales channels. Here, we therefore consider the situation where firms also have some information about how well their product fits a particular search query.

To model those instances, suppose that in addition to the platform receiving a match score $\theta_i \in [\underline{\theta}, \bar{\theta}]$, each firm receives a private signal $t_i \in [\underline{t}, \bar{t}]$ and that the consumer's match value with firm i is drawn from a distribution $G(v_i|t_i + \theta_i)$. The match scores t_i are independently and identically distributed (IID) across firms according to a compactly supported, atomless distribution $\hat{F}(t_i)$. Denoting $y_i = t_i + \theta_i$, we assume, similar to the main model, that $G(v_i|y_i')$ has likelihood ratio dominance over $G(v_i|y_i)$ whenever $y_i' \geq y_i$.

Define a firm's *adjusted bid* $\psi(b_i, \theta_i)$ to be a smooth and strictly increasing function of a firm's bid on $[0, p]$ and score on $[\underline{\theta}, \bar{\theta}]$. Consider an algorithm which awards the sponsored slot to the firm with the highest adjusted bid and uniformly and randomly assigns all other firms to organic positions. Normalize the smallest adjusted bid to $\psi(0, \underline{\theta}) = 0$. We suppose that for all bids $b_i > p$ the adjusted score is the same as if they had bid zero $\psi(b_i, \theta_i) = \psi(0, \theta_i)$. Specify the payment rule, ρ_n, so that the firm that wins the auction pays the value of its bid.[21] As bidding higher than p is dominated by bidding zero, it is without loss of generality to restrict the firms' strategy space to bids in $[0, p]$. Because the distribution of θ_i is atomless, the probability of a tie is zero. Thus, unlike the main model, each firm's expected profit is continuous in the bids on $[0, p]$.

Denote $\bar{b} \equiv (1-q)p(1 - G(\bar{r}|\bar{t} + \bar{\theta}))$ as the expected profit a firm with private signal $\bar{t}$ and a platform's signal $\bar{\theta}$ would make if the consumer would visit that firm at their first search and decide to engage in optimal sequential search afterwards.

We can then state the following proposition, which is the analogue of our main result for the case where firms also have some private information.

Proposition 2. *Suppose firms have private information. There is a sequence of algorithms that uniformly obfuscate firms with higher match scores across premium positions that is asymptotically optimal. Along the sequence the sponsored firm's match value distribution converges to* $G\left(v|\bar{\theta} + \bar{t}\right)$ *and the platform's profit converges to a mass point at* $\bar{b}$.

It is not difficult to see that if n is large and the combined information of the winning firm and the platform about the match score is the same as the platform's information in the main model, the platform earns the same profits, whether or not the firms have private information. The reason is that if firms do not have information as in the main model, they know that the platform uses match scores to allocate the sponsored slots and take this already into account when making the bid. If, on the other hand, firms have some private information in the form of a match score they realize that the firm with the highest combined match score $\bar{t} + \bar{\theta}$ will win the auction and that as far as the allocation of organic slots is concerned the platform ignores all information. Thus, for a given total match score of the winner, sales will be independent of whether firms have private information.

[21] Note that this marks a minor difference with the main result where we considered a second-price auction. Second-price auctions are a bit more difficult to handle as when firms are ranked according to their adjusted bid, the second-highest adjusted bid may have a bid that is higher than the highest-ranked adjusted bid. When n is large, this difference becomes negligible, however, as the second-highest (adjusted) bid is arbitrarily close to the highest (adjusted) bid.

3.2 Heterogeneous Prices

The main model not only stipulates that prices are exogenously determined, but also that they are the same across firms. For an analysis of the welfare consequences of sponsored positions it is important to acknowledge, however, that different firms are likely to have different prices. In this Section, we accommodate this fact. To keep the analysis tractable, we assume there is a finite, but rich set of prices charged by different firms. Moreover, and in line with considering that the sales through a specific search query are a relatively small fraction of the overall sales of a firm, prices remain determined independently of a search query.[22] Allowing for different firms charging different prices, we introduce a finite set of prices $\mathcal{P} = \{p^1, p^2, ..., p^K\}$.[23] Without loss of generality, assume that $p^1 < p^2 < ... < p^K$.

We denote the distribution over $\mathcal{P}$ by H with probability mass function h. For this setting, we redefine the ex ante reservation value $\bar{r}$ as the unique solution to

$$\sum_{p\in\mathcal{P}} h(p) \left(\int_{\bar{r}+p}^{\bar{v}} (v - p - \bar{r})\, \mathrm{d}G(v) \right) = s. \tag{4}$$

Observe that the above definition of $\bar{r}$ differs slightly from the one in the previous Section in so far as the reservation value of a randomly chosen firm accounts for the (distribution of) price(s) a consumer encounters at the firm. For consistency, we let η denote the value of the outside option net of any price the consumer may have to pay for it. Defining $\gamma := \max(\eta, \bar{r})$, if the platform uniformly obfuscates and $n \to \infty$, then a consumer whose best searched firm so far has a match value v at price p continues to search (randomly) if and only if $v - p < \gamma$.

Finally, to focus on the main issues at hand, we introduce two simplifying assumptions. First,

$$\int_{\gamma+p^K}^{\bar{v}} \left(v - p^K - \gamma \right) \mathrm{d}G(v|\bar{\theta}) \geq s. \tag{5}$$

If (5) holds, there are no firms with prices so high that a consumer would not want to pay the search cost s to inspect them, even if she knew that the firm's match score is $\bar{\theta}$. Second,

[22] Janssen and Williams (2022) consider a model where a social influencer recommends followers to inspect a certain product. They consider that firms may change their prices depending on whether or not they are recommended. Their analysis suggests that the conclusions we derive here could be extended to situations where firms' prices are endogenous. In particular, a firm that wins a sponsored slot (as the recommended firm in Janssen and Williams (2022)) will optimally adjust its price (upwards) in response to the favorable news of being awarded the sponsored slot. This will further increase the willingness to bid to get it and boost the platform's profits.

[23] We implicitly assume here that the reason why prices vary is orthogonal to match values. We could further micro-found this assumption by introducing heterogeneous marginal costs. The inclusion of heterogeneous costs, however, would not alter any of the qualitative predictions in this Section. To simplify the exposition of the analysis, we thus continue to assume that marginal costs are zero.

ensuring that the platform benefits from offering a sponsored slot

$$p^K q < \max_{p \in \mathcal{P}} p(1 - G(\gamma + p|\bar{\theta})). \tag{6}$$

The assumption trivially holds for any $q < 1$ if all firms charge the same price as in the main model.

Under heterogeneous prices, the following two implications are reminiscent of our previous analysis. First, by the Mixing Principle of Consumer Search, a consumer will not purchase from the sponsored firm if $v_s - p_s < \gamma$ for n large enough. Second, if the platform wants the consumer to inspect particular slots, then those slots must have a reservation value that is weakly higher than γ for large n.

Therefore, the following algorithm allows the platform to achieve asymptotically maximum profits. Suppose the platform runs an adjusted second-price auction and places the firm with the highest bid in the sponsored slot. For large enough n, the price of firms with the highest willingness to bid for the sponsored slot (and thus the price of a firm that wins the sponsored slot) equals $p^* \in \arg\max_{p \in \mathcal{P}} p\left(1 - G(\gamma + p|\bar{\theta})\right)$. Organic slots are divided into premium and non-premium ones. Importantly, the platform now uses premium positions regardless of the value of the outside option because it wants the consumer to inspect products with the highest prices to maximize commission fees. To incentivize the consumer to inspect firms in premium positions despite the higher prices, the algorithm requires a firm's match score to exceed a minimum threshold. Furthermore, the platform uniformly obfuscates these firms across premium positions to keep the consumer searching. We can thus invoke Definition 2 of uniform obfuscation across premium positions.

Proposition 3. *Given* (5) *and* (6)*, there is a sequence of algorithms that uniformly obfuscate firms with high match scores across premium positions that is asymptotically optimal. Along this sequence, the expected match score of the firm that wins the sponsored position converges to the maximal match score $\bar{\theta}$, and the prices of firms in the premium positions to p^K.*

Thus, obfuscation remains an important tool for the platform even if firms have different prices. Additionally, firms with a match score that, at least for $n \to \infty$, is arbitrarily close to the highest match score have a chance of winning the sponsored slot.

It remains to be understood how the platform can extract almost the entire rent from the winning firm in this case. The relevant intuition is reminiscent of Theorem 1. As $n \to \infty$, the bids of all firms converge to their expected profit when occupying the sponsored slot as for any finite n there is a positive, although vanishing, probability that they win the auction and the probability there is another firm with the same price approaches one. Thus, the auction for the sponsored slot is almost surely decided by the match score as a tiebreaker among firms with price p^*. Since these firms earn equal profits, the adjusted-second price auction underlying Theorem 1 remains asymptotically optimal.

Since the match scores of firms in premium positions and the sponsored approaches the maximal match score, the next result follows immediately from the last part of Proposition 3.

Corollary 1. *Under the asymptotically optimal algorithm, the price of the sponsored firm is lower than the price of a firm in a premium position.*

When assigning firms to premium positions, the platform only cares about the probability the consumer buys from any of these firms. This probability converges to one as long as the probability that the consumer buys at a particular firm in a premium position is positive. In contrast, when deciding which firm to assign the sponsored slot to it is a single firm's profits that the platform seeks to maximize, because this is what firms are willing to bid to obtain the slot. Therefore, the probability of a single firm generating a sale cannot be too small, implying a lower price ($p^* < p^K$) for the firm winning the sponsored slot.

4 Consumer Surplus

This section discusses the important policy question of whether consumers are better or worse off if search platforms (stop) employ(ing) sponsored positions. Clearly, if in the absence of sponsored positions the platform's and consumers' interests are aligned and the platform maximizes consumer surplus (which may be what Brin and Page had in mind when they wrote the phrase quoted in the Introduction) consumers are better off. Asymptotically[24] the platform chooses a perfect ranking, i.e., a ranking where firms with higher match score are ranked above firms with lower match scores.

Most platforms without sponsored slots earn, however, commission fees when a consumer buys from a firm on their website. Using the Mixing Principle for Consumer Search in Lemma 1 it follows that to maximize commission fees the platform continues to have an

[24]An example where this is not the case for small n is available upon request.

incentive to uniformly obfuscate, potentially only over a premium set of firms. Therefore, with a sponsored slot, the overall ranking is more informative as this slot contains a firm with a higher match score. This implies that introducing a sponsored position benefits consumers.

Corollary 2. *If the platform only maximizes sales commission revenue, then introducing a sponsored position increases consumer welfare if n is large.*

Intuitively, the additional information provided to the consumer via the sponsored slot has two effects. First, it ensures that the consumer samples the firm that is most likely to have a high match value. This raises the expected match value of the product the consumer eventually chooses. Second, sampling the firm with the best match score reduces the number of slots the consumer expects to inspect, thereby lowering expected search costs.

If, in addition, firms have different prices there is a third reason why having a sponsored position is better for consumers. As detailed in the previous section, if firms have different prices the platform introduces premium sets to select firms with higher prices to maximize the commission fee. In comparison, the firm in the sponsored slot has a lower price. Thus, the sponsored slot also helps consumers to find a firm with a lower price.

5 Discussion and Conclusion

In this paper, we analyze how selling sponsored positions affects a search platform's ranking of products. When deciding on its ranking, the platform takes into account that consumers are free to choose how to search. The platform faces an incentive to leverage its information about consumers to put the firm it deems most relevant for a consumer in the sponsored slot. This induces the consumer to inspect the sponsored slot first, increasing the sponsored firm's demand. There is a second reason why winning the sponsored slot is positive news to firms as the firm updates its beliefs about its match value. Obfuscation of organic slots also plays a crucial role in creating rankings, since it increases the firms' incentive to acquire the sponsored position by lowering the consumer's benefit of searching organic positions. As a result, the platform increases revenue by introducing sponsored positions.

Importantly, these findings apply when the number of search results (firms) is sufficiently large as is the case for many real-world search platforms. When consumers search for products, and the number of keyword relevant firms is small, learning effects arise and the optimal platform rule varies with the specific details of the environment. We demonstrate the robustness of our main result to (i) firms holding private information about the consumer's match value with their product, and (ii) firms quoting different prices.

Our results have important implications for consumer welfare. In particular, if in the absence of sponsored positions, sales commissions revenue is the only source of revenue, introducing a sponsored slot benefits consumers as it induces a strictly more informative ranking. Moreover, when prices differ across firms the platform has an incentive to create premium sets of organic search results with firms that charge high prices. In contrast, the platform tends to allocate the sponsored position to a firm with a lower price, creating an additional benefit to consumers.

This paper focuses on the incentives of search platforms how to use their large data sets on consumer search and purchase behavior. Importantly, these incentives shape online markets as the actions of search platforms determine how consumers search. We see several fruitful directions for future research. One issue relates to the number of sponsored positions. While we restrict the platform to sell a single sponsored position throughout the paper, many real-world platforms feature multiple sponsored positions. As a result, how to optimally allocate these slots poses an important question.

As a first step towards such an analysis, observe that the results in this paper generalize to a platform selling a fixed number k of sponsored positions. With exogenously given prices the platform cannot do better than designing an algorithm that assigns the k firms with the highest match scores to the k sponsored positions while obfuscating organic ones. A firm conducting the majority of its business through a search platform, however, may set its prices strategically taking into account its likelihood of obtaining a sponsored position. How do the platform's incentive to rank firms and design auctions for sponsored positions change in response? Relatedly, it is important to understand what determines the optimal number of sponsored positions itself.

References

C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer Science & Business Media, 2006.

S. P. Anderson and R. Renault. Pricing, product diversity, and search costs: A Bertrand-Chamberlin-Diamond model. *The RAND Journal of Economics*, 30(4):719–735, 1999.

S. P. Anderson and R. Renault. Search direction: Position externalities and position auction bias. *CEPR DP # 16724*, 2021.

M. Armstrong and J. Zhou. Paying for prominence. *Economic Journal*, 121(556):F368–F395, 2011.

M. Armstrong and J. Zhou. Consumer information and the limits to competition. *American Economic Review*, 112(2):534–77, 2022.

S. Athey and G. Ellison. Position auctions with consumer search. *The Quarterly Journal of Economics*, 126(3):1213–1270, 2011.

E. J. Balder. Generalized equilibrium results for games with incomplete information. *Mathematics of Operations Research*, 13(2):265–276, 1988.

H. Bar-Isaac and S. Shelegia. Monetizing steering. *CEPR DP #17685*, 2022.

M. R. Baye and J. Morgan. Information gatekeepers on the internet and the competitiveness of homogeneous product markets. *American Economic Review*, 91(3):454–474, 2001.

S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30:107–117, 1998.

Y. Chen and C. He. Paid placement: Advertising and search on the internet. *The Economic Journal*, 121(556):F309–F328, 2011.

A. De Corniere and G. Taylor. A model of biased intermediation. *The RAND Journal of Economics*, 50(4):854–882, 2019.

M. Dewatripont, I. Jewitt, and J. Tirole. The economics of career concerns, part I: Comparing information structures. *The Review of Economic Studies*, 66(1):183–198, 1999.

R. Donnelly, A. Kanodia, and I. Morozov. Welfare effects of personalized rankings. *SSRN # 3649342*, 2022.

K. Eliaz and R. Spiegler. A simple model of search engine pricing. *The Economic Journal*, 121(556):F329–F339, 2011.

D. Garcia and S. Shelegia. Consumer search with observational learning. *The RAND Journal of Economics*, 49(1):224–253, 2018.

A. Ghose, P. G. Ipeirotis, and B. Li. Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, 60(7):1632–1654, 2014.

M. Janssen and C. Williams. Influencing search. *CEPR DP # 15811*, 2022.

M. C. W. Janssen, A. Parakhonyak, and A. Parakhonyak. Non-reservation price equilibria and consumer search. *Journal of Economic Theory*, 172:120–162, 2017.

E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6): 2590–2615, 2011.

T. T. Ke, S. Lin, and M. Y. Lu. Information design of online platforms. *HKUST Business School RP # 2022-070*, 2022.

E. Lehmann. Comparing location experiments. *Annals of Statistics*, 16(2):521–533, 1988.

F. Long, K. Jerath, and M. Sarvary. Designing an online retail marketplace: Leveraging information from sponsored advertising. *Marketing Science*, 41:115–138, 2022.

E. Mossel, M. Mueller-Frank, A. Sly, and O. Tamuz. Social learning equilibria. *Econometrica*, 88(3):1235–1267, 2020.

M. Motta and A. Penta. Market effects of sponsored search auctions. *BSE Working Paper 1356*, 2022.

V. Nocke and P. Rey. Consumer search, steering and choice overload. *Journal of Political Economy*, forthcoming, 2023.

N. Persico. Information acquisition in auctions. *Econometrica*, 68(1):135–148, 2000.

M. Rothschild. Searching for the lowest price when the distribution of prices is unknown. *Journal of Political Economy*, 82(4):689–711, 1974.

N. S. Sahni and H. S. Nair. Does advertising serve as a signal? evidence from a field experiment in mobile search. *Review of Economic Studies*, 87(3):1529—-1564, 2020.

R. M. Ursu. The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science*, 37(4):530–552, 2018.

M. L. Weitzman. Optimal search for the best alternative. *Econometrica*, 47(3):641–654, 1979.

A. Wolinsky. True monopolistic competition as a result of imperfect information. *The Quarterly Journal of Economics*, 101(3):493–511, 1986.

C. Yu. The welfare effects of sponsored product advertising. *Stanford University*, 2024.

A Appendix: Preliminaries

To prove our asymptotic results, we consider a framework with infinitely many firms $i \in \mathbb{N}$ in which we embed the model with finitely many firms. Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots) \in \Theta$ be the vector of scores, $\mathbf{v} = (v_1, v_2, \dots) \in V$ the vector of match values, and z the nonatomic random variable with support Z for randomization. We maintain the distributional assumptions of Section 1, denote the probability measure on $\Theta \times V \times Z$ by μ and probabilities and expectations w.r.t. μ by $\mathbb{P}[\cdot]$ and $\mathbb{E}[\cdot]$, respectively. Letting $\mathbf{b} = (b_1, b_2, \dots) \in B$ be the bid vector and X the set of firm permutations, the set of algorithms $\mathcal{A}$ is the set of functions $a : B \times \Theta \times Z \to X$ such that $a(\mathbf{b}, \cdot)$ is measurable for all $\mathbf{b} \in B$. The set of payment rules $\mathcal{P}$ is the set of measurable functions $\rho : B \to B$ with $\rho = (\rho^1, \rho^2, \dots)$ such that $\rho^i(\mathbf{b}) \le b_i$.

To embed the finite firm model into this framework, let $\mathcal{A}_n \subset \mathcal{A}$ be the subset of algorithms which only permute the first n firms, which is to say that $a(\mathbf{b}, \boldsymbol{\theta}, z) = x$ implies that $x(i) = i$ for all $i \ge n + 1$. The game with n firms is obtained by requiring the platform to select an algorithm in $\mathcal{A}_n$ and restricting the consumer to only inspect the first n positions.

We present three lemmas that characterize the consumer's search problem in a general environment where the consumer learns while searching. We begin by establishing an algorithm-independent upper bound on the equilibrium expected match value of a consumer engaging in optimal search. The reservation value for a firm known to have the highest possible match score is the unique value r^* that satisfies $d_i(r^*, \mu(\cdot|\theta_i = \bar{\theta})) = 0$, where $d_i(v, \mu)$ is defined in (1).

Lemma A.1. *In every equilibrium of the game with n firms, the expected match value acquired by the consumer who has made a purchase after searching m firms is less than* $u^* \equiv \mathbb{E}[v_i|\theta_i = \bar{\theta}, v_i \ge r^*]$*, for all* $m \le n$*.*

Proof. We first claim that the consumer halts their search and buys from a firm upon inspecting its good and finding a match value above r^*. Upon inspecting a single firm, the expected utility from continuing to search depends on the consumer's beliefs about the match scores of the remaining firms. Informing the consumer precisely of the remaining firms' match scores yields a Blackwell improvement and thus must weakly increase the expected utility from continued search. Also, setting the remaining firms' match scores to each equal the largest possible score $\bar{\theta}$ weakly increases the expected utility from continued search. In this case, it is optimal to follow Weitzman's rule and thus the consumer will halt their search and buy from the first visited firm when its match value exceeds r^*. Furthermore, if the match values uncovered at the first $m-1$ visited firms are below r^* and the consumer decides to visit an mth firm, then this same argument provides that the consumer will immediately buy from this mth firm whenever its value is above r^*. Thus, the claim inductively follows.

Suppose the consumer buys the good sold by firm i when its match score is θ_i and the consumer's expected utility from choosing optimally among the set of continuation strategies that exclude buying immediately from firm i is w. Then the consumer's expected match value in this event is $\mathbb{E}[v_i|\theta_i, v_i \geq w]$. Due to the fact that this expectation is increasing in w and θ_i as well our previous arguments ensuring that $w \leq r^*$, the expression is bounded above by $u^* \equiv \mathbb{E}[v_i|\theta_i = \bar{\theta}, v_i \geq r^*]$. Thus $u^* < \infty$ serves as an upper bound for the expected match value acquired by the consumer and is independent of the algorithm, equilibrium strategies, number of firms inspected by the consumer, and number of firms in the market. $\square$

This argument also bounds the probability a consumer engages in a lengthy search because if the search lasts too long, the consumer's expected payoff must be negative.

Lemma A.2. *In every equilibrium of the game and at every decision node for the consumer, for each $\varepsilon > 0$ there exists an $m \in \mathbb{N}$ such that the probability that a consumer searches m or more additional firms is less than ε.*

Proof. Consider a decision node for the consumer in an equilibrium of the game and let h denote the consumer's information set. Let M denote the random variable equal to the number of additional firms inspected by a consumer. From Lemma A.1, a consumer's expected match value given that they search m' more firms is less than u^* for all $m' \in \mathbb{N}$. Hence, a consumer's expected utility is bounded above by

$$\sum_{m' \in \mathbb{N}} (u^* - m' \cdot s)\mathbb{P}(M = m'|h).$$

For a given $m \in \mathbb{N}$, the above expression is less than or equal to

$$\mathbb{P}(M \leq m|h)u^* + \mathbb{P}(M > m|h)(u^* - m \cdot s).$$

As the consumer's expected utility from optimal search must be nonnegative, the above expression must likewise be nonnegative, implying

$$\mathbb{P}(M > m|h) \leq \frac{u^*}{m \cdot s}.$$

Thus, regardless of the algorithm or number of firms, if $m > \frac{u^*}{\varepsilon \cdot s}$, then the probability that the consumer searches beyond m more firms is less than ε. □

Define an "m-restricted" consumer as the consumer in our model, except that we restrict the consumer to only be able to search at most m positions. Let U denote the expected utility for the "unrestricted" consumer in our model and let U_m denote the expected utility for the m-restricted consumer. We obtain the following bound on expected utility.

Lemma A.3. *Consider a decision node at which the consumer has visited k firms and let h denote the consumer's information set. For a given $\varepsilon > 0$, let m be large enough so that the probability of searching $m - k$ more firms is less than ε. Then $\mathbb{E}[U|h] \leq \mathbb{E}[U_m|h] + \varepsilon \cdot u^*$.*

Proof. At the specified decision node, let M denote the random variable corresponding to the number of additional firms searched by the consumer. The m-restricted consumer can always mimic the unrestricted consumer's strategy for the next $m - k$ steps of search and then exit the market whenever the unrestricted consumer would search further, implying that $\mathbb{E}[U_m|h] \geq \mathbb{E}[U|M \leq m - k, h]\, \mathbb{P}(M \leq m - k, h)$. Therefore, we obtain

$$\begin{aligned}\mathbb{E}[U|h] &= \mathbb{E}[U|M \leq m-k, h]\,\mathbb{P}(M \leq m-k|h) + \mathbb{E}[U|M > m-k, h]\,\mathbb{P}(M > m-k|h) \\ &\leq \mathbb{E}[U_m|h] + \mathbb{E}[U|M > m-k, h]\,\mathbb{P}(M > m-k|h) \leq \mathbb{E}[U_m|h] + \varepsilon \cdot u^*\end{aligned}$$

completing the proof. □

A.1 Consumer Search and Mixing

A key idea of our argument is that when there are many firms, then for *any* possible algorithm, when the consumer finds that their match value with the sponsored firm lies below $\bar{r}$, they almost certainly will continue to search rather than buy the sponsored firm's product. The tool we use to make this simple idea concrete is the fact that independently and identically distributed (IID) random variables have the property of *mixing* (Mossel, Mueller-Frank, Sly, and Tamuz, 2020, Lemma 1). Intuitively, mixing means that any event E defined on the same probability space as a sequence of IID random variables $\{Y_i\}_{i \in \mathbb{N}}$ can only be strongly related to a finite number of them. Formally, for every $\epsilon > 0$, except for a set $N \subset \mathbb{N}$ with $|N| < 1/\epsilon^2$, each $i \notin N$ has the property that for every event K only depending on Y_i

$$\left|\mathbb{P}(E \cap K) - \mathbb{P}(E)\,\mathbb{P}(K)\right| < \epsilon. \tag{7}$$

Random variables satisfying (7) are called *ϵ-independent of E*. In our model, match values $\{v_i\}_{i\in\mathbb{N}}$ form an IID sequence of random variables and hence have the mixing property.

Proof of Lemma 1. For a given v and neighborhood N of $d_i(v, \mu)$, there exists an $\epsilon_0 > 0$ such that, if $|\lambda(K) - \mu(K)| < \epsilon_0$ for all $\sigma(v_i)$-measurable $K \subset \Omega$, then $d_i(v, \lambda) \in N$. The conclusion follows by noting that, to the contrary, if the conclusion does not hold, then there is a sequence of probability measures $\{\lambda_k\}_{k\in\mathbb{N}}$ satisfying $|\lambda_k(K) - \mu(K)| < 1/k$ for all $\sigma(v_i)$-measurable K and $d_i(v, \lambda_k) \notin N$ for all $k \in \mathbb{N}$. But this implies a contradiction because the induced probability distribution over v_i, λ_k, weakly converges to the one over v_i given by μ and thus $d_i(v, \lambda_k) \to d_i(v, \mu)$ (see Aliprantis and Border, 2006, Theorem 15.3).

Let $\epsilon = \alpha \cdot \epsilon_0$ and consider $n > 1/\epsilon^2$. From the mixing property (Mossel et al., 2020, Lemma 1), at least one firm $i \leq n$ has a score that is ϵ-independent of E_n. It follows that for any $\sigma(v_i)$-measurable event K

$$\left|\mathbb{P}(K|E_n) - \mathbb{P}(K)\right| < \frac{\epsilon}{\mathbb{P}(E_n)} < \frac{\epsilon}{\alpha} = \epsilon_0$$

and thus $d_i(v, \mu(\cdot|E_n)) \in N$. □

Recall that, during search, the consumer's strategy involves choosing positions in the ranking to search. Let $\hat{v}_i = \hat{v}_i(\omega, a_n, \rho_n, \sigma_n)$ be the match value of the firm located in position i and define the function $\hat{d}_i(v, \lambda) = \int_\Omega \max\{\hat{v}_i - v, 0\} d\lambda(\omega) - s$ which reflects the option value of inspecting position i when the outside option is v given the consumer's uncertainty over the state λ.[25] Under the hypotheses of Lemma 1, the conclusion of Lemma 1 extends to ensure the existence of an n^{**} such that $n \geq n^{**}$ implies that $|\hat{d}_i(v, \mu(\cdot|E_n)) - d_i(v, \mu)| < \delta$ for some position $i \leq n$. This can be seen by first letting ϵ be the value which guarantees $|d_i(v, \mu(\cdot|E_n)) - d_i(v, \mu))| < \delta/2$ whenever v_i is ϵ-independent of E_n. The conclusion follows from noting that at least one position has probability of at least $1 - \frac{1}{\epsilon^2 n}$ of containing a match value that is ϵ-independent of E_n and then applying the triangle inequality.

[25]Note that $\hat{d}_i$ differs from d_i defined previously because d_i is the option value of a given firm i whereas $\hat{d}_i$ reflects the option value of the firm in position i taking into account the ranking as well.

B Appendix: Proofs for Section 2

Lemma B.1 (Upper Bound on Platform Profit)**.** *For any sequence of algorithms, payment rules, and corresponding selected equilibria, $\{(a_n, \rho_n, \sigma_n)\}_{n\in\mathbb{N}}$, the platform's limiting profit is bounded above by* $\limsup_{n\to\infty} \Pi(a_n, \rho_n) \leq (1-q)p\left(1 - G(\max\{\bar{r}, \eta\}|\bar{\theta})\right) + pq$.

Proof. First, letting $D(a_n, \rho_n, \sigma_n)$ the demand for the sponsored firm, we claim that $\limsup_{n\to} D(a_n, \rho_n, \sigma_n) \leq 1 - G(\max\{\bar{r}, \eta\}|\bar{\theta})$. Observe that, because a consumer will never purchase from the sponsored firm when if offers a match values below η, the expected demand received by the sponsored firm is bounded above by $1 - G(\eta|\bar{\theta})$.

In the case where $\bar{r} > \eta$, we can tighten this upper bound as follows. Notice that if $S_n \subset \Theta \times V \times Z$ is the event in which the consumer buys from the sponsored firm in equilibrium σ_n, for a given $\delta > 0$ we have

$$D(a_n, \rho_n, \sigma_n) = \mathbb{P}(S_n \cap \{\hat{v}_1 \geq \bar{r} - \delta\}) + \mathbb{P}(S_n \cap \{\hat{v}_1 < \bar{r} - \delta\}). \tag{8}$$

Denoting $E_n \equiv S_n \cap \{\hat{v}_1 < \bar{r} - \delta\}$, we now show that $\mathbb{P}(E_n) \to 0$ as $n \to \infty$. Toward a contradiction, suppose that $\mathbb{P}(E_n)$ does not vanish in the limit, which is to say that there is an $\alpha > 0$ and a subsequence $\{n_k\}_{k\in\mathbb{N}}$ satisfying $\mathbb{P}(E_{n_k}) > \alpha$ for all k. From Lemma 1, there is a k^* such that, for all $k \geq k^*$, there is a position $i \leq n_k$ with $\hat{d}_i(\hat{v}_i, \mu(\cdot|E_{n_k})) > \hat{d}_i(\bar{r} - \delta, \mu(\cdot|E_{n_k})) > 0$. But this implies a contradiction as the consumer can yield a strictly higher expected payoff by deviating and committing to inspect (or possibly paying the search cost to reinspect) position i whenever they know that the event E_{n_k} has occurred. Therefore $\mathbb{P}(E_n)$ must vanish in the limit.

Next, observe that

$$\mathbb{P}(S_n \cap \{\hat{v}_1 \geq \bar{r} - \delta\}) \leq \mathbb{P}(\hat{v}_1 \geq \bar{r} - \delta) \leq 1 - G(\bar{r} - \delta|\bar{\theta}).$$

Thus, it follows that $\limsup_{n\to\infty} D(a_n, \rho_n, \sigma_n) \leq 1 - G(\bar{r} - \delta|\bar{\theta})$. Moreover, as the inequality holds for all $\delta > 0$, we have $\limsup_{n\to\infty} D(a_n, \rho_n, \sigma_n) \leq 1 - G(\bar{r}|\bar{\theta})$. This proves our first claim.

Using this bound placed on the limiting demand, we now bound the platform's limiting profit. The platform's profit is less than or equal to the winning bid. Let β_n denote the expected winning bid in equilibrium σ_n. Whenever a firm has a positive probability of winning the auction, its equilibrium bid cannot exceed the expected profit conditional on winning. That is, its bid is bounded above by $(1-q)\cdot p\cdot D(a_n, \rho_n, \sigma_n)$. The limiting equilibrium bids, therefore, satisfy $\limsup_{n\to\infty}\beta_n \leq \limsup_{n\to\infty}(1-q)\cdot p\cdot D(a_n, \rho_n, \sigma_n) \leq (1-q)\cdot p\cdot(1-G(\bar{r}|\bar{\theta}))$. Because the expected revenue via commission fees is bounded above by pq, the desired conclusion follows. □

Using uniform obfuscation, it becomes possible to approximately guarantee that the consumer buys from the sponsored firm whenever its match value exceeds $\bar{r}$.

Lemma B.2 (Obfuscation Achieves Bound). *Let $\{(a_n, \rho_n, \sigma_n)\}_{n\in\mathbb{N}}$ be a sequence of uniformly obfuscating algorithms, payment rules, and selected equilibrium and suppose that for all sufficiently large n, the consumer begins search at the sponsored position. Then the probability that the consumer receives a match value above $\bar{r}$ and searches at least one organic firm vanishes in the limit.*

Proof. Let $\delta > 0$ and let E_n denote the event in which the consumer receives a match value above $\bar{r}+\delta$ from the sponsored firm and searches at least one organic firm. We begin by proving the claim for the simpler problem where we replace the consumer in our model with an "m-restricted consumer" who shares the same preferences, but can search at most m firms. Combining this conclusion with Lemma A.3 then proves the original claim.

Fix $m \geq 2$ and let $E_{m,n}$ denote the event in which the m-restricted consumer searches m firms and $\hat{v}_1 \geq \bar{r}+\delta$. We claim that $\lim_{n\to\infty}\mathbb{P}(E_{m,n}) = 0$. Toward a contradiction, suppose to the contrary that there is a constant $\alpha > 0$ and a subsequence $\{n_k\}_{k=1}^{\infty}$ along which $\mathbb{P}(E_{m,n_k}) > \alpha$ holds for all k. By Lemma 1, there is an n^* for which $n \geq n^*$ implies that there is at least one position $i \leq n$ satisfying $\hat{d}_i(\hat{v}_1, \mu(\cdot|E_{n_k})) < \hat{d}_i(\bar{r}+\delta, \mu(\cdot|E_{n_k})) < 0$. However, because the algorithm is uniformly obfuscating, the match value distributions are the same for all $i \in \{m, m+1, \dots, n\}$ and thus $\hat{d}_m(\bar{r}+\delta, \mu(\cdot|E_{n_k})) < 0$. This means that buying from the sponsored position is, in expectation, strictly preferred to inspecting the mth position conditional on the event E_{m,n_k}. This implies a contradiction as E_{m,n_k} is an event in which the consumer inspects the mth position. Therefore, $\lim_{n\to\infty}\mathbb{P}(E_{m,n}) = 0$.

Similarly, consider the case with $m \geq 3$ and let $E_{(m-1),n}$ be the event in which $\hat{v}_1 \geq \bar{r}+\delta$ and the m-restricted consumer searches at least $m-1$ firms. Given $E_{(m-1),n}$, the expected difference in the expected utility between buying the good with the highest match value of those which have been inspected in the first $m-2$ positions and continuing search to the $m-1$th position is greater than[26]

$$-\hat{d}_{m-1}(\bar{r}+\delta, \mu(\cdot|E_{(m-1),n})) - \mathbb{P}(E_{m,n}|E_{(m-1),n}) \cdot u^*.$$

By the same argument above, if $\mathbb{P}(E_{(m-1),n})$ has a subsequence that is bounded away from zero, then this expression is eventually positive along the subsequence, implying a contradiction. Thus, $\mathbb{P}(E_{(m-1),n}) \longrightarrow 0$. Continuing the argument, it inductively follows that $\mathbb{P}(E_{2,n}) \longrightarrow 0$ for an m-restricted consumer with $m \geq 2$.

Let $0 < \epsilon < -u^*/d_i(\bar{r}+\delta, \mu)$ and let m be large enough so that the probability of searching $m-1$ more positions from a decision node is less than ϵ, as provided by Lemma A.2. Using Lemma A.3, conditional on the event E_n, the difference between the expected utility of buying the good from the sponsored position and searching the first organic position is greater than

$$-\hat{d}_2\left(\bar{r}+\delta, \mu(\cdot|E_n)\right) - \mathbb{P}(E_{2,n}|E_n) \cdot u^* - \epsilon \cdot u^*. \tag{9}$$

Repeating the above arguments again, if $\mathbb{P}(E_n)$ has a subsequence bounded away from zero, then (9) is eventually positive for a large enough n, implying a contradiction. Therefore, $\mathbb{P}(E_n) \longrightarrow 0$. As the choice of $\delta > 0$ was arbitrary, the desired conclusion holds. □

Our interest is in equilibria in which all firms submit the same bid and the platform awards the sponsored position to the firm with the highest match score and performs uniform obfuscation within the organic positions. It is a straightforward calculation to show that, for such algorithms and bidding strategies, the consumer optimally begins their search at the sponsored position. The calculation can be found at the end of this Section in Subsection B.1. For these equilibria, the demand for a firm with score that wins the sponsored position and has a score of θ_i eventually exceeds $1 - G(\bar{r}+\delta|\theta_i)$ for all $\delta > 0$. Moreover, the distribution of the winning firm's score weakly converges to a point mass on $\bar{\theta}$, thus, the sequence of the uniformly obfuscating algorithms achieves the upper bound given in Lemma B.1.

[26] Let $\mathbf{v}_k \equiv (\hat{v}_1, \ldots, \hat{v}_k)$. In the event $E_{(m-1),n}$, the realized utility is $(\max \mathbf{v}_{m-1} - (m-1)\cdot s)\mathbf{1}_{E^C_{m,n}} + (\max \mathbf{v}_m - m \cdot s)\mathbf{1}_{E_{m,n}} \leq \max \mathbf{v}_{m-1} - (m-1)s + \max \mathbf{v}_m \cdot \mathbf{1}_{E_{m,n}}$. Subtracting the right side from the realized utility if the consumer were to only inspect $m-2$ firms yields $s - \max\{\hat{v}_{m-1} - \max \mathbf{v}_{m-2}, 0\} - \max \mathbf{v}_m \cdot \mathbf{1}_{E_{m,n}}$. Taking the expectation of this expression given $E_{(m-1),n}$ and applying Lemma A.1 obtains $-\hat{d}_{m-1}\left(\max \mathbf{v}_{m-2}, \mu(\cdot|E_{(m-1),n})\right) - \mathbb{E}[\max \mathbf{v}_m \cdot \mathbf{1}_{E_{m,n}}|E_{(m-1),n}] > -\hat{d}_{m-1}(\bar{r}+\delta, \mu(\cdot|E_{(m-1),n})) - \mathbb{P}(E_{m,n}|E_{(m-1),n}) \cdot u^*$.

Notice that if a firm deviates to a higher bid then it wins the auction for sure, but loses information about whether it has the highest match score. It must be that such a deviation offers an expected profit of $p(1-q)(1-G(\bar{r}))$ in the limit.

Proof of Theorem 1. First, observe that when $\bar{r} > \eta$, that a straightforward application of the Mixing Lemma provides that the expected revenue via commission fees converges to pq under any sequence of algorithms. Otherwise, if we let $\tilde{E}_n$ denote the event in which the consumer takes the outside option when there are n firms, if there is a subsequence along which $\mathbb{P}(\tilde{E}_{n_k}) > \alpha > 0$, there is an $n^* \in \mathbb{N}$ such that $d_i(\eta, \mu(\cdot|\tilde{E}_n)) > 0$ whenever $n \geq n^*$.

Denote the leftmost side of (2) by L_n and the rightmost side by R_n so that (2) simplifies to $L_n \leq \beta_n \leq R_n$. We want to show that $L_n < R_n$ as the number of firms grows large.

Given the algorithm and the consumer's optimal search, we establish the following. First, because total industry profit is bounded from above, the product $\frac{1}{n-1}\sum_{m=1}^{n}\pi(m,n)$ vanishes in the limit. Second, as we verify in Claim 1 below, the expected profit when deviating to a higher bid converges to $\hat{\pi}(1,n) \to (1-q)p(1-G(\bar{r}))$ since the consumer will only make a purchase if their match with the sponsored firm exceeds $\bar{r}$ and the distribution of the match value of the upward deviating firm is $G(v_i) = \int_{\underline{\theta}}^{\bar{\theta}} G(v_i|\theta_i)dF(\theta_i)$. Third, if we let $\tilde{\pi}^*(m,n)$ denote the expected profit for a firm deviating to a lower bid and being assigned to a position $m \geq 2$ given that it has the highest score, we have $\tilde{\pi}(m,n) = \frac{1}{n}\tilde{\pi}^*(m,n) + \frac{n-1}{n}\pi(m,n)$. Hence, we can write $R_n = \pi(1,n) - \frac{1}{n-1}\sum_{m=2}^{n}\tilde{\pi}^*(m,n)$ where the bound on industry profit guarantees that the rightmost term vanishes in the limit. Combining these three observations, we find that $L_n \to (1-q)p(1-G(\bar{r}))$ and $R_n \to (1-q)p(1-G(\bar{r}|\bar{\theta})) > (1-q)p(1-G(\bar{r}))$, where the inequality holds because $G(v|\theta)$ satisfies the MLRP. Thus, there exists an n^* such that $n \geq n^*$ implies $L_n \leq R_n$. Therefore, for $n \geq n^*$ there exist symmetric equilibria in which all firms bid $\beta_n \in [L_n, R_n]$. Thus, we have that $R_n \leq \Pi(a_n, \rho_n)$. Because $R_n \to (1-q)p(1-G(\bar{r}|\bar{\theta}))$ it follows that $(\hat{a}_n, \hat{\rho}_n)_n$ is asymptotically optimal.

Claim 1 (Upward Deviation)**.** *Given $(\hat{a}_n, \hat{\rho}_n)_n$, suppose all other firms play the proposed symmetric equilibrium of bidding β_n and the consumer searches optimally given the belief that all firms bid β_n. Let S'_n be the event in which a firm which deviates to a higher bid makes a sale. Then* $\lim_{n\to+\infty}\mathbb{P}(S'_n) \leq 1 - G(\bar{r})$.

Let $S'_n \subset \Theta \times V \times Z$ and $S'_{m,n} \subset \Theta \times V \times Z$ be the events in which the consumer and m-restricted consumer buy the sponsored product given a deviation, respectively. As before, given the proposed strategies, it is without loss of generality to assume that the consumer searches organic firms in order of their ranking. As the continuation value for the unrestricted consumer is always at least that of the m-restricted consumer, we have $S'_n \subset S'_{m,n}$ for all m and n. For a given $k \in \mathbb{N}$, let $\underline{v}_k \in (\underline{v}, \bar{r})$ satisfy $\frac{G(\bar{r}|\theta_i)-G(\underline{v}_k|\theta_i)}{G(\bar{r}|\theta_i)} \geq 1 - \frac{1}{k}$ for all θ_i. Let V^k denote the event in which $v_S \in [\underline{v}_k, \bar{r}]$ given the deviation. Having assumed that the density $g(v_i|\theta_i)$ is positive on the interior and continuous, we know that there is a constant $c > 0$ such that $\frac{g(v_i|\theta_i)}{g(v_i|\theta_j)} \geq c$ whenever $v_i \in V^k$ for all θ_i and θ_j.

Given the bound placed on the likelihood ratio in the event V^k, the m-restricted consumer's beliefs about the first $m-1$ organic firm's match values converges to the true distribution, which is that they are IID $G(v_i)$. At the limiting distribution, the probability that the m-restricted consumer buys from the sponsored firm given V^k is $\int_{V^k} G(v_S)^{m-1} dG(v_S|V^k) < G(\bar{r})^{m-1}$. Using the inequality

$$\begin{aligned}\mathbb{P}(S'_{m,n}|v_S \leq \bar{r}) &\leq \mathbb{P}(S'_{m,n}|V^k)\,\mathbb{P}(V^k|v_S \leq \bar{r}) + 1 - \mathbb{P}(V^k|v_S \leq \bar{r}) \\ &\leq \mathbb{P}(S'_{m,n}|V^k)\left(1 - \frac{1}{k}\right) + \frac{1}{k}\end{aligned}$$

we obtain

$$\lim_{n\to\infty} \mathbb{P}(S'_{m,n}|v_S \leq \bar{r}) \leq \lim_{n\to\infty} \mathbb{P}(S'_{m,n}|V^k)\left(1 - \frac{1}{k}\right) + \frac{1}{k} < G(\bar{r})^{m-1}\left(1 - \frac{1}{k}\right) + \frac{1}{k}.$$

As the above expression holds for all $k \in \mathbb{N}$, taking the limit as $k \to \infty$ we find $\lim_{n\to\infty} \mathbb{P}(S'_{m,n}|v_S \leq \bar{r}) \leq G(\bar{r})^{m-1}$. Because $S'_n \subset S'_{m,n}$, it follows that $\lim_{n\to\infty} \mathbb{P}(S'_n|v_S \leq \bar{r}) \leq G(\bar{r})^{m-1}$ for all $m \in \mathbb{N}$ and therefore $\lim_{n\to\infty} \mathbb{P}(S'_n|v_S \leq \bar{r}) = 0$. The desired conclusion therefore follows from noting that $\mathbb{P}(S'_n) = \mathbb{P}(S'_n|v_S \leq \bar{r})G(\bar{r}) + \mathbb{P}(S'_n|v_S > \bar{r})(1 - G(\bar{r}))$. □

Define $r(\phi)$ to be function satisfying $s = \int_{r(\phi)}^{\bar{v}} (1 - G(v|\theta \geq \phi))dv$. Throughout, we assume that $\eta < r(\bar{\theta})$ and denote $\zeta = r^{-1}(\eta)$.

Proof of Theorem 2. Let $N_\phi(n)$ denote the random variable corresponding to the number of firms assigned to premium positions given the "premium score threshold" $\phi \geq \zeta$ when there are n firms.

Consider now the variation of our model in which we replace the match score distributions of all n firms with the following truncated distribution $F_\phi(\theta) \equiv \frac{F(\theta)-F(\phi)}{1-F(\phi)} \cdot \mathbf{1}(\theta \geq \phi)$. From

Theorem 1, we know that, in this setting, the platform's limiting revenue from uniform obfuscation in this problem is $\ell(\phi) \equiv (1-q)p(1-G(r(\phi)|\bar{\theta})) + pq$.

Return to our model with the match scores of all n firms distributed according to $F(\theta)$, using the *premium obfuscation* algorithm, the platform's expected revenue conditional on $N_\phi(n) = \tilde{n}$ is equal to $\ell(\phi)$ as it is precisely the same as applying *uniform obfuscation* in the setting with $\tilde{n}+1$ firms and match scores drawn from $F_\phi(\theta)$. Hence, denote $\tilde{\Pi}(\tilde{n}, \phi)$ to be the platform's expected revenue conditional on $N_\phi(n) = \tilde{n}$. From the above, $\lim_{\tilde{n}\to\infty} \tilde{\Pi}(\tilde{n}, \phi) = \ell(\phi)$. Noting that $\mathbb{P}(N_\phi(n) \leq n^\dagger | n) \longrightarrow 0$ for all $n^\dagger \in \mathbb{N}$ provides that the assigning firms with quality threshold ϕ to premium positions yields the platform a limiting revenue of $\ell(\phi)$.

Denoting $\phi_j = \zeta + 1/j$, we see that $\ell(\phi_j) \longrightarrow (1-q)p(1-G(\eta|\bar{\theta})) + pq$ as $j \longrightarrow \infty$. Letting y_{nj} denote the platform's expected revenue in the game with n firms when it uses premium quality threshold ϕ_j. Given that $\lim_{n\to\infty} y_{nj} = \ell(\phi_j)$ for all j and $\lim_{j\to\infty} \ell(\phi_j) = \ell(\zeta)$ there must exist a function $j^*(n)$ such that $\lim_{n\to\infty} y_{nj^*(n)} = \ell(\zeta)$. This completes the proof. $\square$

Proof of Proposition 1. As Lemma B.1 identifies the maximal limiting profit under the information structures as $(1-q)p(1-G(\gamma|\bar{\theta})) + pq$ and $(1-q)p(1-\tilde{G}(\gamma|\bar{\theta})) + pq$ respectively with $\gamma = \max\{\bar{r}, \eta\}$, the result follows from showing $\tilde{G}(\gamma|\bar{\theta}) \leq G(\gamma|\bar{\theta})$. Drawing from the argument for Theorem 5.1 in Lehmann (1988), let $\{\alpha_m\}$ be a vanishing sequence of values in $(0,1)$, $\{t_m\}$ the sequence satisfying $F(t_m|\gamma) = 1-\alpha_m$, and $\{\tilde{t}_m\}$ the sequence satisfying $\tilde{F}(\tilde{t}_m|\gamma) = 1-\alpha_m$. Due to the fact that $\tilde{F}$ is more accurate than F, we have $F(t_m|v_i) \leq \tilde{F}(\tilde{t}_m|v_i)$ for all $v_i < \gamma$ and $F(t_m|v_i) \geq \tilde{F}(\tilde{t}_m|v_i)$ for all $v_i > \gamma$. Consider the two posterior probabilities

$$G(\gamma|\theta_i \geq t_m) = \frac{\int_{\underline{v}}^{\gamma} (1-F(t_m|v_i))\, g(v_i)dv}{\int_{\underline{v}}^{\bar{v}} (1-F(t_m|v_i))\, g(v_i)dv}$$

$$\tilde{G}(\gamma|\theta_i \geq \tilde{t}_m) = \frac{\int_{\underline{v}}^{\gamma} \left(1-\tilde{F}(\tilde{t}_m|v_i)\right) g(v_i)dv}{\int_{\underline{v}}^{\bar{v}} \left(1-\tilde{F}(\tilde{t}_m|v_i)\right) g(v_i)dv}.$$

By rearranging terms, we see that $\tilde{G}(\gamma|\theta_i \geq \tilde{t}_m) \leq G(\gamma|\theta_i \geq t_m)$ if and only if

$$\frac{\int_{\underline{v}}^{\gamma}\left(1-\tilde{F}(\tilde{t}_m|v_i)\right)g(v_i)dv}{\int_{\underline{v}}^{\gamma}\left(1-F(t_m|v_i)\right)g(v_i)dv} \leq \frac{\int_{\gamma}^{\bar{v}}\left(1-\tilde{F}(\tilde{t}_m|v_i)\right)g(v_i)dv}{\int_{\gamma}^{\bar{v}}\left(1-F(t_m|v_i)\right)g(v_i)dv} \tag{10}$$

which must hold as the left side is less than one while the right side exceeds one. Thus, $\tilde{G}(\gamma|\theta_i \geq \tilde{t}_m) \leq G(\gamma|\theta_i \geq t_m)$ for all m. At the same time, $G(\gamma|\theta_i \geq t_m) \rightarrow G(\gamma|\bar{\theta})$ and $\tilde{G}(\gamma|\theta_i \geq \tilde{t}_m) \rightarrow \tilde{G}(\gamma|\bar{\theta})$ as $m \rightarrow +\infty$, implying $\tilde{G}(\gamma|\bar{\theta}) \leq G(\gamma|\bar{\theta})$. □

B.1 Calculations

Consider the problem of the m-restricted consumer. Under the proposed equilibria the match value distribution of the sponsored firm converges to $G(v|\bar{\theta})$ and the distributions for the first $m-1$ organic positions converge to the marginal distribution $G(v)$. The optimal search strategy at the limiting distribution is to follow Weitzman's rule, first inspecting the sponsored firm. As m grows large, the expected payoff then approaches

$$\int_{\underline{v}}^{\bar{v}} \max\{v, \bar{r}\} dG(v|\bar{\theta}) - s. \tag{11}$$

Of the strategies that do not involve beginning search at the sponsored position, the optimal one involves beginning search at an organic position and then following Weitzman's rule thereafter. As m grows large, the expected payoff from this strategy converges to

$$\iint \max\{v', \min\{v, \bar{r}^*\}, \bar{r}\} dG(v')dG(v|\bar{\theta}) - (1 + G(r^*))s. \tag{12}$$

Noting that (11) is strictly larger than (12), it follows that an m-restricted consumer strictly prefers to begin search at the sponsored position as n grows large, and that the loss from beginning search at an organic position is bounded away from zero for large m. Combining this with Lemma A.3, it follows that, under the proposed equilibrium and for large enough n, the unrestricted consumer strictly prefers to begin their search at the sponsored position.

C Appendix: Proofs for Section 3

C.1 Privately Informed Firms

For simplicity, this Section proves Proposition 2 under the assumption that the outside option is non-binding. The argument readily extends to the case with a binding outside option and premium positions following the approach used to prove Theorem 2.

Consider an algorithm which awards the sponsored slot to the firm with the highest adjusted bid and uniformly and randomly assigns all other firms to organic positions. Denote $\bar{b} \equiv (1-q)p(1-G(\bar{r}|\bar{t}+\bar{\theta}))$. A firm's **adjusted bid** $\psi(b_i, \theta_i)$ is a smooth and strictly increasing function of firm's bid on $[0, \bar{b}]$ and score on $[\underline{\theta}, \bar{\theta}]$. Normalize the smallest adjusted bid to $\psi(0, \underline{\theta}) = 0$. We suppose that for all bids $b_i > \bar{b}$ the adjusted score is the same as if they had bid zero $\psi(b_i, \theta_i) = \psi(0, \theta_i)$. Specify the payment rule so that the firm that wins the auction pays the value of its bid. As bidding higher than $\bar{b}$ is dominated by bidding zero, it is without loss in generality to restrict the firms' strategy space to bids in $[0, \bar{b}]$. Because the distribution of θ_i is atomless, the probability of tying is zero. Thus, each firm's expected profit is continuous in the bids on $[0, \bar{p}]$. Using Lemma B.2, we pin down a lower bound on the expected profit for a firm from winning the auction when n grows large.

Lemma C.1. *For all $\epsilon' > 0$ and $\bar{v}' \in (\bar{r}, \bar{v})$ there exists $n' \in \mathbb{N}$ such that, if $n \geq n'$, then a firm with signal $y_i = (\theta_i, t_i)$ that wins the auction generates an expected profit that is greater than $(1-q)p(G(\bar{v}'|y_i) - G(\bar{r}|y_i)) - \epsilon'$.*

Proof. Let $\mathcal{V}_n \subset [\bar{r}, \bar{v}']$ denote the subset of match values such that the consumer buys immediately from the sponsored firm if $\hat{v}_1 \in [\bar{r}, \bar{v}'] \backslash \mathcal{V}_n$ and searches at least one more firm $\hat{v}_1 \in \mathcal{V}_n$. From the Lemma B.2, the Lebesgue measure of $\mathcal{V}_n$ must vanish in the limit. Thus, when a firm with signal $y_i = (\theta_i, t_i)$ wins the auction, it makes a sale immediately with probability $G(\bar{v}'|y_i) - G(\bar{r}|y_i) - \int_{\mathcal{V}_n} dG(v_i|y_i)$.

We argue that $\int_{\mathcal{V}_n} dG(v_i|y_i)$ uniformly converges to zero as $n \to \infty$. The density $g(v_i|y_i)$ is finite for all $y_i \in [\underline{y}, \bar{y}]$ and $v_i \in [\bar{r}, \bar{v}']$. Due to compactness and continuity, there must be a constant $c > 0$ such that $\max_{v_i \in [\bar{r}, \bar{v}'], y_i \in [\underline{y}, \bar{y}]} g(v_i|y_i) \leq c$. But then it follows that $\max_{y_i \in [\underline{y}, \bar{y}]} \int_{\mathcal{V}_n} g(v_i|y_i) dv_i \leq c \cdot \lambda(\mathcal{V}_n)$ and uniform convergence follows. Thus, it follows that, for all $\epsilon' > 0$ there exists $n' \in \mathbb{N}$ such that, if $n \geq n'$, then a deviating firm's profit given that it wins the auction and has a signal of y_i is greater than $(1-q)p(G(\bar{v}'|y_i) - G(\bar{r}|y_i)) - \epsilon'$. $\square$

Proof of Proposition 2. Consider the sequence of uniformly obfuscating algorithms described above. Now, assume that the consumer begins search at the sponsored position. The calculation found in Subsection B.1 verifies that this is optimal in the limit. As all payoffs are continuous in each player's strategy and the strategy spaces are compact, Theorem 3.1 in Balder (1988) establishes the existence of an equilibrium. The argument in Lemma B.1 provides that the limit inferior of the profit derived from any sequence of algorithms is bounded above by $\bar{b}$. We want to show that the uniformly obfuscating algorithms described above achieve this upper bound. Towards this end, we begin by proving the following claim. For each $n \geq 2$, let σ_n be an equilibrium strategy profile for the game featuring n firms.

Let Ψ_n be the random variable denoting the highest adjusted bid in equilibrium σ_n. We claim that Ψ_n weakly converges to point mass on $\bar{\psi} \equiv \psi(\bar{b}, \bar{\theta})$. Toward a contradiction, suppose to the contrary that there is a $\psi^* < \bar{\psi}$, an $\epsilon > 0$, and a subsequence $\{n_k\}_{k \in \mathbb{N}}$ such that $\mathbb{P}(\Psi_{n_k} \leq \psi^*) \geq \epsilon$ for all k. Let i_n be a firm with the lowest expected profit in equilibrium σ_n. The expected profit for i_n must vanish in the limit. Let $\hat{\Psi}_n$ correspond to the highest adjusted bid among firms $j \neq i_n$. As $\hat{\Psi}_n \leq \Psi_n$, it must also be that $\mathbb{P}(\hat{\Psi}_{n_k} \leq \psi^*) \geq \epsilon$ for all k. Let $M = \{\lambda \in \Delta([0, \bar{\psi}]) : \lambda([0, \psi^*]) \geq \epsilon\}$. As $\Delta([0, \bar{\psi}])$ is compact and M a closed subset of $\Delta([0, \bar{\psi}])$, M is also compact. We want to show that, for n large enough, firm i_n can achieve a positive payoff whenever the distribution of $\hat{\Psi}_n$ is given by a probability measure $\lambda \in M$.

Let b^* denote the bid that satisfies $\psi(b^*, \bar{\theta}) = \psi^*$. Let $\epsilon' > 0$ be small enough and $\bar{v}'$ large enough so that $(1-q)p(G(\bar{v}'|\bar{t} + \bar{\theta}) - G(\bar{r}|\bar{t} + \bar{\theta})) - \epsilon' > b^*$. From Lemma C.1, there is an $n' \in \mathbb{N}$ such that $n \geq n'$ implies that the expected net profit from winning the auction for a firm i with signal t_i and score θ_i is at least $(1-q)p(G(\bar{v}'|\theta_i + t_i) - G(\bar{r}|\theta_i + t_i)) - \epsilon'$.

Suppose that the distribution of $\hat{\Psi}_n$ is $\lambda \in M$. Notice that if $\lambda\left(\left[0, \psi\left(0, \bar{\theta}\right)\right)\right) > 0$, then firm i_n's expected profit from always bidding zero is positive. Suppose instead that $\lambda\left(\left[0, \psi\left(0, \bar{\theta}\right)\right)\right) = 0$ so that i_n's bid must be positive to ensure a positive chance of winning the auction. Let $\underline{\psi} = \min \operatorname{supp} \lambda$. Notice that if i_n submits the bid $\underline{b}$ that satisfies $\psi(\underline{b}, \bar{\theta}) = \underline{\psi}$, then the probability of i_n winning the auction is zero. All bids higher than $\underline{b}$ deliver a positive probability of winning the auction. Letting $b_{i_n} \to \underline{b}$ from the right, the distribution of i_n's score conditional on winning the auction weakly converges to a point mass on $\bar{\theta}$. When t_{i_n} is in a neighborhood of $\bar{t}$, then $\iint (1-q)p(G(\bar{v}'|\theta_{i_n}+t_{i_n}) - G(\bar{r}|\theta_{i_n}, t_{i_n}))dF(\theta_{i_n}|\psi(b_{i_n}, \theta_{i_n}) \geq x)d\lambda(x) - \epsilon' > b^* > b_{i_n}$. This means that for all $n \geq n'$, if the distribution of $\hat{\Psi}_n$ is λ, then i_n has a positive expected profit. Since M is compact and firm i_n's expected profit is continuous in the adjusted bid distribution for all other firms, then firm i_n's expected profit is bounded away from zero for all $n \geq n'$. But this implies a contradiction and thus $\liminf_{n\to\infty} \mathbb{P}(\Psi_n \geq \psi^*) = 1$ for all $\psi^* < \bar{\psi}$. The desired conclusion follows. □

C.2 Heterogeneous Prices

Proof of Proposition 3. Consider a sequence of algorithms $(a_n)_{n\in\mathbb{N}}$, each of which assigns the sponsored slot to the highest bidding firm, with ties being broken in favor of higher match scores. Moreover, a_n uniformly obfuscates all firms with $p > p^K - \frac{1}{n}$ and $\theta > \zeta + \frac{1}{n}$ across premium positions, where

$$\int_{\gamma}^{\bar{v}} (v-\gamma)\mathrm{d}G(v|\theta > \zeta) = s.$$

We now show that $(a_n)_{n\in\mathbb{N}}$ asymptotically maximizes (1) commission revenue from non-sponsored slots given the consumer does not buy from the sponsored slot, (2) profits from the sponsored slot, and (3) that $(a_n)_{n\in\mathbb{N}}$ thereby maximizes total platform revenue.

(1) Denote commission earnings from non-sponsored positions, conditional on the consumer not buying from the sponsored slot, by R, which is bounded by $p^K q$. By the Mixing Principle of Consumer search, for $n \to \infty$, there exists—under any search history—almost surely a premium slot with a reservation value that exceeds η. It follows that the conditional probability that the consumer buys from a firm in a premium position converges to 1, and, thus $R(a_n) \to p^K q$.

(2) Note that the profits of the firm in the sponsored slot, denoted by $\pi_s(a_n)$, are bounded by $\max_{p\in\mathcal{P}} \left(1 - G(\gamma + p|\bar{\theta})\right)$. We assume that $\mathcal{P}$ is sufficiently dense so that there exists a unique maximizer $p^* \in \mathcal{P}$. It follows that this expression also constitutes a bound for platform profits from the sponsored slot. Due to the independence of prices, match scores, and match values between firms, the Mixing Principle of Consumer Search implies that a consumer buys from the sponsored if and only if $v_s - p^* \geq \gamma$ as $n \to \infty$, and, thus, $\lim_{n\to\infty} \pi_s(a_n) = \left(1 - G(\gamma + p^*|\bar{\theta})\right)$.

Lemmas S.2 and S.3 in the Online Appendix establish (i) that the winning firm's type approaches $\bar{\theta}$, (ii) that the price of that firm is p^*, and (iii) that the platform is indeed able to extract the expected profits of the firm in the sponsored slot.

(3) To see why $(a_n)_{n\in\mathbb{N}}$ asymptotically maximizes total platform profits, note that

$$\lim_{n\to\infty} (\pi_s(a_n) + R(a_n)) = p(1 - G(\gamma + p|\bar{\theta})) + G(\gamma + p|\bar{\theta})p^K q. \tag{13}$$

Any alternative sequence $(a'_n)_{n\in\mathbb{N}}$ (weakly) reduces both π_s and R because of (1) and (2). It follows that for any alternative sequence of algorithms $(a'_n)_{n\in\mathbb{N}}$ to yield higher total limit profits than $(a_n)_{n\in\mathbb{N}}$, the sequence $(a'_n)_{n\in\mathbb{N}}$ must increase the probability that the consumer buys from a non-sponsored slot (which equals $G(\gamma + p|\bar{\theta})$ under a_n), thereby reducing the probability they buy from the sponsored one (which equals $1 - G(\gamma + p|\bar{\theta})$ under a_n) by the same amount. This, however, is only profitable if $p^* \geq qp^K$ contradicting (6).

□

Search Platforms: Big Data and Sponsored Positions

Online Appendix

Maarten Janssen Thomas Jungbauer Marcel Preuss Cole Williams

S.1 Example 1

In this Section, we discuss an example illustrating how uniform obfuscation can fail to be optimal when there are a small number of firms, because of non-monotonicities in the inference consumers draw from observing their match value at the sponsored slot.

Suppose there are three firms $i = 1, 2, 3$. The consumer's match value is either low ℓ, medium m, or high h and a good is only worth purchasing if it provides at least a medium value. A firm's match score is L when the value is low and H when the value is either medium or high, i.e., the platform can distinguish firms with low match scores from other firms, but cannot distinguish firms with medium and high match scores. Let p_H and p_L denote the marginal probability that a firm's score is high and low, respectively.[27]

Suppose the platform employs the following algorithm. The firm with the highest bid is placed in sponsored positions, ties are broken in favor of the firm with the highest match score, further ties are broken with equal probability. For the two nonsponsored firms, if only one of them has a high signal it is placed in the second position with probability $\alpha \geq \frac{1}{2}$, otherwise they are arranged in the organic positions with equal probability.

Given the algorithm, the consumer's optimal search proceeds in the following manner. If the sponsored firm's value is high h, then the consumer buys it immediately since there is no advantage from continuing. If instead the sponsored firm's value is low ℓ, then given the

[27] This example departs from the assumptions of our model in that the distribution of match values conditional on the match scores do not share the same support. This is insignificant to the particular example since we could modify the distributions to $\mathbb{P}(\{\ell\}|L) = \mathbb{P}(\{m, h\}|H) = 1 - \varepsilon$ so that the conclusion continues to hold for $\varepsilon > 0$ sufficiently small.

algorithm, the consumer learns that all remaining firms must likewise have low match values and so the consumer might as well exit the market. If, however, the consumer observes m in the sponsored slot, then it might still be prudent to continue searching as some remaining firm might deliver a higher match value. To describe the consumer's learning over the course of search, let subscripts denote the index of the position so that the list of possible events are $\{(H_1, H_2, H_3), (H_1, L_2, H_3), (H_1, H_2, L_3), (H_1, L_2, L_3)(L_1, L_2, L_3)\}$ which occur with corresponding probabilities $\{p_H^3, 3(1-\alpha)p_H^2 p_L, 3\alpha p_H^2 p_L, 3p_H p_L^2, p_L^3\}$. The probability that slot two has a high score given that the first does is

$$\mathbb{P}(H_2|H_1;\alpha) = \frac{p_H^3 + 3\alpha p_H^2 p_L}{1 - p_L^3}. \tag{S.1}$$

The probability that slot three has a high score given that the first two also do is

$$\mathbb{P}(H_3|H_1, H_2;\alpha) = \frac{p_H^3}{p_H^3 + 3\alpha p_H^2 p_L}. \tag{S.2}$$

The probability that slot three has a high score given that the first does and the second has a low score is

$$\mathbb{P}(H_3|H_1, L_2;\alpha) = \frac{3(1-\alpha)p_H^2 p_L}{3(1-\alpha)p_H^2 p_L + 3p_H p_L^2}. \tag{S.3}$$

Inspecting the above expressions, we find that (S.2)>(S.3) which naturally implies that the consumer is more optimistic about the third slot upon observing a medium match value in the first two slots than if they were to observe a medium in the first and a low in the second. Also, (S.1)>(S.3) holds true, implying the consumer is more optimistic about the second slot after observing a medium in the first than they are about the third slot upon observing a medium and low value in the first two. The comparison between (S.1) and (S.2) depends on α. As Figure 1 illustrates, increasing α makes proceeding to the second firm more attractive, but continuing to the third less so.

Suppose the parameters are such that, under uniform obfuscation (i.e. when $\alpha = 1/2$), the consumer continues searching when observing m in the first firm and m in the second firm, but halts otherwise. To be concrete, assume that the parameters satisfy $\frac{1}{2}\mathbb{P}(H_2|H_1;\alpha = \frac{1}{2})(h - m) = s$ which leads the consumer to follow the desired search pattern. We compare uniform obfuscation against a nonuniformly obfuscating algorithm with $\alpha = \alpha^* > \frac{1}{2}$ whereby α^* is large enough to ensure that a consumer will inspect the second slot if the first provides

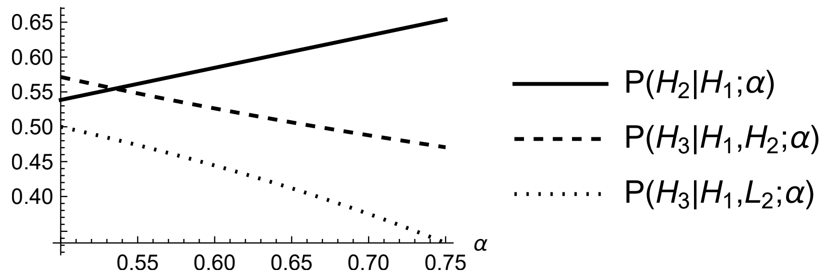
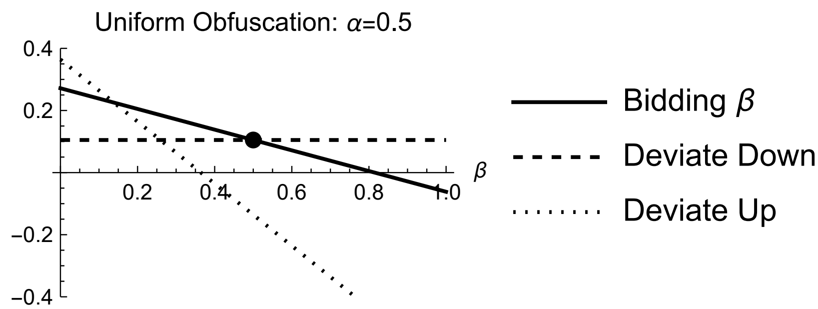


Figure 1: The figure plots the conditional probabilities as a function of α given that match values each occur with equal probability.

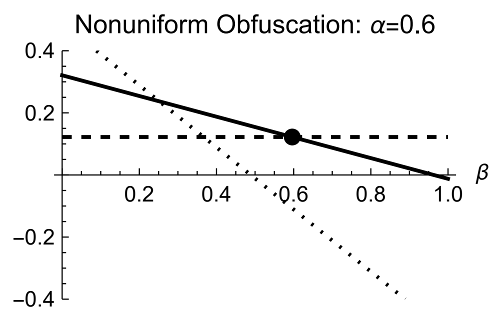
a medium match value, but will search no further. For example, in Figure 1, given our assumptions on parameters, setting $\alpha = 0.6$ guarantees that $\mathbb{P}(H_3|H_1, H_2; \alpha^*) < \mathbb{P}(H_2|H_1; \frac{1}{2})$ and thus it is never optimal for the consumer to inspect the third firm. Notice that by providing some information in the organic slots, the nonuniformly obfuscating algorithm makes inspecting the second firm more desirable, but increases a sponsored firm's return demand as consumers will not inspect the third firm.

For each algorithm, consider a symmetric equilibrium in which each firm bids β. Figure 2 plots the tentative expected equilibrium and deviation profits under the two proposed algorithms for different values of the bid β. Naturally, the expected profit from playing the tentative equilibrium strategy is decreasing in the bid with a slope of $-\frac{1}{3}$ as each of the three firms win with equal probability. Placing a higher bid, no matter how high, ensures that a firm wins the auction, but also makes winning uninformative as it does not provide information about the firm's match score. The slope of the expected profit given an upward deviation is thus -1. Thus, this deviation risks winning in case the firm has a medium match score and consumers continue searching after visiting the sponsored slot. The deviation is optimal for low values of the candidate equilibrium bid, but not for higher values. Offering a bid less than β ensures that a firm does not win the sponsored slot; hence, the deviation profit is simply the expected profit in an organic slot and does not depend on the bid.

We can use Figure 2 to compare the profitability of the two algorithms by comparing the range of bids firms are willing to place. The figure shows that, in principle, there can be a continuum of equilibrium bids. However, we only need to compare the equilibria with the highest bids as the platform can secure a profit equal to this bid by setting a reserve price equal to the highest possible equilibrium bid. As illustrated by Figure 2, using a uniform



(a)



(b)

Figure 2: The figures plot an individual firm's expected profit in a tentative symmetric equilibrium in which all firms bid β under the respective algorithms from likewise bidding β, deviating to a lower bid, and deviating to a higher bid (given that match values each occur with equal probability).

obfuscation algorithm, the platform can achieve a profit of exactly $\frac{1}{2}$, which is identified by finding the highest value for the bid at which no firm wishes to deviate from placing that bid. On the other hand, using the nonuniformly obfuscating algorithm, the platform can secure a profit of approximately 0.6. Thus, due to the learning from match values at the sponsored slots, the platform is better off choosing the nonuniformly obfuscating algorithm.

S.2 Example 2

In this Section, we discuss an example illustrating how the platform may deviate from assigning the firm with the highest match score to the sponsored slot when there are a small number of firms. To this end, consider a similar setup as before in Example 1 where there are three distinct match values l, m, h and where the platform cannot distinguish medium and high values, but perfectly recognizes a low value firm. Thus, when the match value is l, the match score is L and when the match values is either m or h the platform has a match score of H. Now there are four firms, however. Moreover, suppose the consumer would choose any match value, including the lowest l, over their outside option.

Consider two different algorithms, one where the platform always puts the firm with the highest match score in the sponsored position and uniformly obfuscates the organic positions and another, second, algorithm where the platform almost always follows the same algorithm apart from the case where two match scores are L and two are H. In that case the platform puts a firm with the lowest match score in the sponsored position and uniformly obfuscates the organic positions.

We show the conditions under which the second algorithm gives the winning firm a higher probability of selling. In particular, consider that with the first algorithm the consumer buys immediately from the sponsored position whenever it contains an h or l value, but they continue to search if he sees an m value. On the other hand, with the second algorithm the consumer buys immediately from the sponsored position whatever its match value. As the second algorithm makes sure that the consumer will always buy from the sponsored position and never buys from an organic position, it is clear that platform gets its highest possible profit as firms bid maximally to get into the sponsored position.

We first consider under what conditions after observing an m in the sponsored slot the consumer continues to search under the first algorithm, but not under the second, and we start the analysis with the second algorithm. We use backward induction to determine the pay-off for the consumer *if* he decides to continue to search after observing an m. So, consider that after observing an m in the first round, the consumer observes either ll, ml (or lm) or mm in the subsequent two rounds (and we thus focus on whether the consumer wants to inspect the last object or not). If, in the meantime a consumer has found an h object, they will of course buy immediately).

(i) After observing mll in the first three search rounds, the consumer obviously stops searching, as they must update their beliefs in such a way that the last object is also an l. Their pay-off in this case is $m - 3s$.

(ii) After observing mlm or mml in the first three search rounds, we postulate the consumer continues searching as under the second algorithm the only pattern that is consistent with this is for there to be an H object in the last round. Thus, the consumer's pay-off, then would be $\frac{1}{2}(m+h) - 4s > m - 3s$. Thus, we implicitly assume that $h - m > 2s$.

(iii) Finally, after observing mmm in the first three search rounds, we postulate the consumer continues searching. Updating their beliefs, if the consumer continues to search they encounter an L on the last search with probability $\frac{\frac{1}{3}\cdot 4p_H^3 p_L}{\frac{1}{3}\cdot 4p_H^3 p_L + p_H^4} = \frac{4p_L}{4p_L+3p_H}$ and an H with probability $\frac{3p_H}{4p_L+3p_H}$. Thus, the consumer's pay-off would then be $\frac{8p_L+3p_H}{8p_L+6p_H}m + \frac{3p_H}{8p_L+6p_H}h - 4s$ (and this is better than stop searching with a payoff of $m - 3s$, if $\frac{3p_H}{8p_L+6p_H}(h-m) > s$). Thus, we implicitly assume that

$$\frac{h-m}{s} > 2 + \frac{8p_L}{3p_H}. \tag{S.4}$$

Note that this condition also implies that $h - m > 2s$.

Let us then go back one search round and consider that the consumer after observing an m in the first round, they observes either l or m in the next round.

(i) After observing ml in the first two search rounds, we postulate the consumer stops searching, as they must update their beliefs in such a way that the last two objects are either ll or HH. Given the second algorithm, the total ex ante probability that the consumer observes ml in the first two search rounds is $\frac{1}{3}\cdot 4p_H^3 p_L + 4p_L^3 p_H$ and thus continuing to search yields an expected pay-off of

$$\frac{4p_L^3 p_H}{\frac{1}{3}\cdot 4p_H^3 p_L + 4p_L^3 p_H}(m-3s) + \frac{\frac{1}{3}\cdot 4p_H^3 p_L}{\frac{1}{3}\cdot 4p_H^3 p_L + 4p_L^3 p_H}\left(\frac{1}{2}(h-3s) + \frac{1}{2}\left(\frac{1}{2}(m+h) - 4s\right)\right).$$

This is smaller than $m-2s$ (the pay-off if they stop searching) if, and only if, $\frac{1}{2}\left(h-2s\right)+\frac{1}{2}\left(\frac{1}{2}(m+h)-3s\right)<m-2s+\frac{\frac{1}{3}\cdot 4p_H^3p_L+4p_L^3p_H}{\frac{1}{3}\cdot 4p_H^3p_L}s$, or $\frac{3}{4}(h-m)<\left(\frac{3}{2}+\frac{3p_L^2}{p_H^2}\right)s$, which is true if

$$\frac{h-m}{s}<2+\frac{4p_L^2}{p_H^2}. \tag{S.5}$$

(ii) After observing mm in the first two search rounds, we postulate the consumer continues to search all the way until they found an h object, as they must update their beliefs in such a way that the last two objects are either HL, LH or HH (which happens with a total ex ante probability of $\frac{2}{3}\cdot 4p_H^3p_L+p_H^4$) and thus they get an additional pay-off of continuing to search of

$$\begin{aligned}
& \frac{p_H}{p_H+\frac{2}{3}\cdot 4p_L}\left(\frac{1}{2}\left(h-m-s\right)+\frac{1}{2}\left(\frac{1}{2}(h-m)-2s\right)\right) \\
+ \quad & \frac{\frac{2}{3}\cdot 4p_L}{p_H+\frac{2}{3}\cdot 4p_L}\left(\frac{1}{2}(h-m-\frac{1}{2}s)-\frac{3}{4}\cdot 2s\right) \\
= \quad & \frac{3p_H}{3p_H+8p_L}\left(\frac{3}{4}(h-m)-1\frac{1}{2}s\right)+\frac{8p_L}{3p_H+8p_L}\left(\frac{1}{2}(h-m)-\frac{7}{4}s\right) \\
= \quad & \frac{\frac{9}{4}p_H+4p_L}{3p_H+8p_L}(h-m)-\frac{\frac{9}{2}p_H+14p_L}{3p_H+8p_L}s,
\end{aligned}$$

which should be larger than 0 for them to prefer to continue searching, which is the case if

$$\frac{h-m}{s}>\frac{\frac{9}{2}p_H+14p_L}{\frac{9}{4}p_H+4p_L}. \tag{S.6}$$

So, now we can give a condition under which the consumer stops searching after observing an m in the first slot as their overall additional pay-off (incorporating learning and optimal search) of continuing to search after observing an m in the first round instead of stopping immediately is (as the overall probability of this event happening is $p_H^4+4p_H^3p_L+4p_L^3p_H$):

$$\begin{aligned}
& -\frac{4p_L^3p_H+\frac{1}{3}\cdot 4p_H^3p_L}{p_H^4+4p_H^3p_L+4p_L^3p_H}s+ \\
& \frac{p_H^4+\frac{2}{3}\cdot 4p_H^3p_L}{p_H^4+4p_H^3p_L+4p_L^3p_H}\left\{\frac{1}{2}(h-m)+\frac{1}{2}\left(\frac{\frac{9}{4}p_H+4p_L}{3p_H+8p_L}(h-m)-\frac{\frac{9}{2}p_H+14p_L}{3p_H+8p_L}s\right)-s\right\}<0,
\end{aligned}$$

which after combining terms yields

$$
\begin{aligned}
& p_H^2(3p_H+8p_L)\frac{h-m}{6}\left(1+\frac{\frac{9}{4}p_H+4p_L}{3p_H+8p_L}\right) \\
< \quad & \left(\frac{1}{3}p_H^2(3p_H+8p_L)\left(1+\frac{\frac{9}{4}p_H+7p_L}{3p_H+8p_L}\right)+4p_L\left(p_L^2+\frac{1}{3}p_H^2\right)\right)s,
\end{aligned}
$$

or

$$3p_H^2\frac{h-m}{24}(7p_H+16p_L) < \left(\frac{1}{3}p_H^2\left(\frac{21}{4}p_H+19p_L\right)+4p_L^3\right)s,$$

or

$$\frac{h-m}{s} < \frac{42p_H^3+152p_H^2p_L+96p_L^3}{21p_H^3+48p_H^2p_L} \tag{S.7}$$

Let us now do the same exercise for the first algorithm and use backward induction to determine the pay-off for the consumer *if* he decides to continue to search after observing an m. So, consider that after observing an m in the first round, the consumer observes either ll, ml (or lm) or mm in the subsequent two rounds.

(i) After observing ll in search rounds two and three, the consumer encounters an L with probability $\frac{4p_L^3p_H}{4p_L^3p_H+2p_H^2p_L^2}=\frac{2p_L}{2p_L+p_H}$ on the last search and an H with probability $\frac{p_H}{2p_L+p_H}$.

(ii) After observing lm (the same analysis applies to ml) in search rounds two and three, the total ex ante probability of $HHLH$ and $HHLL$ is $\frac{1}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2$, where for example $6p_H^2p_L^2$ is the ex ante probability that there are two H and two L and in that case (under the first algorithm) the sponsored slot is an H and the chance that the first organic slot contains an L is 1/3. So the conditional probability of the last one being H, resp. L, is $\frac{4p_H^3p_L}{4p_H^3p_L+6p_H^2p_L^2}=\frac{2p_H}{2p_H+3p_L}$ and $\frac{3p_L}{2p_H+3p_L}$.

(iii) Finally, after observing mm in search rounds two and three, the total ex ante probability of $HHHH$ and $HHHL$ is $p_H^4+\frac{1}{3}\cdot 4p_H^3p_L$ so the conditional probability of the last one being H, resp. L, is $\frac{p_H}{p_H+\frac{4}{3}p_L}$ and $\frac{\frac{4}{3}p_L}{p_H+\frac{4}{3}p_L}$.

If the consumer continues to search in the first case, they certainly continue to search in all other cases and this is the case if $\frac{p_H}{p_H+2p_L}\frac{m+h}{2}+\frac{2p_L}{p_H+2p_L}m-s>m$, or $\frac{p_H}{p_H+2p_L}\frac{h-m}{2}>s$ or

$$\frac{h-m}{s} > 2\left(1+\frac{2p_L}{p_H}\right). \tag{S.8}$$

It is clear that (S.8) implies (S.4) and (S.6).

What we will do in the subsequent analysis is the following. We show that if (S.8) holds, then the consumer wants to continue searching in all previous search rounds until they have found an h. The easiest way to do so is to show that if (S.8) holds the consumer prefers to search in round t even if in subsequent round $t+1$ they stop searching. As the pay-off of continuing to search in round t is higher than that (as the consumer actually continues to search as this yields a higher pay-off), they certainly want to continue searching if that higher continuation pay-off is taken into account.

With this in mind, let us then go back one period and consider that the consumer after observing an m in the first round either observes l or m in the next round.

(i) After observing l in search round two, the consumer believes that the last two objects are either LL, HL or HH and so overall, they believe the objects are either $HLHH$, $HLHL$, $HLLH$ or $HLLL$ and the ex ante total probability that one of these events happens is $\frac{1}{3}\cdot 4p_H^3p_L+\frac{2}{3}\cdot 6p_H^2p_L^2+4p_Hp_L^3$.(For example, the ex ante probability that three products are H and one is L is $4p_H^3p_L$ and out of these cases under the first algorithm the probability that the first objects ranked is an H and the next one is an L is $\frac{1}{3}$.) Thus, the conditional probability that the third object searched is an H equals $\frac{\frac{1}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2}{\frac{1}{3}\cdot 4p_H^3p_L+\frac{2}{3}\cdot 6p_H^2p_L^2+4p_Hp_L^3}$ and therefore the consumer prefers searching in round 3 even if they do not continue searching in round 4 if

$$\frac{\frac{1}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2}{\frac{1}{3}\cdot 4p_H^3p_L+\frac{2}{3}\cdot 6p_H^2p_L^2+4p_Hp_L^3}\frac{h-m}{2}>s.$$

This condition can be rewritten as $\frac{h-m}{s}>2\frac{2p_H^2+6p_Hp_L+6p_L^2}{2p_H^2+3p_Hp_L}=2\left(1+\frac{3p_Hp_L+6p_L^2}{2p_H^2+3p_Hp_L}\right)$. The RHS of this inequality is smaller than the RHS of (S.8) if $\frac{3p_H+6p_L}{2p_H+3p_L}<2$, which is clearly the case.

(ii) After observing m in search round two, the ex ante total probability of the first two objects being H equals $p_H^4+\frac{2}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2$. Thus, the conditional probability that the third object searched is an H equals $\frac{p_H^4+\frac{1}{3}\cdot 4p_H^3p_L}{p_H^4+\frac{2}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2}$ and therefore the consumer prefers searching in round 3 even if they do not continue searching in round 4 if

$$\frac{p_H^4+\frac{1}{3}\cdot 4p_H^3p_L}{p_H^4+\frac{2}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2}\frac{h-m}{2}>s,$$

which can be written as $\frac{h-m}{s} > 2\frac{3p_H^2+8p_Hp_L+6p_L^2}{3p_H^2+4p_Hp_L} = 2\left(1+\frac{4p_Hp_L+6p_L^2}{3p_H^2+4p_Hp_L}\right)$. The RHS of this inequality is smaller than the RHS of (S.8) if $\frac{4p_H+6p_L}{3p_H+4p_L} < 2$, which again is clearly the case.

So, now we go to the first search round where the consumer observes an m. The ex ante total probability of this state equals $p_H^4 + 4p_H^3p_L + 6p_H^2p_L^2 + 4p_Hp_L^3$. Thus, the conditional probability that the second object searched is an H equals $\frac{p_H^4+\frac{2}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2}{p_H^4+4p_H^3p_L+6p_H^2p_L^2+4p_Hp_L^3}$ and therefore the consumer prefers searching in round 2 even if they do not continue searching in round 3 if

$$\frac{p_H^4+\frac{2}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2}{p_H^4+4p_H^3p_L+6p_H^2p_L^2+4p_Hp_L^3}\frac{h-m}{2} > s,$$

which can be written as $\frac{h-m}{s} > 2\frac{p_H^4+4p_H^3p_L+6p_H^2p_L^2+4p_Hp_L^3}{p_H^4+\frac{2}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2} = 2\left(1+\frac{\frac{1}{3}\cdot 4p_H^3p_L+\frac{2}{3}\cdot 6p_H^2p_L^2+4p_Hp_L^3}{p_H^4+\frac{2}{3}\cdot 4p_H^3p_L+\frac{1}{3}\cdot 6p_H^2p_L^2}\right)$. The RHS of this inequality is smaller than the RHS of (S.8) if $\frac{\frac{1}{3}\cdot 4p_H^3+\frac{2}{3}\cdot 6p_H^2p_L+4p_Hp_L^2}{p_H^3+\frac{2}{3}\cdot 4p_H^2p_L+\frac{1}{3}\cdot 6p_Hp_L^2} < 2$, which again is clearly the case.

Thus, combining (S.5), (S.7) and (S.8), a sufficient condition for it to be possible that after observing m the consumer continues searching after the first algorithm, but not under the second is that

$$2\left(1+\frac{2p_L}{p_H}\right) < \frac{h-m}{s} < \min\left\{\frac{42p_H^3+152p_H^2p_L+96p_L^3}{21p_H^3+48p_H^2p_L}, 2+\frac{4p_L^2}{p_H^2}\right\}. \tag{S.9}$$

As the LHS and the RHS are function of p_H only (as $p_L = 1 - p_H$), it is possible to choose values of $h - m$, s such that both inequalities hold if there exist values of p_H such that the RHS is larger than the left hand side. Figure 3 draws $\frac{h-m}{s}$ on the vertical axis and p_H on the horizontal axis. The term of the LHS is represented by the black curve and the two terms of the RHS are represented by the green and red curve, respectively. It is clear that for (roughly) $p_H < 0.27$ the black curve is below the green curve and the green curve is the constraining factor on the RHS, reflecting the first term on the RHS. Thus, for the area in between the black and the green curve inequality (S.9) holds.

Finally, we consider the conditions under which it is optimal to stop searching under the second algorithm after observing an l on the first search. (It is clear it is optimal to stop searching after observing an l under the first algorithm). After observing an l under the

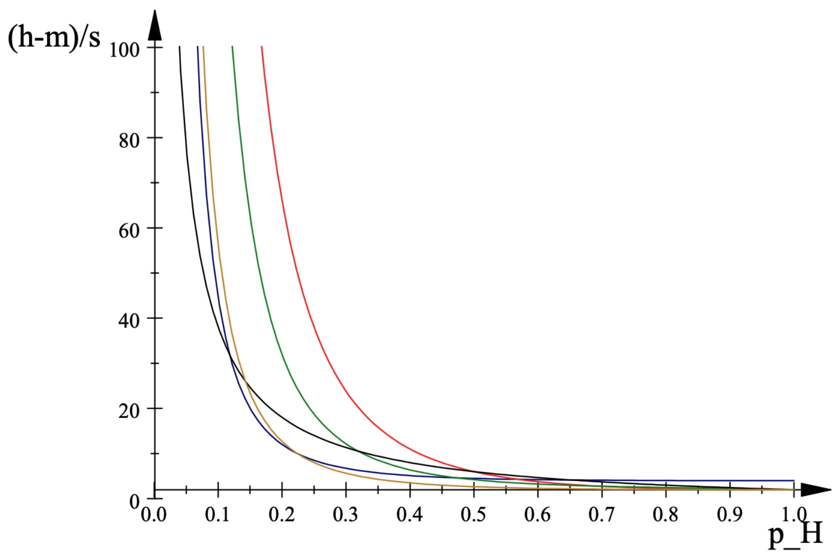


Figure 3

second algorithm, the consumer knows there are either two H products or all are L. It is clear that after observing three l the consumer stops searching. Also, after observing two L (which happens with a total ex ante probability of $\frac{1}{3} \cdot 6p_H^2 p_L^2 + p_L^4$) the consumer knows the continuation is either LL or HH and then the consumer stops searching if

$$\frac{2p_H^2}{2p_H^2 + p_L^2}\left((\frac{m+h}{2} - s) + \frac{1}{2}(\frac{h-m}{2} - s) \right) + \frac{p_L^2}{2p_H^2 + p_L^2}(l - s) < l,$$

or if $l \approx m$

$$\frac{(h-m)}{s} < 2 + \frac{2p_L^2}{3p_H^2}, \tag{S.10}$$

which is depicted by the beige curve in the figure.

Finally, after observing lm the consumer knows that there is still one L and one H option and it is then optimal to continue to search if the H option on the second search turned out to be an m if $\frac{1}{2}\left(\frac{1}{2}(h-m) - s\right) + \frac{1}{2}\left(\frac{1}{2}(h-m) - 2s\right) > 0$, which is the case if $\frac{h-m}{s} > 3$, which is clearly implied by the previous conditions in the figure. (The pay-off formula follows from the fact that with probability $\frac{1}{2}$ the next option is an H and then the consumer certainly stops searching afterwards, while with the remaining probability $\frac{1}{2}$ the next option is an L and then the consumer certainly continues searching).

Thus, after observing an L in the first position, the pay-off of continuing to search is given by

$$\frac{2p_H^2+p_L^2}{6p_H^2+p_L^2}(l-s)+\frac{4p_H^2}{6p_H^2+p_L^2}\left(\frac{1}{3}(l-s)+\frac{1}{3}(h-s)+\frac{1}{3}(m+\frac{1}{2}(h-m)-2\frac{1}{2}s)\right).$$

For $l \approx m$ this is smaller than l if (approx.)

$$\frac{4p_H^2}{6p_H^2+p_L^2}\left(\frac{1}{3}(m-s)+\frac{1}{3}(h-s)+\frac{1}{3}(\frac{1}{2}(m+h)-2\frac{1}{2}s)\right)<\frac{4p_H^2}{6p_H^2+p_L^2}m+\frac{2p_H^2+p_L^2}{6p_H^2+p_L^2}s,$$

or

$$\frac{4p_H^2}{6p_H^2+p_L^2}\left(-\frac{3}{2}s+\frac{1}{2}(h-m)\right)<\frac{2p_H^2+p_L^2}{6p_H^2+p_L^2}s,$$

or

$$\frac{h-m}{s}<3+\frac{p_L^2+2p_H^2}{2p_H^2}=4+\frac{p_L^2}{2p_H^2}, \tag{S.11}$$

which is depicted by the blue curve in the graph.

Thus, we can conclude that there is a parameter region where the stipulated search behavior under both algorithms is optimal at every stage, which is given by the black curve as the lower bound and the blue curve as the upper bound. In this region, the consumer always stops searching at the sponsored slot under the second algorithm but continues to search after observing an m in the sponsored slot under the first algorithm. It is clear from the analysis that we have provided sufficient conditions for this to be the case that are by no means necessary conditions. The total area of parameter values where the consumer always stops searching at the sponsored slot under the second algorithm but continues to search after observing an m in the sponsored slot under the first algorithm is larger than the area between the black and blue curves depicted in the figure.

S.3 Auxiliary Analysis for Heterogenous Prices

S.3.1 Auxiliary Analysis

We here provide the missing details underlying the argument in the proof of Proposition 3 (2). Specifically, we first show that if the platform runs an adjusted second-price auction placing the firm with the highest bid in the sponsored slot and breaking ties in favor of higher match scores, then the probability that a firm with price p^* wins the auction approaches 1.

Second, we argue that this auction also allows the platform to capture the entire rents of the winning firm, thus asymptotically achieving the upper bound of profits that can be earned form the sponsored slot.

Lemma S.2. *There exists an $\bar{n}$ such that firms with price p^* bid more in equilibrium than any firm with $p \in \mathcal{P} \setminus \{p^*\}$ and $h(p) > 0$ if $n \geq \bar{n}$.*

Proof. The expected profit difference from obtaining and not obtaining the sponsored slot is

$$(p)\left(1 - G(\gamma + p|\hat{\theta}^p(n, H)\right) - \frac{1}{n-1}\sum_{m=2}^{n} \pi\left(m, p, \hat{\theta}^{c,p}(n, H)\right) \tag{S.12}$$

where $\hat{\theta}^p(n, H)$ denotes the expected highest match score among all firms with the price (p), given the distribution H and n and where $\pi(m, \cdot)$ is the profit from position m. Clearly, firms will never bid more than (S.12). This upper bound increases in θ. Thus, the most that a firm with $p \neq p^*$ bids is

$$(p)\left(1 - G(\gamma + p|\bar{\theta}\right) - \frac{1}{n-1}\sum_{m=2}^{n} \pi\left(m, p, \bar{\theta}\right). \tag{S.13}$$

Note that the second term vanishes in the limit because the joint profits of all firms ares bounded. Then, by the properties of $G(\cdot)$ there is a $\theta^* < \bar{\theta}$ such that

$$(p)\left(1 - G(\gamma + p|\theta)\right) < (p^*)\left(1 - G(\gamma + p^*|\theta')\right) \tag{S.14}$$

for all $p \in \mathcal{P} \setminus \{p^*\}$ with $h(p) > 0$, $\theta \in \mathbf{\Theta}$ and $\theta' > \theta^*$. Since, by assumption

$$(p)\left(1 - G(\gamma + p|\bar{\theta})\right) < (p^*)\left(1 - G(\gamma + p^*|\bar{\theta})\right) \tag{S.15}$$

for all $p \in \mathcal{P} \setminus \{p^*\}$ with $h(p) > 0$, and because $(p)\left(1 - G(\gamma + p|\theta)\right)$ is continuous in θ, and $\mathcal{P}$ is finite, there is a $\theta^* < \bar{\theta}$ such that

$$(p)\left(1 - G(\gamma + p|\bar{\theta})\right) < (p^*)\left(1 - G(\gamma + p^*|\theta'\right) \tag{S.16}$$

for all $p \in \mathcal{P} \setminus \{p^*\}$ with $h(p) > 0$, $\theta \in \mathbf{\Theta}$ and $\theta' > \theta^*$. Since the left-hand side of the inequality above is monotonically increasing in θ, inequality (S.14) follows.

Since $\hat{\theta}^{p^*}(n, H)$ increases in n, there is an $\bar{n}$ large enough so that $\hat{\theta}^{p^*}(n, H) > \theta^*$. Considering (S.12) and noting that the second term vanishes, we conclude that the upper bound of what $(p^*)-$ types are willing to bid exceeds that of other types if $n \geq \bar{n}$. Firms with p^* will never bid less than any firm with $p \neq p^*$ in any symmetric equilibrium because if not, then every firm with type p^* would find it profitable to raise its bid. $\square$

Lemma S.3. *In the limit, the platform earns* $(p^*)\left(1 - G(\bar{r} + p^*|\bar{\theta})\right)$ *a.s.*

Proof of Lemma S.3. As $n \to \infty$, the probability that there are at least two firms with cost-price pair (c^*, p^*) converges to 1. By Lemma S.2, these firms with (c^*, p^*) compete only against each other asymptotically almost surely. Notably, all firms with a cost price pair (c^*, p^*) are ex ante identical as in our base model. The same argument as in the proof of Theorem 1 then guarantees an equilibrium bid to exist, which in the limit converges to $(p^*)\left(1 - G(\gamma + p^*|\bar{\theta})\right)$. Since there are at least 2 bidders a. s., platform profits equal this bid. This is equal to the upper bound on the platform's profits, implying asymptotic optimally. Naturally, the match score of the winning firm approaches $\bar{\theta}$ as $n \to \infty$. $\square$