Search, Learning and Tracking

Marcel Preuss*

September 2019

Abstract

Oftentimes search across sellers helps consumers not only to find a good price and product that matches their needs, but also to learn about their willingness-to-pay for a match. In this case I show that tracking the number of search attempts is informative about a consumer's type and, in turn, influences prices. Search history-based price discrimination reduces profits and welfare when search costs are low under mild conditions, yet it is the unique equilibrium outcome when tracking is feasible. This remains true even if consumers can opt out of tracking, or if sellers can publicly commit to a tracking policy.

Keywords: Tracking, Consumer Search, Price-Discrimination, Privacy

Using current technology the vast majority of online sellers cannot access a user's online search history on their own (see Bujlow, Carela-Español, Sole-Pareta, and Barlet-Ros, 2017, for a recent survey). Yet third-party tracking networks can make such information available to sellers, and are becoming increasingly pervasive.[1] Thus far, however, little is known about the effects of personalized pricing in a dynamic equilibrium search environment, even though this practice may become one of the main uses for search history data.

*SC Johnson Graduate School of Management, Cornell University, Sage Hall, Ithaca, NY, email: preuss@cornell.edu. I am very grateful to my former advisor Martin Peitz for his guidance while working on this project. I also want to thank Robert Crowell, Justin Johnson, Stefan Lauermann, Espen Moen, Andrew Rhodes, Sandro Sheliga, Michael Waldman, Chengsi Wang, Chris Wilson and Asher Wolinksy for helpful comments. Financial support from the Deutsche Forschungsgemeinschaft (CRC TR 224) is gratefully acknowledged.

[1] Recent work by Li, Hang, Faloutsos, and Efstathopoulos (2015) shows that around 46% of the 10,000 most popular websites according to the Alexa rankings are monitored by at least one third-party tracker, and that Google monitors 25% of them.

In this paper, I address the question how the ability of sellers to track search histories of consumers, who search sequentially to learn about prices and product suitability, affects search behavior, equilibrium prices and welfare. A key observation is that if the willingness-to-pay for a suitable product, called a match, varies across consumers, then the length of a consumer's search history can be informative about her willingness-to-pay for a match. To this end, I propose a sequential search model where consumers are ex ante heterogeneous and sellers are able to set search history-contingent prices.

The first major finding is that in the unique equilibrium tracking leads to prices which increase with the number of search attempts. This result follows from two novel assumptions. The first is that prior to search, consumers are uncertain about their willingness-to-pay for any product both because they do not know whether it is a match and because they do not know how much they value a match. The second is that search serves a twofold role that can resolve each type of uncertainty, sometimes independently. More precisely, while consumers always learn their willingness-to-pay for a match if they encounter one, I assume that they sometimes do learn it even if the sampled product is not a match.

The idea that we often do not know exactly what we want prior to searching appears uncontroversial. For example, Kahneman and Thaler (2006) conclude that "people do not always know what they like".[2] Accounting for the possibility that search resolves this uncertainty is one of my contributions.

As an illustration imagine a consumer who wants to buy alloy wheels. Most likely, several designs exist that she will like. However, without experience, she barely knows what these designs look like and therefore cannot say how much she values them. As is standard for search goods, sampling and inspecting a product helps her find out how much she values it. If it is a match, she also learns how much she values a suitable design in general. Moreover, even sampling a design of the wrong color, or wrong size, say, may help her realize what design she actually likes and what she would be willing to pay for it. This match-independent source of learning about her willingness-to-pay for a match is present in my model.

[2]Dzyabura and Hauser (2019) have recently modeled learning about preferences to study optimal recommendation systems in a non-equilibrium search model. Santos, Hortaçsu, and Wildenbeest (2017) emphasize the role of learning about the utility distribution while searching as well and provide a method to estimate search costs in this case.

To develop an intuition for why prices increase, observe that because search is costly, consumers who are informed about their willingness-to-pay for a match continue searching only if it is strictly larger than some of the remaining sellers' expected prices. This implies the existence of a cut-off type. Additionally, search with learning makes consumers more likely to be informed, the more search attempts they have already made. Thus, if a consumer has a longer search history and decides to continue searching, that decision is a stronger signal that her willingness-to-pay exceeds the cut-off type than if she has a shorter search history.

As in Diamond (1971), sellers whose expected price lies below the cut-off type might want to hold up informed consumers by setting a price above what these expect. However, this cannot happen under rational expectations. Hence, when consumers are more probable to be informed, a mechanism that counteracts the increasing incentive to hold up consumers is necessary to restore an equilibrium. An increasing price path achieves exactly this because a higher expected price lowers the probability that an informed consumer continues searching. Thus, counterintuitively, sellers hardly benefit from being able to predict the willingness-to-pay of consumers with long search histories. Instead, the equilibrium price is so high that barely any informed consumer samples another seller after too many search attempts.

This intuition underpins the second main finding. Search history-based pricing lowers both industry profits and welfare when search costs are low. Yet no seller wants to disregard search histories and charge a uniform price in equilibrium. Interestingly, the effect on consumers is ambiguous. However, tracking can also raise profits and total surplus when search costs are high as shown in Section 5. In short, this is because search may break down completely when search costs are high unless sellers are able to discriminate between different search histories.

Informed by models of first and third-degree price discrimination, where consumer data are exogenously given, personalized pricing is often seen as welfare-enhancing.[3] My results thus suggest that one must be cautious when applying

[3] For instance, the contribution by the US to the joint meeting between the OECD Competition Committee and the Committee on Consumer Policy on Nov 28, 2018, acknowledges that "to the extent that personalized pricing allows firms to engage in something resembling perfect first-degree price discrimination [...], it will tend to improve total welfare [...and] may raise consumer welfare by enhancing competition" (see note DAF/COMP/WD(2018)140). Indeed,

these models to an environment where consumers anticipate that their search behavior can affect the price of the products they search.

In the EU, the sale of personal browsing data is restricted by the General Data Protection Regulation ("GDPR") and usually requires consent by the user.[4] In contrast, the US Congress has taken a different approach, allowing even Internet service providers, who can track their customers' complete browsing activity, to use the data commercially.[5] These opposite solutions suggest that more theoretical guidance is needed in dealing with tracking and its consequences. I therefore analyze three policy-relevant extensions in Section 6. These are (1) granting consumers the right to block tracking at any time, (2) enabling sellers to commit to not tracking and (3) introducing a platform that controls access to search data.

I find that consumers would never opt out of tracking as they prefer to distinguish themselves from consumers with longer search histories, similarly to the unravelling argument of Milgrom and Roberts (1986). Analogously, consumers with short search histories prefer sampling sellers who have committed to tracking. As a result, tracking sellers are sampled earlier, receive more demand, and make more profits, implying that all sellers want to track in equilibrium.

Interestingly, an intermediary platform is able to internalize these negative demand externalities and benefits from restricting all sellers' access to search data. Since this not only raises profits but also welfare, it implies a so-far neglected advantage of dominant online marketplaces. In light of the current investigation by the European Commission into whether Amazon's data sharing policy regarding third-party sellers is too restrictive and prevents sellers from being more competitive, this seems particularly relevant.[6]

---

search history-based pricing in my model resembles first-degree price discrimination to the extent that conditional on actively searching, consumers with long search histories do have a higher expected willingness-to-pay.

[4] See the Regulation (EU) 2016/679 of the European Parliament and of the Council.

[5] Marshall, Jack. 2017. "With Washington's Blessing, Telecom Giants Can Mine Your Web History". *The Wall Street Journal,* March 30. https://www.wsj.com/articles/with-washingtons-blessing-telecom-giants-can-mine-your-web-history-1490869801

[6] Schechner, Sam. 2019. "Amazon Faces Probe in Europe Over Use of Merchant Data". *The Wall Street Journal,* July 17. https://www.wsj.com/articles/amazon-to-overhaul-marketplace-terms-as-part-of-german-settlement-11563353036

1 Literature

Since tracking lets sellers price discriminate based on consumer behavior, this article relates to the literature on behavior-based price discrimination (BBPD). In the standard BBPD framework consumers make a purchase in multiple periods and sellers have a way to recognize old customers. Extending the early work by Hart and Tirole (1988) on BBPD by a monopolist, Fudenberg and Tirole (2000) and Villas-Boas (1999) analyze the case of competing sellers. Despite the modeling differences between their work and mine, they also find that firms are in many cases better off if they are not able to track their customers.

Few papers explicitly address tracking in search markets. A notable exception is Armstrong and Zhou (2016). They find that a seller seeks to deter search by charging higher prices from returning consumers when it is possible to observe whether a consumer is visiting for the first time, which differs from how I model tracking.

Nevertheless, my setting is closely related to models of ordered search, in particular to those where all consumers search in the same order, which is as if sellers can track consumers. Arbatskaya (2007) has made an early contribution to this stream of research, showing that search cost heterogeneity leads to prices which increase in the order of search in a market where products are homogeneous.

More closely related to this paper, Zhou (2011) presents a search model based on the framework of Wolinsky (1986) and Anderson and Renault (1999), where consumers search for both price and product fit. Because consumers never value two products equally in this model, they may always encounter better matches at other sellers. This is why he finds that, when consumers have sampled most sellers, the remaining sellers enjoy greater monopoly power and, thus, set higher prices. This insight continues to hold when search is only partially ordered as in models of prominence (see Armstrong, Vickers, and Zhou, 2009).[7]

The papers above have in common that consumers' valuations for all products are drawn from a common distribution, which is known ex ante.[8] Instead, I focus

[7]For instance, Haan and Moraga-González (2011) and Moraga-González and Petrikaitė (2013) have studied how advertising or merging with a competitor can make a subset of sellers more salient, who are then visited first and find it profitable to set lower prices than their competitors.

[8]Some have focused on other types of consumer heterogeneity. For instance, Armstrong et al.

on consumers who draw their valuations from heterogeneous (dichotomous) distributions, which are a priori unknown. Accounting for this type of consumer heterogeneity and learning renders a seller's position informative about a consumer's valuation for a suitable product, which is absent in those models.[9]

In many instances, consumers may search through their options in a deliberate order because they have prior information about their valuations or prices (see Armstrong, 2017, for an extensive discussion). A closely related example of such a market is the duopoly model by Anderson and Renault (2000), who assume that a fraction of consumers is informed about all their match values before search and, thus, sample their preferred product first.[10] Importantly, prior information about idiosyncratic match values implies that consumers are ex ante heterogeneous as in this paper. Contrary to the assumption about tracking, however, it also implies that search orders differ between consumers, preventing sellers from knowing their position in a consumer's search process.

A notable exception is Rhodes (2011), who analyzes the effect of prominence on profits in search markets. In his model, consumers know their valuation for all products but not where to find them. Since the prominent firm is sampled first in equilibrium, Rhodes (2011) essentially combines ex ante heterogeneous consumers with sellers who have an imperfect tracking device. This makes his setting somewhat similar to mine, albeit several major differences exist. First, I consider a perfect tracking device which informs sellers about their exact position in a consumer's search process. Second, his finding that the prominent firm sets a lower price is due to a different mechanism.[11] Third, I additionally conduct a

(2009) consider search cost heterogeneity in Section 4 of their paper, which can reverse the result obtained for homogeneous search costs.

[9]In contrast, sellers can infer a consumer's reservation value from their position in Zhou (2011), which depends on the number of remaining sellers. This is why, when the number of sellers becomes large, reservation values vary little and the difference between ordered and random search vanishes. In contrast, tracking always changes the equilibrium outcome of my model, independently of the number of sellers.

[10]Notice that informed consumers in my model are different as they still do not know what seller to visit for a match. Besides, the availability of match value information prior to search is also a key assumption in Haan, Moraga-González, and Petrikaitė (2018) and in Choi, Dai, and Kim (2018).

[11]In his model, non-prominent firms know that a consumer who continues searching beyond the prominent firm has not yet found her most-preferred product, which is why non-prominent firms always have much more market power. In contrast, there is no change in market power

welfare analysis of tracking and alternative data access regimes.

Moreover, this article is one of the few that explicitly embeds learning about future match values into an equilibrium search model.[12] Another recent contribution in this area is Mauring (2017), who studies buyers who can learn about the distribution of match values by observing whether a seller has traded in the past. However, sellers are non-strategic in her model and do not price discriminate.

Finally, this work is related to the research on consumer privacy, which often builds on models of BBPD. Taylor (2004) considers the possibility that sellers can share a consumer's purchase history. He finds that privacy protection policies are necessary only if consumers are naive, but not so otherwise. Acquisti and Varian (2005) reach similar conclusions.[13] Contrary to these results, my analysis suggests that even if consumers can maintain privacy at no cost and if giving it up reduces their surplus, they will likely disclose their search history to sellers.

2 Search with Tracking

There is a unit mass of consumers and N horizontally differentiated single-product sellers, with $\infty > N > 1$. To learn a product's price and their valuation for it, consumers must sample a seller at a cost $s > 0$. Each seller supplies its product at constant marginal cost, which is normalized to zero. Besides, prices can be conditioned on the number of sellers that a consumer has previously sampled. That is, the price offered to a consumer is set after the consumer incurs the search cost to sample a particular product. Notably, the possibility of tracking does not imply price discrimination per se as sellers could still ignore search history information and stick to a single price.

across positions in my model.

[12] When using the Hotelling or the spokes model of Chen and Riordan (2007) to model preferences, valuations are negatively correlated and learning about the value of one option comes as a by-product of sampling another (e.g. as in Armstrong and Zhou, 2011; Rhodes, 2011). Adam (2001) has also studied directed search with learning about the options' utility distributions but without endogenous prices.

[13] In a related setting where consumers can block tracking, Conitzer, Taylor, and Wagman (2012) find that they may even do so too often from a consumer surplus standpoint.

2.1 Consumers

Consumers have unit demand and sample products in random order and with free recall. That is, sellers commit to a single price per consumer and each consumer receives at most N different offers. In what follows, I will often refer to "seller k" as the kth seller a particular consumer samples. Consequently, seller k is not the same for all consumers. After sampling a product, consumers can buy the product, buy from a previous seller, continue searching, or stop searching. If they stop without buying anything, they receive a payoff of zero. Alternatively, a consumer obtains utility $u_k = v_k - p_k$ if she purchases from seller $k \leq N$, where p_k is the price and v_k her seller-specific product valuation. Valuations for sellers' products are independently and identically distributed according to

$$v_k = \begin{cases} \theta & \text{with probabilty } m \\ 0 & \text{with probability } 1 - m, \end{cases}$$

where $m \in (0, 1)$. I refer to the event $v_k = \theta$ as a match at seller k and to m as the match probability. Thus, θ represents a preference parameter that equals a consumer's valuation conditional on a match. It differs across consumers and is distributed according to $\theta \sim F$ with density f and support $[0, \bar{v}]$. I make the following standard assumption regarding f.[14]

Assumption 1 *The density f is log-concave and twice differentiable.*

In addition, I assume that consumers know $\theta \sim F$ ex ante, but not the realization of their individual type θ. Crucially, there are two ways for consumers to learn θ. Firstly, by finding a match, and secondly, by receiving a perfectly informative signal $\sigma = \theta$ following a search attempt. This signal is realized with probability $\lambda > 0$, independently of a consumer's type. Thus,

$$\mu_k = 1 - (1 - \lambda)^{k-1} \tag{1}$$

is the probability that a consumer has received σ after $k - 1$ search attempts.

[14] I maintain Assumption 1 for the remainder of this paper. Any restriction on f stated in one of the forthcoming results constitutes an additional assumption and not a replacement of it.

2.2 Equilibrium Analysis

I use Perfect Bayesian Nash Equilibrium (PBE). This entails that consumers employ a possibly non-stationary search rule given the already sampled, as well as all expected future prices. Sellers set profit-maximizing prices conditional on the observed search history, the expected search rule and their beliefs about a consumer's type. In equilibrium, the prices consumers expect must be consistent with the sellers' pricing strategy, and each seller's belief about a consumer's type must be consistent with the equilibrium search rule. Moreover, I assume that consumers have passive beliefs. That is they do not adjust their expectations about forthcoming sellers' prices when they observe an out-of equilibrium price. Finally, I restrict attention to symmetric equilibria.

To construct an equilibrium, I begin by conjecturing that consumers expect prices to increase the longer they search. I will then show that there can be no other equilibrium. Hence, suppose that

$$p_1^e \leq p_2^e \leq ...p_N^e, \tag{2}$$

where the superscript indicates expected prices.

The optimal stopping behavior is characterized by a simple cut-off rule. Provided that seller k sets a price $p_k \leq p_{k+1}^e + s/m$, a consumer who has found a match at seller k buys if and only if $\theta \geq p_k$ and stops without making a purchase otherwise.[15] This is due to two reasons. First, she cannot expect $v_l > v_k$ at any seller $l > k$. Second, expected price savings do not justify incurring the search cost again as $m(p_l^e - p_k) \leq s$ for all $l > k$.

To summarize, consumers always stop searching after they find a match and otherwise continue searching if the continuation value from search is non-negative. The continuation value, in turn, depends on whether consumers are informed.

Informed consumers. Let $\hat{V}_k(\theta)$ be an informed consumer's expected value of continuing search when her type equals θ and she has not encountered a match before. Then, she samples seller k if and only if $\hat{V}_k(\theta) \geq 0$, which holds if and only if $m(\theta - p_k^e) \geq s$ as $p_l^e \geq p_k^e$ for all $l > k$. Let $K_\theta = \max\left\{k \leq N : m(\theta - p_k^e) \geq s\right\}$

[15] In the proof of Theorem 1, I show that deviating to a price strictly above $p_{k+1}^e + s/m$ is not profitable.

denote the endogenous number of sellers after which an informed consumer of type θ always stops searching. Explicitly, $\hat{V}_k(\theta)$ can be written as

$$\hat{V}_k(\theta) = \sum_{j=k}^{K_\theta}(1-m)^{j-k}\big(m(\theta - p_j^e) - s\big). \tag{3}$$

Uninformed consumers. An uninformed consumer additionally takes into account the expected option value of learning her type after future search attempts. Recall that (3) represents the continuation value of an informed consumer without a match. Since the joint probability of learning θ but not finding a match is $(1-m)\lambda$, the option value of becoming informed when sampling seller k equals

$$(1-m)\lambda \int_0^{\bar{v}} \hat{V}_{k+1}(\theta) f(\theta)\mathrm{d}\theta. \tag{4}$$

Adding to (4) the expected surplus from sampling the next seller k yields an uninformed consumer's continuation value of sampling k, given by

$$\begin{aligned} V_k = {} & m \int_{p_k^e}^{\bar{v}} (\theta - p_k^e) f(\theta)\mathrm{d}\theta - s \\ & + (1-m)\left(\lambda \int_0^{\bar{v}} \hat{V}_{k+1}(\theta) f(\theta)\mathrm{d}\theta + (1-\lambda)\max\big\{V_{k+1}, 0\big\}\right). \end{aligned} \tag{5}$$

The search rule when expectations satisfy (2) can be summarized as follows.

Lemma 1 *If $p_k \le p_{k+1}^e + s/m$, consumers use the following search rule:*

1. *If $v_k = \theta$, buy from seller k if and only if $\theta \ge p_k$, otherwise stop searching.*
2. *If $v_k = 0$ and θ is known, continue searching if and only if $\theta \ge p_{k+1}^e + \frac{s}{m}$.*
3. *If $v_k = 0$ and θ is unknown, continue searching if and only if $V_{k+1} \ge 0$.*

Let $\mathcal{R}^*$ denote this search rule. By Lemma 1, an informed consumer without a match samples seller k with probability $1 - F(p_k^e + s/m)$. Thus, conditional on facing an informed consumer who follows $\mathcal{R}^*$, the probability of a sale equals

$$m \cdot \mathbb{P}\left(\theta \ge p_k \middle| \theta \ge p_k^e + s/m\right) = m \cdot \min\left\{1, \frac{1 - F(p_k)}{1 - F(p_k^e + s/m)}\right\}. \tag{6}$$

In contrast, uninformed consumers either sample seller k with probability one or not at all, depending only on $V_k \geq 0$. Consequently, the probability of selling to a searching and uninformed consumer as seller k is $m \cdot \mathbb{P}(\theta \geq p_k) = m\big(1 - F(p_k)\big)$.

To derive demand, it remains to calculate the share of informed consumers who have learned θ prior to sampling seller k. According to the optimal search rule $\mathcal{R}^*$, only consumers without a match continue searching. Thus, the probability that a consumer has learned θ before sampling k is given by the probability that she has received σ, which equals μ_k as in (1). After using some algebra and accounting for both informed and uninformed consumers, expected demand in position k can be written as

$$D_k(p_k, p_k^e) = (1 - \mu_k)m\big[1 - F(p_k)\big] + \mu_k m\big[1 - F\big(\max\{p_k^e + \frac{s}{m}, p_k\}\big)\big]. \tag{7}$$

An additional qualifier is needed if all uninformed consumers stop searching before sampling N sellers. Denote by $\mathcal{K}$ the equilibrium search persistence of uninformed consumers, that is the number of search attempts after which an uninformed consumer that has not encountered a match stops searching. Search persistence is characterized by (i) $V_k \geq 0$ for all $k \leq \mathcal{K}$ and (ii) $V_{\mathcal{K}+1} < 0$. Because increasing expected prices imply that $V_l \geq 0$ if $V_k \geq 0$ for all $l < k$, (i) simplifies to $V_{\mathcal{K}} \geq 0$. Moreover, since $\int_{p_k^e}^{\bar{v}} (\theta - p_k^e) f(\theta) \mathrm{d}\theta$ and $\int_0^{\bar{v}} \hat{V}_{k+1}(\theta) f(\theta) \mathrm{d}\theta$ are decreasing in k, V_k is decreasing in k. Thus, for each sequence of expected prices satisfying (2), $\mathcal{K}$ is uniquely determined by

$$\mathcal{K} = \max\big\{k \leq N | V_k \geq 0\big\}. \tag{8}$$

That is, demand is characterized by (7) only if $k \leq \mathcal{K}$. Knowing that uninformed consumers do not search more than $\mathcal{K}$ times, seller $k > \mathcal{K}$ could profitably exploit informed consumers by deviating to a price $p_k \in (p_k^e, p_k^e + s/m)$, which cannot be an equilibrium. Hence, consumers expect $p_k^e \geq \bar{v} - s/m$ if $k > \mathcal{K}$ and at most a zero mass of informed consumers sample more than $\mathcal{K}$ sellers in any equilibrium.[16]

[16] If $p_k^e > \bar{v} - s/m$, this equilibrium relies on sellers believing with probability one that they face an informed consumer if they observe that they are in position $k > \mathcal{K}$, which can happen only off path. It is possible to relax this assumption on sellers' beliefs through a minor modification of the model. If there is an arbitrarily small but strictly positive mass of uninformed consumers who always continue searching unless they find a match, seller $k > \mathcal{K}$ strictly prefers a price close

As is standard in search models, there exist equilibria in which consumers expect the prices of sellers $k \geq K$ for any $K \geq 1$ to be that high that all uninformed consumers stop searching before seller K. I explicitly exclude those equilibria, here and in the following. That is, suppose that if uninformed consumers sampled K, seller Kth best-response to that would be a price at which $V_K \geq 0$. In this case I restrict attention to equilibria where $\mathcal{K} \geq K$.

Since sellers can make only one offer per consumer, their problem as seller $k \leq \mathcal{K}$ is independent of that as seller $k' \neq k$. Hence, seller k chooses p_k to maximize

$$p_k D_k(p_k, p_k^e). \tag{9}$$

While I analyze all cases in the appendix, I focus on $p_k < p_k^e + s/m$ here. Taking the derivative with respect to p_k, the first-order condition is

$$p_k = \frac{1 - F(p_k)}{f(p_k)} + \frac{\mu_k}{1 - \mu_k} \frac{1 - F(p_k^e + s/m)}{f(p_k)}. \tag{10}$$

Notice that p_k is decreasing in p_k^e. This is intuitive since the higher the price consumers expect, the more informed consumers do not sample seller k, which reduces the share of inelastic demand. Further, (10) implies that sellers would like consumers to expect a low price only to hold them up by setting a higher price after consumers have incurred the search cost. However, expectations must be correct in equilibrium and therefore $p_k^e = p_k$.

Log-concavity of f implies that (10) has a unique fixed point, i.e. a solution with $p_k^e = p_k$. Let p_k^+ denote this fixed point price. Then, uniqueness of p_k^+ ensures that if there is an equilibrium with increasing prices, it is unique in satisfying this property. Notably, I can also show that no alternative equilibrium exists. Yet, a unique fixed point alone does not guarantee equilibrium existence. To rule out profitable deviations, the right-hand side of (10) must be decreasing in p_k on $[0, p_k^e + s/m)$ for a given p_k^e. This holds if f is non-decreasing or if λ is sufficiently small, as shown in Lemma A.6. Precisely, for every f, there is a value $\lambda^* > 0$ such

to $p_k = \bar{v} - s/m$ over any other price and (almost) no consumer samples any seller $k > \mathcal{K}$ as in the model presented above.

that any deviation from $p_k^e = p_k^+$ is not profitable if $\lambda \leq \lambda^*$.[17]

Theorem 1 *If f is non-decreasing or if λ is sufficiently small, there is a unique equilibrium in which prices satisfy $p_k^+ < p_{k+1}^+$ for all $k < \mathcal{K}$.*

Though the property of rising prices is reminiscent of the famous hold-up problem in search markets, one must be careful with interpreting it. Counterintuitively, prices are not increasing in k because sellers hold up an increasing share of informed consumers whose demand is locally inelastic. This is because in order to do so, sellers would need to set prices above the price consumers expect, which is ruled out in equilibrium. Instead, the equilibrium price p_k^+ must be high enough such that seller k has no incentive to exploit the remaining probability of facing an informed consumer at the risk of selling less to an uninformed one. Thus, the more likely a consumer has learned θ, the more often she must be kept from searching, which is why p_k^+ increases in k. As a result, the probability of actually facing an informed consumer barely increases in k.

To understand intuitively why prices cannot be anything but increasing in k, suppose that consumers expect a sequence of prices which features $p_{k-1}^e > p_k^e$ for some k. Consider the last seller k whose expected price satisfies $p_k^e < p_{k-1}^e$ and call him seller j. Of course, if consumers have rational expectations, seller j and $j-1$ must set their prices equal to p_j^e and p_{j-1}^e in equilibrium. Suppose that the expected price difference is small. Then, consumers with a match will always buy from $j-1$ or stop searching and only consumers without a match at $j-1$ sample seller j. Of the latter group, a higher share will be informed after seller $j-1$ has been sampled due the reception of σ. Moreover, $p_j^e < p_{j-1}^e$ implies that conditional on being informed, consumers are more likely to sample seller j than seller $j-1$. Consequently, seller j faces a higher share of informed consumers than seller $j-1$ and, other things equal, has a larger incentive to hold up consumers. Because equilibrium prices must offset these incentives, $p_j^e < p_{j-1}^e$ cannot hold.

Finally, if the expected price difference is large, the share of informed consumers sampling j increases further as consumers who become informed through finding

[17] As shown in the appendix, λ^* can be derived from the maximum value of λ that ensures that seller $\mathcal{K}$ has no incentive to deviate. Take the logit-normal distribution with mean 1/2 and $\sigma = 1/2$ as an example and assume that $N = 5$. Then, if $p_5^+ + s/m \leq 0.73$, the equilibrium exists for all $\lambda \leq \lambda^* \approx 4\%$.

a match at $j-1$ will continue searching as well. This implies an even higher share of inelastic demand for seller j, leading to the same contradiction as above.

3 Banning Tracking

The analysis above shows that in an environment in which ex ante uninformed consumers learn while searching, sellers ask for higher prices from consumers who have been searching for longer. In this section, I examine how banning tracking affects the equilibrium.

3.1 Equilibrium Analysis

Without tracking, sellers cannot make their prices contingent on consumers' search histories. Hence, consumers expect all sellers to set the same price in a symmetric equilibrium, i.e. $p_k^e = p_{k+1}^e = p^e$ for all $k = 1, \ldots, N$. By Lemma 1, the stopping rule $\mathcal{R}^*$ thus remains optimal in any such equilibrium.

As in the previous section, denote by $\mathcal{K}$ the maximum number of sellers that an uninformed consumer is willing to sample before she stops searching. As opposed to random search in Wolinsky (1986) and Anderson and Renault (1999), it is a real possibility in my model that a consumer stops searching without making a purchase and before exploring all options. This is because the continuation value of uninformed consumers depends on the option value of learning θ, which is increasing with the number of remaining sellers. Thus, even though consumers expect the same price at all sellers in equilibrium, the stopping rule of uninformed consumers is non-stationary and $V_1 \geq 0$ does not imply $V_N \geq 0$.[18]

Because consumers follow $\mathcal{R}^*$, they always stop searching after encountering a match. Thus, consumers sampling seller k are informed about their type only if they have observed the independent signal σ at least once before, which happens with probability μ_k. With probability $(1-\lambda)^{\mathcal{K}}$ however, a consumer does not receive σ during her first $\mathcal{K}$ search attempts and stops searching after seller $\mathcal{K}$.

[18]Besides, in contrast to search with tracking analyzed in the previous section, informed consumers may sample more than $\mathcal{K}$ sellers because sellers cannot observe when they are in position $k > \mathcal{K}$.

Besides, notice that $p_k^e = p_l^e = p^e$ guarantees that if sampling seller k is worthwhile for an informed consumer, sampling any seller $l < k$ is too. Finally, the probability of not encountering a match before seller k conditional on searching is $(1-m)^{k-1}$. Thus, the mass of informed consumers sampling k equals

$$\begin{array}{ll}(1-m)^{k-1}\left[1-F\left(p^e+s/m\right)\right]\mu_k & \text{if } k \leq \mathcal{K} \\ (1-m)^{k-1}\left[1-F\left(p^e+s/m\right)\right]\left(1-(1-\lambda)^{\mathcal{K}}\right) & \text{if } k > \mathcal{K},\end{array} \tag{11}$$

while the probability of selling to an informed consumer conditional on a match is given by $\mathbb{P}\left(\theta \geq p | \theta \geq p^e + s/m\right)$, where p is the actual price.

Besides, the mass of uninformed consumers sampling k equals $(1-m)^{k-1}(1-\mu_k)$ if $k \leq \mathcal{K}$ and zero otherwise. Since it is optimal to buy immediately conditional on a match if both $\theta \geq p$ and $p \leq p^e + s/m$ hold, the probability of selling to an uninformed consumer is $1 - F(p)$. Then, demand weighted over all positions $k = 1, ..., N$ can be written as

$$\begin{aligned}D(p,p^e) = \frac{m}{N}\Bigg\{ & \sum_{k=1}^{\mathcal{K}}(1-m)^{k-1}\Big((1-\mu_k)\left[1-F(p)\right] + \left[1-F\left(p^e+s/m\right)\right]\mu_k\Big) \\ & + \sum_{k=\mathcal{K}+1}^{N}(1-m)^{k-1}\left[1-F\left(p^e+s/m\right)\right]\left(1-(1-\lambda)^{\mathcal{K}}\right)\Bigg\},\end{aligned} \tag{12}$$

if $p \leq p^e + s/m$.[19] Next, recall that the equilibrium value of $\mathcal{K}$ is determined by the value function of uninformed consumers as follows. The continuation value from sampling any seller $k \leq \mathcal{K}$ is non-negative and the one from sampling seller $\mathcal{K}+1$ is strictly negative. The first condition simplifies to $V_{\mathcal{K}} \geq 0$ because all sellers charge the same price and because the option value of learning θ is decreasing in k. Thus, as before, there is a unique $\mathcal{K}$ for every p^e. Without tracking, however, $\mathcal{K}$ and the equilibrium price, denoted by $\tilde{p}(\mathcal{K})$, are determined jointly. That is, $\mathcal{K}$ has to satisfy

$$\mathcal{K} \in E := \left\{k \leq N : \; V_k\big(p^e(k)\big) \geq 0 \text{ and } V_{k+1}\big(p^e(k)\big) < 0\right\}, \tag{13}$$

[19]The case of deviating to $p > p^e + s/m$ is dealt with in the proof of Theorem 2.

where the argument $p^e(k)$ in $V_k(\cdot)$ denotes the consumers' expected price when $\mathcal{K} = k$, which has to equal $\tilde{p}(k)$ under rational expectations. Since $\tilde{p}(\mathcal{K})$ is discontinuous in $\mathcal{K}$, an equilibrium with $\mathcal{K} > 0$ need not always exist.[20] To this end, I provide a sufficient condition for the existence of an equilibrium in which $\mathcal{K} = N$ at the end of this section.

Moreover, multiple equilibria with different levels of search persistence may exist since $\tilde{p}(\mathcal{K})$ is decreasing in $\mathcal{K}$ if $m \leq 0.25$ and $\lambda \leq 0.5$, which seem to be reasonable assumptions (see Lemma A.10 in the Appendix).[21] Though seemingly different, this is similar to the equilibrium multiplicity in the random search framework by Wolinsky (1986) and Anderson and Renault (1999). There, two equilibria usually exist. One where consumers always continue sampling one of the remaining sellers if they do not make a purchase and one where nobody searches. These equilibria correspond to $\mathcal{K} = N$ and $\mathcal{K} = 0$. In addition, any $1 \leq \mathcal{K} < N$ may constitute an equilibrium here, where a lower $\mathcal{K}$ is associated with consumers sampling fewer sellers on average.

Crucially, if $N \in E$, other equilibria with $\mathcal{K} < N$ exist only because sellers and consumers may be pessimistic about $\mathcal{K}$ and $\tilde{p}(\mathcal{K})$, as they are in the equilibrium where nobody searches. As before, I therefore do not consider those equilibria and restrict attention to the equilibrium with the highest $\mathcal{K}$.[22] Additionally, define

$$\phi_k(\mathcal{K}) := \frac{(1-m)^{k-1}\big[\mathbb{1}_{k\leq\mathcal{K}} + \big(1-(1-\lambda)^{\mathcal{K}}\big)\mathbb{1}_{k>\mathcal{K}}\big]}{\sum_{j=1}^{N}(1-m)^{j-1}\big[\mathbb{1}_{j\leq\mathcal{K}} + \big(1-(1-\lambda)^{\mathcal{K}}\big)\mathbb{1}_{j>\mathcal{K}}\big]} \tag{14}$$

for the probability of a seller being in position k if informed consumers without a match never stopped searching before sampling all N sellers. As shown in the proof of the next theorem, a seller's profits are proportional to

$$p\Bigg(\big(1-\bar{\mu}(\mathcal{K})\big)\Big[1-F(p)\Big] + \bar{\mu}(\mathcal{K})\Big[1-F\big(p^e+s/m\big)\Big]\Bigg), \tag{15}$$

where $\bar{\mu}(\mathcal{K}) = \sum_{k=1}^{\mathcal{K}}\phi_k(\mathcal{K})\mu_k + \sum_{k=\mathcal{K}+1}^{N}\phi_k(\mathcal{K})$ and $1-\bar{\mu}(\mathcal{K}) = \sum_{k=1}^{\mathcal{K}}\phi_k(\mathcal{K})(1-\mu_k)$.

[20]If $\mathcal{K} = 0$, informed consumer do not search either as they would always be held up.

[21]Except for very few values of $\mathcal{K}$, $\tilde{p}(\mathcal{K})$ is decreasing when $\lambda > 0.5$ or $m > 0.25$ as well.

[22]When Lemma A.10 applies, this equilibrium is the one with the lowest price. Then, one can also show that it achieves the highest profits and consumer surplus among all equilibria without tracking.

The first order-condition characterizing the equilibrium price is given by

$$p = \frac{1 - F(p)}{f(p)} + \frac{\bar{\mu}(\mathcal{K})}{1 - \bar{\mu}(\mathcal{K})} \frac{1 - F(p^e + s/m)}{f(p)}. \tag{16}$$

Observe that (16) resembles condition (10) from the previous section as $\bar{\mu}(\mathcal{K})/\big(1-\bar{\mu}(\mathcal{K})\big)$ represents the ratio of informed to uninformed consumers too. Hence, its right-hand side is decreasing in p upon substituting $p^e = p$, thereby guaranteeing uniqueness of the fixed point $p = p^e = \tilde{p}(\mathcal{K})$ for all $\mathcal{K} \geq 1$. Moreover, holding p^e constant, the right-hand side of (16) is decreasing if f is non-decreasing or if λ is sufficiently small, ensuring that no seller wants to deviate from $\tilde{p}(\mathcal{K})$. Notably, when fixing $\mathcal{K}$, the critical threshold value $\tilde{\lambda}^*(\mathcal{K})$ lies above the threshold that ensures existence of an equilibrium in the regime with tracking.

Theorem 2 *Suppose that f is non-decreasing or that λ is sufficiently small. Then, there exists a unique equilibrium price $\tilde{p}(\mathcal{K})$ for each $\mathcal{K} \in E$.*

A key property of any equilibrium in which uninformed consumers stop searching after $\mathcal{K} \geq 1$ sellers is that consumers who learn θ after $\mathcal{K}$ search attempts continue searching for a match until no seller remains if $\theta < \tilde{p}(\mathcal{K}) + s/m$. This contrasts with the equilibrium outcome with tracking, where informed consumers search weakly less than uninformed ones. We will see in Section 5 that this results in a relatively high share of informed consumers if tracking is not available and $\mathcal{K}$ is small, which can sometimes cause a breakdown of the search market that would not occur with tracking.

Finally, (13) implies that an equilibrium in which uninformed consumers are willing to sample all N sellers exists if and only if

$$V_N\big(\tilde{p}(N)\big) \geq 0 \Leftrightarrow m \int_{\tilde{p}(N)}^{\bar{v}} \big(\theta - \tilde{p}(N)\big) f(\theta) \mathrm{d}\theta - s \geq 0. \tag{17}$$

Notably, $V_N\big(\tilde{p}(N)\big)$ need not decrease in s as the positive indirect effect due to a lower price may dominate the negative direct effect due to higher search costs. However, it is easy to see from (17) that $V_N\big(\tilde{p}(N)\big) > 0$ when $s \to 0$, implying a cut-off condition in terms of s that is sufficient to ensure existence.

Remark 1 *There exists a value $s^* > 0$ such that $N \in E$ if $s \leq s^*$.*

4 Surplus Effects of Banning Tracking

The analysis in the preceding section shows that banning tracking significantly changes the equilibrium outcome. Here, I analyze how tracking affects profits as well as welfare when search costs are low. Restricting attention to low search costs allows me to obtain unambiguous predictions under otherwise fairly general assumptions. The consequences of tracking for consumers are ambiguous as briefly shown at the end of this section. The case of larger search costs is dealt with in the next section, where I assume that θ is uniformly distributed.

I begin by comparing the price without tracking to the average of all prices with tracking. One difficulty of this exercise lies in defining the average price. This is because in the regime with tracking, informed consumers likely stop searching before the Nth seller and may thus neither pay nor see the high prices that sellers ask for after several search attempts. To compare "apples to apples", I therefore define the average price as the price that a consumer who always buys if she finds a match, and always continues searching otherwise, expects to pay ex ante.

Observe that $\phi_k(N)$ as defined in (14) captures the relative frequency of a kth search attempt if both informed and uninformed consumers without a match always continued searching. Hence, the here defined average price is given by

$$\bar{p} = \sum_{k=1}^{N} \phi_k(N) p_k^+. \tag{18}$$

An appealing feature of $\bar{p}$ is that it weighs each price p_k^+ not only by those consumers who buy from seller k, but by all consumers that p_k^+ has an effect on. That is, it also accounts for consumers who only see seller k's price but decide to not buy as well as all consumers who refrain from sampling seller k because they expect a too high price. When comparing $\bar{p}$ to the equilibrium price without tracking for $\mathcal{K} = N$, I find the following.

Theorem 3 *The equilibrium price without tracking, $\tilde{p}(N)$, satisfies $\tilde{p}(N) < \bar{p}$.*

Theorem 3 is due to the fact that the term $\frac{\mu}{1-\mu}$ in the sellers' first-order condition (10) is convex. This results in the equilibrium prices being convex in μ as well, meaning that the hold-up problem, which drives up equilibrium prices, intensifies at an increasing rate in the share of informed consumers.

Henceforth, I use the subscripts T and NT to indicate search persistence with and without tracking, respectively. As mentioned above, this section focuses on the case of small search costs such that $\mathcal{K}_{NT} = N$ in equilibrium. By Remark 1, this holds if $s \leq s^*$. To proceed, define

$$\omega(\mu, p) := (1-\mu)\left(m\int_p^{\bar{v}} \theta f(\theta)\mathrm{d}\theta - s\right) + \mu\int_{p+s/m}^{\bar{v}} \big(m\theta - s\big) f(\theta)\mathrm{d}\theta, \tag{19}$$

representing the surplus generated when a seller meets a consumer who follows $\mathcal{R}^*$ for any informedness-price pair (μ, p). Importantly, my analysis relies on the assumption that $\alpha(p) < 1$, where $\alpha(p) := -pf'(p)/f(p)$.[23] Broadly speaking, $\alpha(p)$ is a measure of the convexity of direct demand faced by a monopolist in this market (see Aguirre, Cowan, and Vickers, 2010, for another application of this measure). Take the profit function $p\,(1 - F(p))$ as an example, which is concave if and only if $\alpha(p) < 2$. Moreover, it can be shown that $\alpha(p) < 1$ is a sufficient condition for ω being concave in p when holding μ constant. In terms of f, $\alpha(p) < 1$ requires that f not be too concave and sufficiently larger than zero over its entire support.[24]

Then, the welfare comparison hinges on the following inequalities,

$$\sum_{k=1}^{N} \phi_k(N)\omega\big(\mu_k, p_k^+\big) \leq \omega\big(\bar{\mu}(N), \bar{p}\big) \;< \omega\big(\bar{\mu}(N), \tilde{p}(N)\big), \tag{20}$$

where $\bar{\mu}(N) = \sum_{k=1}^{N} \phi_k(N)\mu_k$. The first inequality holds true only for certain values of s because $\omega(\mu, p)$ is not a concave function, i.e. not jointly concave in μ and p. However, by letting s approach zero and using $\alpha(p) < 1$, Lemma B.2 in the online appendix provides a value $\hat{s} > 0$ such that $\omega(\mu, p)$ is indeed

[23] Observe that $\alpha(p) < 1$ does not imply log-concavity of f. However, $\alpha(p) < 2$ implies that a seller's profit $pD(p, p^e)$ given by (9) in the case with and by (15) in the case without tracking, respectively, is the sum of two functions which are concave in p. As a result, sellers have a unique best-response price for any given expected price and existence as well as uniqueness of the equilibrium prices are guaranteed by the log-concavity of f and $\alpha(p) < 2$. That is, I do not need to assume that λ is sufficiently small or that f is non-decreasing.

[24] Even though the latter restriction is not an unrealistic assumption, it is the reason why most standard probability distributions violate $\alpha(p) < 1$.

concave for all $s \leq \hat{s}$. If this condition is met, the first inequality follows from an appropriate version of Jensen's inequality presented in the online appendix. The second inequality readily follows from Theorem 3 and the fact that $\omega(\mu, p)$ is strictly decreasing in p; it is valid for any level of search costs.

Multiplying the inequalities in (20) by $(1 - (1 - m)^N)/m$ yields welfare with tracking for $\mathcal{K}_T = N$ on the left-hand side and welfare without tracking for $\mathcal{K}_{NT} = N$ on right-hand side, respectively. This is formally shown in the proof of Theorem 4. The result ultimately follows from choosing an s_N^T small enough such that $\omega(\mu_k, p_k^+) \geq 0$ for all k if $s \leq s_N^T$, which implies that welfare with tracking for any $\mathcal{K}_T < N$ is bounded from above by the left-hand side of (20). The joint condition of $s \leq \min\{s^*, \hat{s}_\omega, s_N^T\}$ is also what is meant when saying that the following statement, Theorem 4, holds for sufficiently small search costs.

Theorem 4 *Suppose that $\alpha(p) < 1$. Then, tracking reduces welfare if search costs are sufficiently small.*

Notice that search history-based pricing in my model resembles first-degree price discrimination to the extent that conditional on actively searching, consumers with long search histories do have a higher expected willingness-to-pay. It thus follows from Theorem 4 that the standard intuition regarding first-degree price discrimination cannot easily be extended to dynamic environments. Likewise, Armstrong and Zhou (2016) find that the ability of competing sellers to recognize returning consumers lowers welfare as it makes consumers search too little in equilibrium as well. Zhou (2011) reaches the same conclusion regarding the effect of ordered search, albeit his analytical proof relies on large search costs.

Perhaps surprisingly, tracking reduces welfare mainly because sellers fare worse than if they could not observe search histories. To see why, observe that

$$\pi(\mu, p) := p\left((1 - \mu)\,[1 - F(p)] + \mu\big[1 - F\big(p + s/m\big)\big]\right) m \tag{21}$$

captures the expected profit per visiting consumer given an informedness-price pair (μ, p). Analogously to the argument outlined above, I require that $\alpha(p) < 1$

to establish that

$$\sum_{k=1}^{N} \phi_k(N)\pi\big(\mu_k, p_k^+\big) \leq \pi\big(\bar{\mu}(N), \bar{p}\big) < \pi\big(\bar{\mu}(N), \tilde{p}(N)\big). \tag{22}$$

Lemma B.3 in the online appendix shows that the first inequality holds true. It is based on the same version of Jensen's inequality as Lemma B.2, yet it does not require that s be small. The second inequality results from the strict concavity of $\pi(\mu, p)$ in p for a constant μ (Lemma A.12), Theorem 3, and Lemma A.11, which implies that $\pi(\mu, p)$ is decreasing from $\tilde{p}(N)$ to $\bar{p} > \tilde{p}(N)$. Finally, multiplying (22) by $(1-(1-m)^N)/m$ yields profits with tracking on the left-hand side and without tracking on the right-hand side, respectively. Since industry profits with tracking are highest when $\mathcal{K}_T = N$, the case of $\mathcal{K}_T < N$ requires no further consideration. This is why the next theorem relies only on $s \leq s^*$, which is to guarantee that $\mathcal{K}_{NT} = N$.

Theorem 5 *Suppose that $\alpha(p) < 1$. Then, tracking reduces profits if $s \leq s^*$.*

Note that consumers and sellers are exposed to the hold-up problem both with and without tracking, leading to equilibrium prices above the profit-maximizing level in either case as formally shown in Lemma A.11. Yet, the theorem implies that sellers are better off with a constant hold-up problem than with one that ranges from nonexistent when $\mu_1 = 0$ to extremely severe when $\mu_N \to 1$.

Related papers from the literature on consumer search have so far found mixed results regarding the effects of tracking or ordered search on industry profits. In Armstrong and Zhou (2016), recognizing returning consumers is detrimental to profits under competition. In contrast, Zhou (2011) finds that industry profits are higher when search is ordered.

Lastly, observe that no seller unilaterally prefers to disregard search history information even though total profits would be higher under non-discriminatory pricing. Since the loss in profits does not result from fiercer competition between sellers, this implies the presence of negative externalities, which I explore in more detail in section 6.2.

In contrast to profits and welfare, consumer surplus is convex in p. Additionally, equilibrium prices are convex in μ by Theorem 3. This renders the effect of tracking on consumer surplus ambiguous, which can be seen from the following example. Fix $N = 10$ as well as $m = 0.25$ and let $s \to 0$ such that $\mathcal{K}_T = \mathcal{K}_{NT} = 10$. Then, tracking raises consumer surplus if and only if $\lambda > \bar{\lambda} \approx 0.28$.[25]

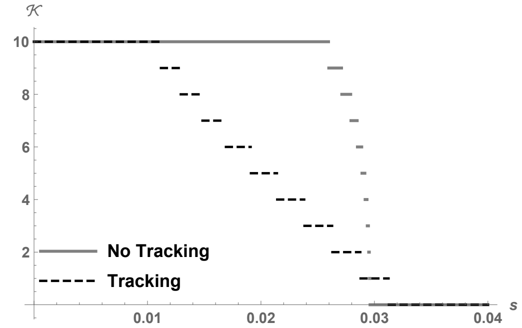
5 Comparative Statics and Uniformly Distributed Types

This section complements the preceding analysis for a wider range of search costs. It also develops more intuition for the surplus results presented before. To find explicit expressions for prices, I assume that F is the uniform distribution with support $[0, 1]$. Consequently, the density f is log-concave and non-decreasing, implying that the existence and uniqueness results from Sections 2 and 3 apply. In addition, I assume that $N = 10$, $m = 0.25$ and $\lambda = 0.1$.[26]

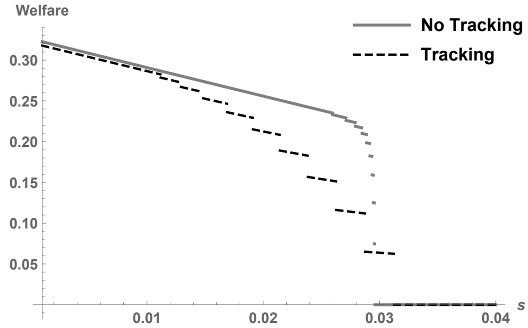
Figure 2a shows that $\mathcal{K}_{NT}$ as well as $\mathcal{K}_T$ are gradually decreasing in s. As expected, uninformed consumers in the tracking regime sample up to N sellers only if search costs are small and otherwise search less than when tracking is banned for a wide range of s. This is because tracking leads to high priced offers for long search histories, which are worth sampling only if search costs are sufficiently small. Importantly, we see that $\mathcal{K}_{NT} = 0$ while $\mathcal{K}_T \geq 1$ if search costs fall within a certain range. In this case the search market is active if and only if tracking is feasible.

The reason why this is possible is the following. A consumer's expected surplus from initiating search is positive as long as the surplus from sampling the first seller is. Notably, the price of the first seller consumers sample when tracking is feasible, p_1^+, is independent of λ. Without tracking in contrast, sellers would always charge a price above p_1^+. This is because some consumers learn their type after sampling the first seller and then continue sampling up to N sellers. As these informed consumers cannot be distinguished from uninformed ones, their presence makes sellers want to hold up consumers unless the expected price is already sufficiently higher than p_1^+. Consequently, when search costs increase, the

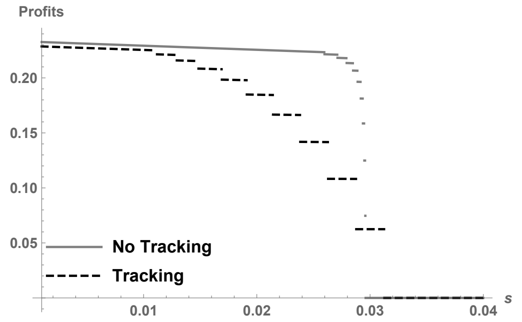
[25] A necessary and sufficient threshold rule in terms of λ does not exist for all parameter values.
[26] The results remain qualitatively unchanged if N, m or λ take on other values.



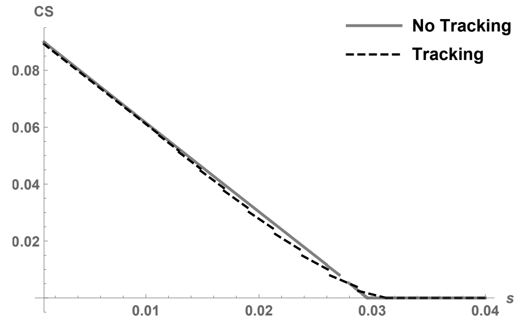
(a) Search Persistence



(b) Welfare



(c) Industry Profits



(d) Consumer Surplus

Figure 1 Effects of changing Search Costs.

active search equilibrium ceases to exist without tracking first.[27]

Figure 2b illustrates how changes in search persistence have a significant effect on total surplus. As predicted by Theorem 4, tracking reduces welfare when search costs are low and $\mathcal{K}_{NT} = \mathcal{K}_T = N$. In addition, if $\mathcal{K}_{NT} > \mathcal{K}_T$ for $s \in (0.012, 0.029)$, the welfare loss due to tracking is even larger. However, we also see that tracking is welfare-enhancing when the market is inactive without it, which holds approximately for $s \in (0.03, 0.031)$.

In line with Theorem 5, Figure 2c shows that tracking reduces profits when search costs are low. Similarly to the welfare comparison, it also illustrates how

[27]Notice that there neither exists an active search mixed strategy equilibrium. To see why, suppose that consumers expect a distribution of prices in equilibrium. Consumers would then use a reservation price rule to decide whether they continue searching or not. Since all consumers value money equally, there would exist a single reservation price. Given such a reservation price, charging a price equal to the reservation price strictly dominates mixing over a set of prices. The key difference to Stahl (1989) is that sellers cannot attract a positive mass of shoppers by undercutting all their competitors.

the difference in profits due to tracking increases when the gap between $\mathcal{K}_T$ and $\mathcal{K}_{NT}$ widens. Moreover, it shows that tracking raises profits when it prevents the market from shutting down as for approximately $s \in (0.03, 0.031)$.

Finally, Figure 2d shows that consumers are barely affected by tracking, even when it makes them search less. The underlying reason is that when consumers with a long search history refrain from sampling another seller, they are barely giving up on their own surplus but on generating profits for the seller, who would extract most of the surplus.

6 Extensions

The previous analysis suggests that banning tracking or price discrmination can have desirable welfare implications. However, such a policy might change factors not modeled here. For instance, tracking may enable firms to target consumers with more suitable ads, which can be desirable to some extent (as in Johnson, 2013). Besides, price discrimination can be welfare-improving and distinguishing search history-based pricing from favorable types of price discrimination may be impossible for the regulator. In this section, I investigate whether alternative data access and tracking regimes can mitigate search history-based pricing without brute-force regulatory intervention. For simplicity, I assume that search costs are sufficiently low such that an equilibrium with $\mathcal{K} = N$ exists, with tracking as well as without.

6.1 The Right to Opt Out

The GDPR grants data subjects in the European Union the right "to object [...] at any time to processing of personal data".[28] To assess the effect of such a policy, I extend the model from Section 2 by allowing consumers to either opt out of tracking or opt in to it on a per seller basis.

The key assumption is that opting out of tracking prevents sellers from observing search histories. Though this could be relaxed, further assume that when a consumer opts in to tracking at seller k after having opted out before, seller

[28] See "The Right to Object" (Art. 21 GDPR).

k observes the complete search history. Consequently, the decision to opt out of tracking has no influence on what forthcoming sellers may observe, implying that the myopically best choice is also sequentially optimal.

If consumers opt out of tracking, sellers cannot discriminate between different search histories and thus charge a single non-discriminatory price in any symmetric equilibrium. Yet, sellers still condition that price on *when* consumers opt out, leading to the following equilibrium property.

Lemma 2 *Consumers never opt in to tracking after having opted out once.*

I prove this result here, but omit steps which follow immediately from the previous analysis. Suppose that consumers' expectations are such that the stopping rule $\mathcal{R}^*$ is optimal and that seller $\hat{k} \leq N$ is the first seller prior to which consumers opt out of tracking. Further, let $\hat{p}(\hat{k})$ denote seller $\hat{k}$'s non-discriminatory price conditional on the consumers' opt-out strategy.

Consumers opt out of tracking prior to sampling seller $\hat{k}$ if and only if they expect $\hat{p}^e(\hat{k}) < p^e_{\hat{k}}$, where $p^e_{\hat{k}}$ represents the expected price conditional on opting in. Moreover, if they enables tracking when sampling seller $l > \hat{k}$, they must expect $p^e_l > p^e_{\hat{k}} > \hat{p}^e(\hat{k}) = \hat{p}(\hat{k})$ in any equilibrium as a result of Theorem 1. Consequently, consumers never enable tracking again after having opted out once.

Conditional on opting in to tracking, Theorem 1 states that prices are increasing in k. Thus, since consumers never switch back and forth between enabling and disabling tracking, prices are weakly increasing with the number of search attempts, regardless of tracking. Consequently, consumers indeed find it optimal to use the search rule $\mathcal{R}^*$ as postulated above. In addition, prices conditional on opting in are characterized by the same equilibrium condition as in Section 2, which is

$$p_k = \frac{1 - F(p_k)}{f(p_k)} + \frac{\mu_k}{1 - \mu_k} \frac{1 - F\big(p_k + \frac{s}{m}\big)}{f(p_k)}, \tag{23}$$

The price for consumers who opt out is derived analogously to the equilibrium price without tracking in Section 3, except its derivation conditions on $k \geq \hat{k}$. The expected demand function of a seller when facing a consumer who opts out (see

(A.47) in the appendix) implies the following equilibrium condition for $\hat{p}(\hat{k})$:

$$\hat{p}(\hat{k}) = \frac{1 - F\big(\hat{p}(\hat{k})\big)}{f\big(\hat{p}(\hat{k})\big)} + \frac{\hat{\mu}(\hat{k})}{1 - \hat{\mu}(\hat{k})} \frac{1 - F\big(\hat{p}(\hat{k}) + \frac{s}{m}\big)}{f\big(\hat{p}(\hat{k})\big)}, \tag{24}$$

where $\hat{\mu}(\hat{k}) = \frac{1}{\sum_{j=\hat{k}}^{N} \phi_j(N)} \left(\sum_{k=\hat{k}}^{N} \phi_k(N) \mu_k \right)$. Since $\hat{\mu}(\hat{k}) > \mu_{\hat{k}}$, the right-hand side of (23) for $k = \hat{k}$ is smaller than the right-hand-side of (24) if $\hat{p}(\hat{k}) = p_k$. Consequently, $p_{\hat{k}}^{+} < \hat{p}(\hat{k})$ for any $\hat{k} < N$, which leads to the next result.

Lemma 3 *Opting out of tracking at any seller $\hat{k} < N$ is no equilibrium strategy.*

The underlying intuition is reminiscent of the unravelling mechanism discovered by Milgrom and Roberts (1986). If consumers opt out of tracking at and after seller $\hat{k}$, the share of informed consumers conditional on opting out exceeds $\mu_{\hat{k}}$. Since the price must be higher when consumers are more likely to know their type, $\hat{p}(\hat{k})$ always exceeds $p_{\hat{k}}^{+}$, the price a consumer can obtain if she does not opt out at seller $\hat{k}$. A consumer would therefore always deviate to opting in unless $\hat{k} = N$, when she is indifferent. Proposition 1 summarizes these findings.

Proposition 1 *In equilibrium, consumers strictly prefer to opt in to tracking unless they are sampling the last remaining seller.*

Since welfare and even consumer surplus can be negatively affected by tracking, Proposition 1 implies that consumers may block too little in equilibrium. This stands in contrast to Conitzer, Taylor, and Wagman (2012), who find that if consumers can maintain anonymity for free in a BBPD framework, they may hurt themselves by choosing this option too often.[29]

6.2 Commitment

Can the power to commit to not tracking search histories (e.g. by advertising to honor privacy) resolve the sellers' dilemma? To shed light on this question, I extend the main model of Section 2 by a new first stage during which sellers

[29] Customer recognition raises consumer surplus in their model because rational consumers delay their purchases unless sellers lower their prices in the first period.

choose a tracking policy once and for all. After observing all sellers' tracking policies, consumers decide what type of seller (tracking or non-tracking), if any, to sample. Conditional on selecting a type, the order of search is random. Consumers can switch between different types as they continue searching.

Since a consumer's choices concerning the type of seller she visits are independent of one another, the myopically best choice is also sequentially optimal. Analogously to the previous section, non-tracking sellers in equilibrium know *when* consumers sample a non-tracking seller and charge a price conditional on that event. Besides, tracking sellers set the prices derived in Theorem 1.

The first prediction is that consumers would sample tracking sellers exclusively if there were sufficiently many of them due the same unravelling argument discussed before. However, if some sellers commit to not tracking, consumers without a match have to sample a non-tracking seller at some point if they want to continue searching. In that case, as it turns out, it would always be profitable for a non-tracking seller to deviate and commit to tracking instead.

To see why, suppose that $n > 1$ sellers commit to not tracking consumers.[30] As argued before, a non-tracking seller will not be sampled before all $N - n$ tracking sellers and may even be sampled last, if at all. By committing to tracking instead, a formerly non-tracking seller can make sure to be among the first $N - (n + 1)$ sellers a consumer samples. Then, the next result follows if $\alpha(p) < 2$, which implies that sellers earn higher profits, the earlier they are sampled by consumers.

Proposition 2 *If $\alpha(p) < 2$, all sellers commit to tracking in equilibrium.*

The statement is formally proven in the appendix. While it is intuitive that sellers prefer to have more information about consumers once these have incurred the search cost and are locked-in, it is surprising that sellers do not benefit from commitment power even though tracking makes them worse off. Intuitively, this is because a seller can "jump the queue" by committing to tracking when other sellers do not, which exerts a negative externality on his competitors.

[30] If $n = 1$, the non-tracking seller correctly anticipates his position in equilibrium.

6.3 Search on a Platform

Platforms on which buyers can shop for products of multiple sellers are ubiquitous on the Internet. Moreover, prices on these platforms seem to vary little with the order of search even though a platform can easily track all search activities on its website. Introducing a common platform for sellers and buyers to the model presented in Sections 2 and 3 reconciles its predictions with this observation.

Suppose that there is an intermediary platform which sellers can join to sell their products and that consumers search exclusively on the platform. In addition, assume that it employs a non-distortionary fee structure and seeks to maximize total platform revenue. Finally, assume that the platform automatically tracks consumers' search activities and that it can commit either to sharing or to withholding all search history data before consumers begin searching.

The subgame that follows if the platform agrees to sharing search histories is identical to search with tracking as analyzed in Section 2. In contrast, sellers set a price equal to $\tilde{p}(N)$ as derived in Section 3 if the platform chooses to not share its data. If total revenue is higher in the absence of search history-based pricing, as has been shown to be true in many cases, a profit-maximizing platform will thus not share search history data with sellers.

More generally, a dominant platform will likely prevent search history-based pricing if it has both the power to commit to a data sharing policy ex ante and the ability to internalize the afore-mentioned demand externalities.

7 Concluding Remarks

I have proposed a framework to study the role of tracking when consumers learn about their preferences while searching for the right product and price. The analysis has revealed that tracking is likely welfare-reducing, in particular when search costs are small. Moreover, sellers price discriminate in equilibrium even though it often reduces industry profits overall, regardless of their ability to commit to a tracking policy ex ante. In the remainder of this section, I discuss two implications of the model beyond consumer search.

First, the combination of tracking and learning, which leads to the main mechanism identified in this paper, may prevail in a special case of sequential bargaining. To illustrate, imagine a multi-product monopolist and a single buyer, bargaining over multiple rounds. The buyer, who may cease bargaining after any round and receive a payoff of zero, can obtain one quote per round at a sampling cost $s > 0$ from the monopolist, who is committed to offer a different product in each round. The buyer is otherwise identical to a representative consumer of the model laid out in Section 2. In particular, she is uninformed about her valuation for a match before bargaining starts but may learn about it while bargaining.

Other things held constant, buyers with higher match values are, of course, more likely to make a purchase. Hence, one might expect that prices decrease over time as conjectured by Coase (1972). Yet, low value buyers are not forced to continue bargaining when they have an outside option like in this model. As Board and Pycia (2014) show, this property prevents the negative selection mechanism, which usually puts downward pressure on prices.

Instead, the price the monopolist would set in any round k is equal to p_k^+, implying that prices even increase the longer negotiations last in this environment. To see why, notice that the only difference to the search framework with competing sellers is that the monopolist internalizes the effect of the price set in round k on demand in round $k' \neq k$. As it is possible to show that the monopolist cannot prepone a buyer's purchase and never wants to postpone one, the price p_k^+ remains his best-response if he cannot commit to a price path ex ante.

In addition, learning may play an important role in the context of job search, where employers can usually infer a candidate's unemployment duration from seeing the CV. If candidates learn about their reservation wage during unemployment, the mechanism explored in this paper implies a novel link between wage offers and unemployment duration, the neglect of which may be one reason for the so far unexplained portion of the empirically observed wage dispersion (see Hornstein, Krusell, and Violante, 2011).

For instance, candidates may become better informed about their reservation wage during unemployment by learning more about working conditions or by researching job opportunities in other sectors that might require retraining. Then, the longer a candidate's unemployment duration, the larger the hold-up problem if applying for jobs is costly. Consequently, wages offered to applicants with longer unemployment duration would need to be lower to mitigate the incentive for employers to hold up candidates.

This theory is consistent with the recent evidence by Schmieder, von Wachter, and Bender (2016), who find a negative effect of non-employment duration on wages which is not driven by differences in reservation wages. It also matches recent findings by Krueger and Mueller (2016), who observe that reservation wages decline little over the spell of unemployment and, thus, cannot be the sole reason for the negative effect of unemployment duration on wage offers.

A Mathematical Appendix

The Proof of Theorem 1 is divided into several parts. Proposition A.3 gives equilibrium existence, whereas Lemma A.8 and Lemma A.9 together ensure uniqueness. The uniqueness result is underpinned by the following observations. Firstly, the prices $(p_k^+)_{k\leq\mathcal{K}}$ are the unique equilibrium prices in any equilibrium that exhibits weakly increasing prices. Secondly, consumers must expect increasing prices in any PBE. For future reference define

$$A(p) := \frac{1-F(p)}{f(p)}, \quad B(p) := \frac{1-F(p+s/m)}{f(p)}, \quad \psi_k(p) := A(p) + \frac{\mu_k}{1-\mu_k}B(p)$$
$$\text{and} \qquad \hat{\psi}_k(p,p^e) := \frac{1-F(p)}{f(p)} + \frac{\mu_k}{1-\mu_k}\frac{1-F(p^e+s/m)}{f(p)}.$$

Lemma A.4 *Consumers expect $p_k^e \geq \bar{v} - s/m$ for $k > \mathcal{K}$ in any equilibrium.*

The proof of this statement is given in the main text. Any expected price p_k^e must satisfy the inequality or else informed consumers would sample beyond $\mathcal{K}$ and face prices that would be inconsistent with their expectations.

Lemma A.5 *For each k, the fixed point p_k^+ is unique. Besides, these prices satisfy $p_k^+ < p_{k+1}^+$.*

PROOF. Using notation from above, $\psi_k(p_k) - p_k = 0$ combines the necessary first-order condition (10) for position k with the rational expectations condition $p_k = p_k^e$. Clearly, $\psi_k(p_k^+) - p_k^+ = 0$ must hold in equilibrium for all $k \leq \mathcal{K}$.

I first show that $\psi_k(p)$ is a decreasing function in p. By Corollary 2 in Bagnoli and Bergstrom (2005) log-concavity of f implies that $A(p)$ is monotone decreasing and it remains to show that $B(p)$ is decreasing. Let $z(p) := p + \frac{s}{m}$ and note that $\partial_p z(p) = 1$. Hence $B(p)$ is decreasing if and only if $\partial_p \frac{1-F(z(p))}{f(p)} = \partial_z \frac{1-F(z)}{f(z-\frac{s}{m})} \leq 0$. From the latter derivative we see that

$$\partial_z \frac{1-F(z)}{f(z-\frac{s}{m})} \leq 0 \quad \text{if and only if} \quad -\frac{f(z)}{1-F(z)} \leq \frac{f'(z-\frac{s}{m})}{f(z-\frac{s}{m})}. \tag{A.25}$$

Moreover, recall that log-concavity of f implies that

$$\partial_z \frac{1-F(z)}{f(z)} \leq 0 \quad \text{and, thus} \quad -\frac{f(z)}{1-F(z)} \leq \frac{f'(z)}{f(z)}. \tag{A.26}$$

Evidently (A.25) follows from (A.26) if

$$\frac{f'(z)}{f(z)} \leq \frac{f'(z-\frac{s}{m})}{f(z-\frac{s}{m})}, \tag{A.27}$$

or, equivalently $\partial_z\big(f'(z)/f(z)\big) \leq 0$ for all z, i.e. if f is log-concave.

Next, I show that condition (10) has a unique fixed point. Since ψ_k is continuous in p and $\lim_{p\to 0}[\psi_k(p)-p] > 0$, while $\lim_{p\to\bar{v}}[\psi_k(p)-p] < 0$, the intermediate value theorem guarantees the existence of p_k^+ with $\psi_k(p_k^+) = p_k^+$. Moreover, since ψ_k is monotone, p_k^+ is unique.

Finally, I prove the claimed ordering of fixed points by contradiction. Assume $p_{k+1}^+ < p_k^+$ for some k. Since $\lambda > 0$ implies $\mu_{k+1} > \mu_k$ we see that $\psi_{k+1}(p_{k+1}^+) > \psi_k(p_{k+1}^+)$. Moreover, since ψ_k is decreasing, we have $\psi_k(p_{k+1}^+) \geq \psi_k(p_k^+)$. Upon combining these inequalities and imposing the equilibrium condition we arrive at $p_{k+1}^+ = \psi_{k+1}(p_{k+1}^+) > \psi_k(p_k^+) = p_k^+$, a contradiction

Lemma A.6 *Assume $p_k^e = p_k^+$ for all $k \leq \mathcal{K}$ and $p_k^e \geq \bar{v} - s/m$ for all $k > \mathcal{K}$. If f is either non-decreasing, or satisfies*

$$\frac{f(p_k)}{1-F(p_k)} \leq \frac{2(1-\mu_k)f(p_k)}{(1-\mu_k)\left[1-F(p_k)\right]+\mu_k} \qquad \text{for } p_k < p_k^+ + s/m \text{ and } k \leq \mathcal{K}, \tag{A.28}$$

then no seller has a profitable deviation from p_k^+ to any p_k in $[0, p_k^+ + s/m)$. More-

over, there is a value $\lambda^*(\mathcal{K}) > 0$ such that (A.28) holds if $\lambda \leq \lambda^*(\mathcal{K})$.

PROOF. In equilibrium, no consumer samples any seller $k > \mathcal{K}$ by Lemma A.4. Hence, seller $k > \mathcal{K}$ has no incentive to deviate. In addition, footnote 16 in the main text describes a simple way to break the underlying indifference.

By Lemma A.5, $p_k^e = p_k^+$ implies that consumers expect rising prices for all $k \leq \mathcal{K}$, i.e. for all sellers that are sampled with positive probability. Consider a seller $K \leq \mathcal{K}$ deviating to some p_K in $[0, p_K^e + s/m)$. By Lemma 1, this implies that consumers still use $\mathcal{R}^*$ and that demand equals $D_K(p_K, p_K^e)$ as in (7).

As the derivative of $p_K D_K(p_K, p_K^e)$ is defined for all p_K in $[0, p_K^e + s/m)$, the local first-order condition is given by $\hat{\psi}_K(p_K, p_K^e) - p_K = 0$. By Lemma A.5, $\hat{\psi}_K(p_K, p_K^e) - p_K = 0$ if $p_K = p_K^e = p_K^+$, implying that p_K^+ is indeed an extreme point. Besides, $p_K = p_K^+$ globally maximizes $p_K D_K(p_K, p_K^e)$ for a given p_K^e if $\hat{\psi}_k(p_K, p_K^e) - p_K \leq 0$ for all $p_K \in (p_K^e, p_K^e + s/m)$, and if $\hat{\psi}_k(p_K, p_K^e) - p_K \geq 0$ for all $p_K \in [0, p_K^e)$. Given $p_K^e = p_K^+$, both conditions hold if $\partial_{p_K} \hat{\psi}_K(p_K, p_K^+) - 1 \leq 0$ for all any p_K in $[0, p_K^+ + s/m)$. Some algebra shows that the latter condition is equivalent to

$$\frac{f'(p_K)}{f(p_K)} \geq -\frac{2(1-\mu_K)f(p_K)}{(1-\mu_K)\big[1-F(p_K)\big] + \mu_K\big[1-F(p_K^e + s/m\big]}, \tag{A.29}$$

on $[0, p_K^+ + s/m)$. This clearly holds if f is non-decreasing. Alternatively, inequality (A.29) is implied by the validity of

$$\frac{f(p_K)}{1-F(p_K)} \leq \frac{2(1-\mu_K)f(p_K)}{(1-\mu_K)\left[1-F(p_K)\right] + \mu_K} \text{ for } p_K < p_K^+ + s/m. \tag{A.30}$$

This is the inequality appearing in the lemma. To see why (A.30) is sufficient, note that

$$(1-\mu_K)\left[1-F(p_K)\right] + \mu_K\big[1-F(p_K^e + s/m\big] \leq (1-\mu_K)\left[1-F(p_K)\right] + \mu_K$$

and that log-concavity implies that $f'(p_K)/f(p_K) \geq -f(p_K)/(1-F(p_K))$.

To prove the last part of Lemma A.6, note that inequality (A.30) holds strictly for $\mu_K = 0$. Moreover, the RHS is continuous and monotone decreasing in μ_K. Thus, for every log-concave density f, there exists exactly one value $\mu_K^* > 0$ such that (A.30) holds for seller K if and only if $\mu_K \leq \mu_K^*$. Setting $\lambda^*(k) := 1 - (1 - \mu_k^*)^{\frac{1}{k-1}}$, we see that (A.30) holds for all sellers jointly if $\lambda \leq \min\left[\{\lambda^*(k)\}_{k\leq\mathcal{K}}\right] = \lambda^*(\mathcal{K})$.

Lemma A.7 *Assume $p_k^e = p_k^+$ for all $k \leq \mathcal{K}$ and $p_k^e \geq \bar{v} - s/m$ for all $k > \mathcal{K}$. Further, suppose that f satisfies one of the conditions from Lemma A.6. Then no seller has a profitable deviation from p_k^+ to any p_k in $[p_k^+ + s/m, \infty)$.*

First, consider deviations by seller $K \leq \mathcal{K}$ to $p_K \in [p_K^+ + s/m, p_{K+1}^+ + s/m]$. Then, consumers still buy upon encountering a match if $\theta \geq p_K$ by Lemma 1. However, demand from informed consumers is elastic now since the lowest type among informed consumers sampling K is $\theta = p_K^+ + s/m \leq p_K$. Using the expression for $D_K(p_K, p_K^+)$ in (7), it follows from $p_K \leq p_K^+ + s/m$ that a marginal change of p_K changes profits by $m\left((1 - F(p_K)) - p_K f(p_K)\right)$. Additionally, Lemma A.5 implies

$$0 = \psi_k(p_K^+) - p_K^+ > \psi_k(p_K) - p_K > \frac{1 - F(p_K)}{f(p_K)} - p_K \qquad \text{(A.31)}$$

since $p_K > p_K^+$. Thus, $(1 - F(p_K)) - p_K f(p_K) < 0$ and since $p_K D_K(p_K, p_K^+)$ is continuous in p_K, any deviation to $p_K \in [p_K^+ + s/m, p_{K+1}^+ + s/m]$ yields lower profits than $p_K = \lim_{\epsilon\to 0} p_K^+ + s/m - \epsilon$, which is not profitable by Lemma A.6.

Second, consider deviations to $p_K \in (p_{K+1}^+ + s/m, \infty)$. Then, the expected price savings from the next search attempt for a consumer with a match are at least $m(p_K - p_{K+1}^+) > s$.[31] Hence, consumers with $v = \theta \geq p_K$ never buy immediately and return only if they do not find another match at the following $n \geq 1$ sellers whose prices satisfy $p_j^+ < p_K - s/m$ for at least some $j > K$. Consequently, the probability of selling to a consumer with a match is just $(1 - m)^n$.

Were $(1 - m)^n$ independent of p_K, the derivative of profits with respect to p_K would be $\big(\big(1 - F(p_K)\big) - p_K f(p_K)\big)\, m(1 - m)^n$ and it would follow from (A.31) that seller K can attain higher profits by choosing a lower price. As $(1 - m)^n$ is in fact weakly decreasing in p_K, it makes a price $p_K \in (p_{K+1}^+ + s/m, \infty)$ even less profitable.

[31] In principle, consumers might sample many more firms before returning to seller k, depending on how large the deviation $p_K - p_K^+$ is. This reduces the probability of consumers returning to k further, which reinforces the following argument.

Proposition A.3 *Suppose that $\lambda \leq \lambda^*(\mathcal{K})$ or that f is non-decreasing. An equilibrium with $p_k = p_k^+ = p_k^e$ for all $k \leq \mathcal{K}$ and $p_k = p_k^e \geq \bar{v} + s/m$ for all $k > \mathcal{K}$ exists.*

PROOF. If $p_k^e = p_k^+$ for all $k \leq \mathcal{K}$ and $p_k^e \geq \bar{v} - s/m$ for all $k \geq \mathcal{K}$, consumers use the search rule $\mathcal{R}^*$ and it follows from Lemma A.4, Lemma A.6 and Lemma A.7 that no seller wants to unilaterally deviate from p_k^e. Thus consumers' expectations are consistent with sellers' pricing strategies, and consumer search behavior is consistent with sellers' beliefs about it.

Uniqueness of the equilibrium constructed above is obtained by combining the following lemmata.

Lemma A.8 *The equilibrium constructed in Proposition A.3 is unique among all equilibria with weakly increasing prices.*

PROOF. In any equilibrium in which consumers expect weakly increasing prices, the price by any seller $k \leq \mathcal{K}$ must satisfy the first-order condition $\hat{\psi}_k(p_k, p_k^e) - p_k = 0$ as well as $p_k = p_k^e$, or equivalently, $\psi_k(p_k) - p_k = 0$. By Lemma A.5, p_k^+ is the unique price for which $\psi_k(p_k) - p_k = 0$ holds.

Lemma A.9 *Expected prices which are not weakly increasing are inconsistent with the resulting best-response prices by sellers.*

PROOF. Let $\mathbf{p}^e = (p_1^e, \ldots, p_N^e)$ be a fixed sequence of expected prices, and suppose that $p_k^e < p_{k-1}^e$ for some $k \leq N$. Let j denote the largest index with $p_j^e < p_{j-1}^e$. Without loss of generality, assume that p_l^e for all $l \leq j$ is low enough such that seller j is sampled with positive probability. Since $p_k^e \geq p_j^e$ for all $k > j$ by construction, the following must be true.

$$\text{Consumers who sample } j \text{ will always buy if } v_j = \theta \geq p_j. \tag{A.32}$$

Despite the similarity between (A.32) and $\mathcal{R}^*$, demand at seller j may be different from $D_j(p_j, p_j^e)$ as derived in (7). On the one hand, consumers may continue searching despite a match at some seller $k < j$, potentially increasing the share of informed consumers seller j faces. On the other hand, informed consumers without a match may stop searching before sampling seller j because of $\hat{V}_k(\theta) < 0$ for some $k < j$ even though $\theta \geq p_j^e + s/m$. To account for both possibilities, let $\delta_j(\theta)$ denote the endogenous probability that a consumer of type θ has sampled seller $j-1$ as well as all sellers $k < j-1$ and that she has become informed.

Notice that $\delta_j(\theta)$ is a non-decreasing step-function, because given $\mathbf{p}^e$ and actual prices, a higher value of θ implies that a consumer is less likely to stop searching if she becomes informed, while the probability of becoming informed is independent of θ. In particular $\delta_j(\theta)$ has a jump at the unique value $\hat{\theta} \in (p_j^e + s/m, p_{j-1}^e + s/m)$ satisfying $\hat{V}_{j-1}(\hat{\theta}) = 0$.[32] Given (A.32) and consumers' expectation p_j^e, the price of seller j has to maximize

$$p_j\left((1-m)^{j-1}(1-\mu_j)\left[1-F(p_j)\right] + \int_{p_j^e+s/m}^{\bar{v}} \delta_j(\theta) f(\theta)\mathrm{d}\theta\right)m \qquad \text{or, equivalently}$$

$$p_j\left((1-m)^{j-2}(1-\mu_j)\left[1-F(p_j)\right] + \int_{p_j^e+s/m}^{\bar{v}} \frac{\delta_j(\theta)}{1-m} f(\theta)\mathrm{d}\theta\right)m, \tag{A.33}$$

where $(1-m)^{j-1}(1-\mu_j)$ represents the mass of uninformed consumers sampling j. To show that $p_j^e < p_{j-1}^e$ always leads to a contradiction, I consider three exhaustive cases: $p_{j-1}^e < p_j^e + s/m$ and $p_{j-1}^e \geq p_j^e + s/m$, with the latter one being analyzed separately for $\hat{\theta} > p_{j-1}^e$ and $\hat{\theta} \leq p_{j-1}^e$.

Case 1: If $p_{j-1}^e < p_j^e + s/m$, then the demand from informed consumers is locally inelastic as it implies that $\hat{\theta} > p_{j-1}^e$. Besides, consumers buy immediately from $j-1$ when they encounter a match if $\theta \geq p_{j-1}$. Thus, $p_{j-1} = p_{j-1}^e$ has to maximize

$$p_{j-1}\left((1-m)^{j-2}(1-\mu_{j-1})\left[1-F(p_{j-1})\right] + \int_{\hat{\theta}}^{\bar{v}} \delta_{j-1}(\theta) f(\theta)\mathrm{d}\theta\right)m, \tag{A.34}$$

while $p_{j-1}^e = p_{j-1}$ must hold as well. Because $\hat{\theta} > p_j^e + s/m$, all informed consumers sampling $j-1$ must be willing to sample seller j as well if they do not find a match at $j-1$. Moreover, additional consumers may not find a match but learn their type after sampling $j-1$. Thus, $\delta_j(\theta) > (1-m)\delta_{j-1}(\theta)$ for all $\theta \geq p_j^e + s/m$, which

[32] This is because an uninformed consumer with type $\theta < \hat{\theta}$ may sample $j-1$ and thereby learn θ.

includes all types of informed consumers who sample j with positive probability. Thus, we have

$$\int_{p_j^e+s/m}^{\bar{v}} \frac{\delta_j(\theta)}{1-m} f(\theta)\mathrm{d}\theta > \int_{\hat{\theta}}^{\bar{v}} \delta_{j-1}(\theta) f(\theta)\mathrm{d}\theta. \tag{A.35}$$

Since $1-\mu_j < 1-\mu_{j-1}$ as well, I can apply the following result, the proof of which can be found in the online appendix.

Remark 2 *Let F be a CDF with positive support and $\underline{\beta} : \mathbb{R}^+ \to \mathbb{R}^+$ be a weakly decreasing function. Further, define*

$$\pi_1(p) := \left(\bar{\alpha}\,(1-F(p)) + \underline{\beta}(p)\right) p, \quad \text{and} \quad \pi_2(p) := \left(\underline{\alpha}\,(1-F(p)) + \bar{\beta}\right) p,$$

and let p_1 maximize π_1 and p_2 maximize π_2. Besides, assume that the constants satisfy $\bar{\alpha} > \underline{\alpha} \geq 0$ and $\bar{\beta} > \underline{\beta}(p_1) \geq 0$. Then, $p_2 \geq p_1$.

Setting $\underline{\beta}(p)$ equal to the RHS of (A.35) and $\bar{\beta}$ equal to its LHS as well as $\bar{\alpha} = 1-\mu_{j-1}$ and $\underline{\alpha} = 1-\mu_j$, we see that if p_j^e maximizes (A.33) and $p_{j-1}^e > p_j^e$ maximizes (A.34), then $p_j^e \geq p_{j-1}^e$, a contradiction. ($\underline{\beta}(p)$ is a constant here.)

Case 2: If $\hat{\theta} > p_{j-1}^e > p_j^e + s/m$, then demand by informed consumers is perfectly inelastic around $p_{j-1} = p_{j-1}^e$ as in Case 1. However, consumers with a match at seller $j-1$ sample (at least) seller j as well and buy from $j-1$ only with probability $(1-m)$, i.e. if they find no match at seller j.[33] That is, profits of seller $j-1$ are given by $(1-m)$ times expression (A.34) if he sets a price equal to p_{j-1}^e. Consequently, seller $j-1$ chooses p_{j-1}^e in equilibrium only if it maximizes expression (A.34). In addition, the fact that no consumer buys from $j-1$ immediately given $p_{j-1} = p_{j-1}^e$ implies that $\delta_j(\theta) > \delta_{j-1}(\theta)$. Analogously to Case 1, invoking Remark 2 then implies a contradiction .

[33]Consumers may sample more than just one seller before returning to seller $j-1$. The probability of selling a good as seller $j-1$ is decreasing in the number of additional search attempts consumers conduct. Hence, the profits derived here for seller $j-1$ are a best-case scenario.

Case 3: If seller $j-1$ sets a price $p_{j-1} = p^e_{j-1} \geq \hat{\theta} > p^e_j + s/m$, then in addition to consumers who first sample seller j before possibly returning to $j-1$, there are informed consumers who sample $j-1$ without the intention to ever buy from $j-1$ because $\theta \leq p_{j-1}$. Their demand is thus elastic, too. Consequently, if seller $j-1$ chooses $p_{j-1} = p^e_{j-1}$ in equilibrium, it has to maximize

$$p_{j-1}\left((1-m)^{j-2}(1-\mu_{j-1})\,[1-F(p_{j-1})] + \int_{p_{j-1}}^{\bar{v}} \delta_{j-1}(\theta) f(\theta)\mathrm{d}\theta\right). \tag{A.36}$$

Comparing (A.36) with (A.33) and invoking Remark 2 immediately leads to a contradiction if $p^e_j < p^e_{j-1}$ because (i) $p_{j-1} = p^e_{j-1} \geq \hat{\theta} > p^e_j + s/m$, (ii) $\delta_j(\theta) > \delta_{j-1}(\theta)$ and (iii) $1-\mu_{j-1} > 1-\mu_j$. This concludes the proof.

PROOF OF THEOREM 2. To abbreviate the denominator of $\phi_k(\mathcal{K})$, define

$$\Phi(\mathcal{K}) := \sum_{j=1}^{N}(1-m)^{j-1}\big[\mathbb{1}_{j\leq\mathcal{K}} + \big(1-(1-\lambda)^{\mathcal{K}}\big)\mathbb{1}_{j>\mathcal{K}}\big]. \tag{A.37}$$

Step 1: There is a unique equilibrium candidate price for each $\mathcal{K}$.

If consumers expect a single non-discriminatory price p^e, a seller who charges $p \leq p^e + s/m$ expects the following demand:

$$\begin{aligned} D(p,p^e) =& \frac{\Phi(\mathcal{K})}{N} m \sum_{k=1}^{\mathcal{K}} \phi_k(\mathcal{K})\left((1-\mu_k)\,[1-F(p)] + \mu_k\left[1-F\big(p^e+s/m\big)\right]\right) \\ &+ \frac{\Phi(\mathcal{K})}{N} m \sum_{k=\mathcal{K}+1}^{N} \phi_k(\mathcal{K})\left[1-F\big(p^e+s/m\big)\right] \\ =& \frac{\Phi(\mathcal{K})}{N} m\Big((1-\bar{\mu}(\mathcal{K}))\,[1-F(p)] + \bar{\mu}(\mathcal{K})\,\big[1-F\big(p^e+s/m\big)\big]\Big), \end{aligned} \tag{A.38}$$

where $\bar{\mu}(\mathcal{K}) = \sum_{k=1}^{\mathcal{K}} \phi_k(\mathcal{K})\mu_k + \sum_{k=\mathcal{K}+1}^{N} \phi_k(\mathcal{K})$. Thus, profits $(pD(p,p^e))$ are proportional to the expression in (15), implying the validity of the first-order condition given in (16). Then, the equilibrium candidate price $\tilde{p}(\mathcal{K})$ is found by substituting $p = p^e$ into (16) and searching for its fixed point. Except for the difference in constants between μ_k and $\bar{\mu}(\mathcal{K})$, the resulting equation is equivalent to $\psi_k(p) - p = 0$.

Thus, existence and uniqueness of $\tilde{p}(\mathcal{K})$ follow from Lemma A.5.

Step 2: Suppose that $\lambda \leq \tilde{\lambda}^*(\mathcal{K})$ or that f is non-decreasing. Then, no seller wants to deviate from $\tilde{p}(\mathcal{K})$ if $p^e = \tilde{p}(\mathcal{K})$.

The first-order condition in (16) is defined for all $p < p^e + s/m$. Then, given $p^e = \tilde{p}(\mathcal{K})$, no seller has a profitable deviation from $\tilde{p}(\mathcal{K})$ to $p \in [0, \tilde{p}(\mathcal{K}) + s/m)$ if

$$\partial_p \left(\frac{1 - F(p)}{f(p)} + \frac{\bar{\mu}(\mathcal{K})}{1 - \bar{\mu}(\mathcal{K})} \frac{1 - F(\tilde{p}(\mathcal{K}) + s/m)}{f(p)} - p \right) \leq 0 \tag{A.39}$$

for $p \in [0, \tilde{p}(\mathcal{K}) + s/m)$. This holds if f is non-decreasing. In addition, analogously to Lemma A.6, it can be verified that inequality (A.39) holds if

$$\frac{f(p)}{1 - F(p)} \leq \frac{2(1 - \bar{\mu}(\mathcal{K}))f(p)}{(1 - \bar{\mu}(\mathcal{K}))\,[1 - F(p)] + \bar{\mu}(\mathcal{K})} \text{ for } p < \tilde{p}(\mathcal{K}) + s/m. \tag{A.40}$$

As before, there is a $\mu^*(\mathcal{K})$ such that (A.40) holds if and only if $\bar{\mu}(\mathcal{K}) \leq \mu^*(\mathcal{K})$. Since $\bar{\mu}(\mathcal{K})$ is monotone in λ, this implies that there exists a unique threshold value $\tilde{\lambda}^*(\mathcal{K})$ as well. Next, replace p_K^+ and p_{K+1}^+ in Lemma A.7 by $\tilde{p}(\mathcal{K})$ and denote deviations by p instead of p_K. Then, an analogous argument shows that no seller wants to deviate to $p \in [\tilde{p}(\mathcal{K}) + s/m, \infty)$ either.

Step 3: Suppose that $\mathcal{K} \in E$ and that $p^e = \tilde{p}(\mathcal{K})$. Then, equilibrium strategies are consistent with sellers' and consumers' expectations.

Since no seller deviates from $\tilde{p}(\mathcal{K})$ if $p^e = \tilde{p}(\mathcal{K})$, consumers' expectations are correct in equilibrium if they expect $p^e = \tilde{p}(\mathcal{K})$. Moreover, given these expectations, all consumers follow $\mathcal{R}^*$, which is consistent with sellers' beliefs based on which they find it optimal to set a price equal to $\tilde{p}(\mathcal{K})$.

Lemma A.10 *Suppose that $\lambda \leq 0.5$ and $m \leq 0.25$. Then, $\tilde{p}(\mathcal{K}) > \tilde{p}(\mathcal{K}')$ if $\mathcal{K} < \mathcal{K}'$ for all $\mathcal{K}' \leq N \leq 100$.*

PROOF. I can show numerically that the maximum of $\Lambda(\mathcal{K}) - \Lambda(\mathcal{K}-1)$ is negative if $\lambda \in [0, 0.5]$ and $m \in [0, 0.25]$ for all $1 < \mathcal{K} \leq N \leq 100$. Thus, $\Lambda(\mathcal{K}) < \Lambda(\mathcal{K}')$ for $\mathcal{K} > \mathcal{K}' \geq 1$. In addition, the first-order condition (16) implies that $\tilde{p}(\mathcal{K}) < \tilde{p}(\mathcal{K}')$ if $\Lambda(\mathcal{K}) < \Lambda(\mathcal{K}')$. The proof of this implication is identical to the argument presented in the last paragraph in the proof of Lemma A.5.

PROOF OF THEOREM 3. I first show that p_k^+ is a strictly convex function of μ. Define

$$Q(p,\mu) := \frac{1-F(p)}{f(p)} + \frac{\mu}{1-\mu}\frac{1-F(p+s/m)}{f(p)} - p$$

and let $p^+(\mu)$ denote the solution to $Q(p,\mu) = 0$ for a given μ. Notice that $Q(p,\mu_k) = 0$ if and only if $\psi_k(p) - p = 0$, implying $p^+(\mu_k) = p_k^+$. Further, define

$$\begin{aligned} q(\mu) :=& (1-\mu)^2\Big(f(p)^2 + f'(p)\big[1-F(p)\big]\Big) \\ &+\mu(1-\mu)\Big(f(p)f(p+s/m) + f'(p)\big[1-F(p+s/m)\big]\Big) + (1-\mu)^2 f(p)^2. \end{aligned}$$

Then, it can be verified that $\partial_\mu p^+(\mu) = -\frac{\partial Q}{\partial \mu}/\frac{\partial Q}{\partial p} = \frac{\big[1-F(p+s/m)\big]f(p)^2}{q(\mu)}$. Taking the second derivative with respect to μ yields

$$\begin{aligned} \partial_\mu^2 p^+(\mu) = \frac{\big[1-F(p+s/m)\big]f(p)^2}{q(\mu)^2}\Bigg\{&(2-2\mu)\underbrace{\Big(f(p)^2 + f'(p)\big[1-F(p)\big]\Big)}_{\geq 0 \text{ as } \partial_p \frac{1-F(p)}{f(p)} \leq 0 \text{ by log-concavity}} \\ -(1-2\mu)&\underbrace{\Big(f(p)f(p+s/m) + f'(p)\big[1-F(p+s/m)\big]\Big)}_{\geq 0. \text{ This follows from (A.25) upon substituting } z=p+s/m.} + (2-2\mu)f(p)^2\Bigg\}. \end{aligned}$$

Clearly, $\partial_\mu^2 p^+(\mu) \geq 0$ for $\mu = 1$. Moreover, $\partial_\mu^2 p^+(\mu) > 0$ for all $\mu \in [0,1)$ if the term in curly brackets is decreasing in μ. Its derivative with respect to μ shows that this holds if and only if

$$\begin{aligned} &\frac{f(p)^2 + f'(p)[1-F(p)]}{f(p)f(p+s/m) + f'(p)[1-F(p+s/m)]} \\ &> 1 - \frac{f(p)^2}{\underbrace{f(p)f(p+s/m) + f'(p)[1-F(p+s/m)]}_{\geq 0 \text{ (see the argument above)}}}. \end{aligned} \tag{A.41}$$

Inequality (A.41) holds with strict inequality for $s = 0$ since the LHS equals one in this case while the RHS is always smaller than one. Thus, (A.41) holds for all

s if the denominator of the LHS in (A.41) is decreasing in s, which is implied by

$$\frac{1}{m}\big[f'(p+s/m)f(p)-f'(p)f(p+s/m)\big] \leq 0 \quad \Leftrightarrow \quad \frac{f'(p+s/m)}{f(p+s/m)} \leq \frac{f'(p)}{f(p)}.$$

The latter inequality holds because log-concavity of f implies that $\frac{f'(p)}{f(p)}$ is decreasing. This shows that $p_k^+(\mu)$ is strictly convex for all $\mu < 1$.

It remains to show that $\bar{p} > \tilde{p}(N)$. Observe that

$$Q\left(p, \bar{\mu}(N)\right) = 0 \quad \Leftrightarrow \quad p = \frac{1-F(p)}{f(p)} + \frac{\bar{\mu}(N)}{1-\bar{\mu}(N)}\frac{1-F(p+\frac{s}{m})}{f(p)} \tag{A.42}$$

where the second equation represents the well-known equilibrium condition without tracking for $\mathcal{K} = N$. Consequently, an equivalent expression for $\tilde{p}(N)$, the equilibrium price without tracking, is given by $p^+\left(\bar{\mu}(N)\right) = p^+\Big(\sum_{k=1}^N \phi_k(N)\mu_k\Big)$. The proof concludes by invoking the strict convexity result:

$$\bar{p} = \sum_{k=1}^{N} \phi_k(N)p^+(\mu_k) > p^+\left(\sum_{k=1}^{N}\phi_k(N)\mu_k\right) = \tilde{p}(N).$$

PROOF OF THEOREM 4. Denote welfare with tracking for a given search persistence $\mathcal{K}_T$ by $W_T\big(\mathcal{K}_T\big)$. Using the definition of $\omega(\mu,p)$ in the main text, I can write $W_T(N)$ as

$$W_T(N) = \sum_{k=1}^{N}(1-m)^{k-1}\omega\big(\mu_k, p_k^+\big) = \frac{1-(1-m)^N}{m}\sum_{k=1}^{N}\phi_k(N)\omega\big(\mu_k, p_k^+\big),$$

where I have used $\phi_k(N) = (1-m)^{k-1}/\big(\frac{1-(1-m)^N}{m}\big)$, which obtains from (14) after replacing $\mathcal{K}$ by N. Likewise, welfare without tracking for $\mathcal{K}_{NT} = N$ is given by

$$W_{NT}(N) = \frac{1-(1-m)^N}{m}\sum_{k=1}^{N}\phi_k(N)\omega\big(\mu_k, \tilde{p}(N)\big) = \frac{1-(1-m)^N}{m}\omega\left(\bar{\mu}(N), \tilde{p}(N)\right)$$

where $\bar{\mu}(N) = \sum_{k=1}^{N} \phi_k(N)\mu_k$ as before. By Lemma B.2,

$$\omega\Bigg(\bar{\mu}(N), \sum_{k=1}^{N} \phi_k(N) p_k^+\Bigg) \geq \sum_{k=1}^{N} \phi_k(N)\omega\big(\mu_k, p_k^+\big) \tag{A.43}$$

if $s \leq \hat{s}$ for some $\hat{s} > 0$. Additionally, the LHS of (A.43) is bounded from above. That is, $\partial_p \omega(\mu, p) < 0$ as well as $\tilde{p}(N) < \bar{p} = \sum_{k=1}^{N} \phi_k(N) p_k^+$ by Theorem 3 imply

$$\omega\left(\bar{\mu}(N), \tilde{p}(N)\right) > \omega\Bigg(\bar{\mu}(N), \sum_{k=1}^{N} \phi_k(N) p_k^+\Bigg).$$

Combining the inequality above with (A.43) proves that $W_{NT}(N) > W_T(N)$. By assumption, $s \leq s^*$ such that $\mathcal{K}_{NT} = N$ always holds. In addition, define $s_N^T := \max\{s \geq 0 : \omega(\mu_N, p_N^+) \geq 0\}$ such that $\omega(\mu_k, p_k^+) \geq 0$ for all $k \leq N$ if $s \leq s_N^T$. Then, $W_T(\mathcal{K}_T) < W_T(N)$ for all $\mathcal{K}_T \leq N$ if $s \leq s_N^T$. Consequently, $s \leq \min\{s^*, \hat{s}, s_N^T\}$ guarantees $W_{NT}(N) > W_T(\mathcal{K}_T)$ for all $\mathcal{K}_T \leq N$.

The following lemmata are used in the proofs of Theorem 3 and Proposition 2.

Lemma A.11 *Define* $\pi(\mu, p) := p\left((1-\mu)\left[1 - F\big(p\big)\right] + \mu\left[1 - F\big(p + s/m\big)\right]\right)$, $p_k^* := \arg\max_p \left(\pi(\mu_k, p)\right)$ *and* $p^*(\mathcal{K}) := \arg\max_p \left(\pi(\bar{\mu}(\mathcal{K}), p)\right)$. *Then,* $p_k^+ > p_k^*$ *if* $k > 1$ *and* $\tilde{p}(\mathcal{K}) > p^*(\mathcal{K})$ *if* $\mathcal{K} \geq 1$, *where* p_k^+ *and* $\tilde{p}(\mathcal{K})$ *denote the equilibrium prices with and without tracking.*

PROOF. Suppose that $p_k^+ \leq p_k^*$ and recall that p_k^+ satisfies

$$p_k^+ = \psi_k(p_k^+) \geq \psi_k(p_k^*) > \frac{(1-\mu)\big[1 - F(p_k^*)\big] + \mu\big[1 - F(p_k^* + s/m)\big]}{(1-\mu) f(p_k^*) + \mu(f(p_k^* + s/m)}, \tag{A.44}$$

where the first inequality follows from Lemma A.5 and the second from the fact that $k > 1$ implies $\mu_k > 0$. Besides, the last term has to equal p_k^* due to $p_k^* = \partial_p \pi(\mu_k, p_k^*)$. This implies $p_k^+ > p_k^*$, which is a contradiction. To prove that $p^*(\mathcal{K}) < \tilde{p}(\mathcal{K})$, replace μ_k by $\bar{\mu}(\mathcal{K})$ and note that $\bar{\mu}(\mathcal{K}) > 0$ for all $\mathcal{K} \geq 1$.

Lemma A.12 *When holding* μ *constant,* $\pi(\mu, p)$ *is strictly concave in* p *over* $p \in [0, \bar{v}]$ *for any* $\mu < 1$ *if* $\alpha(p) < 2$.

PROOF. The statement holds if $\partial_p^2 \pi(\mu, p) < 0$ or equivalently if

$$-(1-\mu)\Big(2f(p) + f'(p)p\Big) - \mu\Big(2f(p+s/m) + f'(p+s/m)p\Big) < 0.$$

The first term is negative for any $\mu < 1$ because of $\alpha(p) < 2$ and it remains to show that $2f(p+s/m) + f'(p+s/m)p \geq 0$. The latter inequality is implied by

$$\Leftrightarrow -p\frac{f'(p+s/m)}{f(p+s/m)} \leq 2 \tag{A.45}$$

Again, since $\alpha(p) < 2$, we know that $-(p+s/m)\frac{f'(p+s/m)}{f(p+s/m)} < 2$ holds. Thus, if $f'(p+s/m) \geq 0$, the LHS is negative and (A.45) holds with strict inequality. Else if $f'(p+s/m) < 0$, $-(p)\frac{f'(p+s/m)}{f(p+s/m)} < -(p+s/m)\frac{f'(p+s/m)}{f(p+s/m)}$ and (A.45) holds with strict inequality as well.

PROOF OF THEOREM 5. Denote industry profits with tracking for a given $\mathcal{K}_T$ by $\Pi_T\big(\mathcal{K}_T\big)$. Using the definition of $\pi(\mu, p)$ in the main text, $\Pi_T(N)$ can be written as

$$\Pi_T(N) = \sum_{k=1}^{N}(1-m)^{k-1}\pi\big(\mu_k, p_k^+\big) = \frac{1-(1-m)^N}{m}\sum_{k=1}^{N}\phi_k(N)\pi\big(\mu_k, p_k^+\big).$$

Likewise, profits without tracking when $\mathcal{K}_{NT} = N$ are given by

$$\Pi_{NT}(N) = \frac{1-(1-m)^N}{m}\sum_{k=1}^{N}\phi_k(N)\pi\big(\mu_k, \tilde{p}(N)\big) = \frac{1-(1-m)^N}{m}\pi\left(\bar{\mu}(N), \tilde{p}(N)\right).$$

where $\bar{\mu}(N) = \sum_{k=1}^{N}\phi_k(N)\mu_k$ as before. By Lemma B.3,

$$\pi\left(\bar{\mu}(N), \sum_{k=1}^{N}\phi_k(N)p_k^+\right) \geq \sum_{k=1}^{N}\phi_k(N)\pi(\mu_k, p_k^+). \tag{A.46}$$

Additionally, the LHS of (A.46) is bounded from above. To see this, note that $\pi(\mu, p)$ is concave in p by Lemma A.12 and that $\arg\max_p \pi\left(\bar{\mu}(N), p\right) > \tilde{p}(N)$ by

Lemma A.11. Thus, using $\sum_{k=1}^{N} \phi_k(N) p_k^+ > \tilde{p}(N)$ from Theorem 3, I obtain that

$$\pi\Big(\bar{\mu}(N), \tilde{p}(N)\Big) > \pi\Big(\bar{\mu}(N), \sum_{k=1}^{N} \phi_k(N) p_k\Big).$$

Combining the inequality above with (A.46) proves that $\Pi_{NT}(N) > \Pi_T(N)$. Since $s \leq s^*$ by assumption, $\mathcal{K}_{NT} = N$ such that industry profits without tracking always equal $\Pi_{NT}(N)$. Additionally, note that $\pi(\mu_k, p_k^+) \geq 0$ for all $k \leq N$ without further assumptions, implying that $\Pi_{NT}(N) > \Pi_T(\mathcal{K}_T)$ for all $\mathcal{K}_T \leq N$.

DERIVATION of Demand in Section 6.1. Expected demand per consumer conditional on observing an opt-out decision for any $p \leq \hat{p}^e(\hat{k}) + s/m$ is given by

$$\begin{aligned}\frac{\Phi(N)m}{N} \sum_{k=\hat{k}}^{N} \frac{\phi_k(N)}{\sum_{j=\hat{k}}^{N} \phi_j(N)} &\Big((1-\mu_k)\left[1 - F(p)\right] + \mu_k \Big[1 - F\big(\hat{p}^e(\hat{k}) + \frac{s}{m}\big)\Big]\Big) \\ = \frac{\Phi(N)m}{N} &\Big(\big(1 - \hat{\mu}(\hat{k})\big)\left[1 - F(p)\right] + \hat{\mu}(\hat{k}) \Big[1 - F\big(\hat{p}^e(\hat{k}) + \frac{s}{m}\big)\Big]\Big), \qquad \text{(A.47)}\end{aligned}$$

where $\hat{\mu}(\hat{k})$ is defined in the main text and $\Phi(N) = \sum_{j=1}^{N} (1-m)^{j-1}$ in (A.37).

PROOF of Lemma 3/ Proposition 1. The statements eventually follow from $p_{\hat{k}}^+ < \hat{p}(\hat{k})$, which is due to $\hat{\mu}(\hat{k}) > \mu_{\hat{k}}$. The proof of this implication is analogous to the argument presented in the last paragraph in the proof of Lemma A.5.

PROOF of Proposition 2. Suppose that $n > 1$ sellers do not track and let $\pi_{nt}(n)$ denote a non-tracking seller's expected profits in that case. Analogously, let $\pi_t(n)$ denote a tracking seller's profits when there are $N - n$ of them. Further, let $\hat{p}(n)$ denote the price chosen by non-tracking sellers conditional on when consumers sample non-tracking sellers.

If n sellers do not track, a consumer samples the first non-tracking seller during her $N - n + 1$th search attempt. Thus, the equation characterizing $\hat{p}(n)$ can be obtained by replacing $\hat{k}$ in $\hat{\mu}(\hat{k})$ by $N-n+1$ and $\hat{p}(\hat{k})$ by $\hat{p}(n)$ in (24), the condition characterizing the price without tracking in Section 6.1. Since $\hat{\mu}(N-n+1) > \mu_k$ for all $k \leq N-n+1$, we know that $\hat{p}(n) > p_k^+$ for all $k \leq N-n+1$, which is why consumers sample tracking sellers first. Moreover, Lemma A.11 and the concavity of π in p implied by Lemma A.12 (due to $\alpha(p) < 2$) thus lead to

$$\pi\big(\mu_k, p_k^+\big) > \pi\big(\mu_k, \hat{p}(n)\big) \text{ for all } k \leq N - n + 1. \tag{A.48}$$

Using $\pi(\mu, p)$ as defined in the main text allows me to write $\pi_{nt}(n)$ as

$$\pi_{nt}(n) = \frac{1}{n} \sum_{k=N-n+1}^{N} (1-m)^{k-1} \pi\big(\mu_k, \hat{p}(n)\big). \tag{A.49}$$

In contrast, if one of the non-tracking sellers deviates to tracking, he earns

$$\pi_t(n-1) = \frac{1}{N-n+1} \sum_{k=1}^{N-n+1} (1-m)^{k-1} \pi\big(\mu_k, p_k^+\big). \tag{A.50}$$

Further, $1 - F\big(p + s/m\big) < 1 - F\big(p\big)$ implies that $\pi(\mu, p)$ is decreasing in μ. Thus, $\pi\big(\mu_k, \hat{p}(n)\big) > \pi\big(\mu_j, \hat{p}(n)\big)$ for all $k \leq N - n$ and $j \geq N - n + 1$. By combining the latter inequality with (A.48), it follows that

$$\frac{1}{N-n+1} \sum_{k=1}^{N-n+1} \pi\big(\mu_k, p_k^+\big) \geq \frac{1}{n} \sum_{k=N-n+1}^{N} \pi\big(\mu_k, \hat{p}(n)\big). \tag{A.51}$$

with strict inequality since $n > 1$. Given (A.49) and (A.50), $\pi_{nt}(n) < \pi_t(n-1)$ obtains because $(1-m)^{k-1} > (1-m)^{j-1}$ for all $k \leq N - n$ and $j \geq N - n + 1$.

References

Alessandro Acquisti and Hal R Varian. Conditioning prices on purchase history. *Marketing Science*, 24(3):367–381, 2005.

Klaus Adam. Learning while searching for the best alternative. *Journal of Economic Theory*, 101(1):252–280, 2001.

Inaki Aguirre, Simon Cowan, and John Vickers. Monopoly price discrimination and demand curvature. *American Economic Review*, 100(4):1601–15, 2010.

Simon P Anderson and Regis Renault. Pricing, product diversity, and search costs: A bertrand-chamberlin-diamond model. *The RAND Journal of Economics*, pages 719–735, 1999.

Simon P Anderson and Régis Renault. Consumer information and firm pricing: negative externalities from improved information. *International economic review*, 41(3):721–742, 2000.

Maria Arbatskaya. Ordered search. *The RAND Journal of Economics*, 38(1): 119–126, 2007.

Mark Armstrong. Ordered consumer search. *Journal of the European Economic Association*, 15(5):989–1024, 2017.

Mark Armstrong and Jidong Zhou. Paying for prominence. *The Economic Journal*, 121(556), 2011.

Mark Armstrong and Jidong Zhou. Search deterrence. *The Review of Economic Studies*, 83(1):26–57, 2016.

Mark Armstrong, John Vickers, and Jidong Zhou. Prominence and consumer search. *The RAND Journal of Economics*, 40(2):209–233, 2009.

Mark Bagnoli and Ted Bergstrom. Log-concave probability and its applications. *Economic theory*, 26(2):445–469, 2005.

Simon Board and Marek Pycia. Outside options and the failure of the coase conjecture. *American Economic Review*, 104(2):656–71, 2014.

Tomasz Bujlow, Valentín Carela-Español, Josep Sole-Pareta, and Pere Barlet-Ros. A survey on web tracking: Mechanisms, implications, and defenses. *Proceedings of the IEEE*, 105(8):1476–1510, 2017.

Yongmin Chen and Michael H Riordan. Price and variety in the spokes model. *The Economic Journal*, 117(522):897–921, 2007.

Michael Choi, Anovia Yifan Dai, and Kyungmin Kim. Consumer search and price competition. *Econometrica*, 86(4):1257–1281, 2018.

Ronald H Coase. Durability and monopoly. *The Journal of Law and Economics*, 15(1):143–149, 1972.

Vincent Conitzer, Curtis R Taylor, and Liad Wagman. Hide and seek: Costly consumer privacy in a market with repeat purchases. *Marketing Science*, 31(2): 277–292, 2012.

Peter A Diamond. A model of price adjustment. *Journal of Economic Theory*, 3 (2):156–168, 1971.

Daria Dzyabura and John R Hauser. Recommending products when consumers learn their preference weights. *Marketing Science*, 2019.

Drew Fudenberg and Jean Tirole. Customer poaching and brand switching. *RAND Journal of Economics*, pages 634–657, 2000.

Marco A Haan and José L Moraga-González. Advertising for attention in a consumer search model. *The Economic Journal*, 121(552):552–579, 2011.

Marco A Haan, José L Moraga-González, and Vaiva Petrikaitė. A model of directed consumer search. *International Journal of Industrial Organization*, 61:223–255, 2018.

Oliver D Hart and Jean Tirole. Contract renegotiation and coasian dynamics. *The Review of Economic Studies*, 55(4):509–540, 1988.

Andreas Hornstein, Per Krusell, and Giovanni L Violante. Frictional wage dispersion in search models: A quantitative assessment. *American Economic Review*, 101(7):2873–98, 2011.

Justin P Johnson. Targeted advertising and advertising avoidance. *The RAND Journal of Economics*, 44(1):128–144, 2013.

Daniel Kahneman and Richard H Thaler. Anomalies: Utility maximization and experienced utility. *Journal of Economic Perspectives*, 20(1):221–234, 2006.

Alan B Krueger and Andreas I Mueller. A contribution to the empirics of reservation wages. *American Economic Journal: Economic Policy*, 8(1):142–79, 2016.

Tai-Ching Li, Huy Hang, Michalis Faloutsos, and Petros Efstathopoulos. Trackadvisor: Taking back browsing privacy from third-party trackers. In *International Conference on Passive and Active Network Measurement*, pages 277–289. Springer, 2015.

Eeva Mauring. Learning from trades. *The Economic Journal*, 127(601):827–872, 2017.

Paul Milgrom and John Roberts. Relying on the information of interested parties. *The RAND Journal of Economics*, pages 18–32, 1986.

José L Moraga-González and Vaiva Petrikaitė. Search costs, demand-side economies, and the incentives to merge under bertrand competition. *The RAND Journal of Economics*, 44(3):391–424, 2013.

Andrew Rhodes. Can prominence matter even in an almost frictionless market? *The Economic Journal*, 121(556):F297–F308, 2011.

Babur De Los Santos, Ali Hortaçsu, and Matthijs R Wildenbeest. Search with learning for differentiated products: Evidence from e-commerce. *Journal of Business & Economic Statistics*, 35(4):626–641, 2017.

Johannes F Schmieder, Till von Wachter, and Stefan Bender. The effect of unemployment benefits and nonemployment durations on wages. *American Economic Review*, 106(3):739–77, 2016.

Dale O Stahl. Oligopolistic pricing with sequential consumer search. *The American Economic Review*, pages 700–712, 1989.

Curtis R Taylor. Consumer privacy and the market for customer information. *RAND Journal of Economics*, pages 631–650, 2004.

J Miguel Villas-Boas. Dynamic competition with customer recognition. *The Rand Journal of Economics*, pages 604–631, 1999.

Asher Wolinsky. True monopolistic competition as a result of imperfect information. *The Quarterly Journal of Economics*, pages 493–511, 1986.

Jidong Zhou. Ordered search in differentiated markets. *International Journal of Industrial Organization*, 29(2):253–262, 2011.

Appendix B: Online Appendix

PROOF OF REMARK 2. Assume that $p_1 > p_2$. By hypothesis $\pi_1(p_1) \geq \pi_1(p_2)$, or equivalently $\left(\bar{\alpha}\,(1 - F(p_1)) + \underline{\beta}(p_1)\right) p_1 \geq \left(\bar{\alpha}\,(1 - F(p_2)) + \underline{\beta}(p_2)\right) p_2$. As $\underline{\beta}$ is weakly decreasing, it follows that

$$\left(\bar{\alpha}\,(1 - F(p_1)) + \underline{\beta}(p_1)\right) p_1 \geq \left(\bar{\alpha}\,(1 - F(p_2)) + \underline{\beta}(p_1)\right) p_2,$$

or equivalently $\underline{\beta}(p_1)(p_1 - p_2) \geq \bar{\alpha}((1-F(p_2))p_2 - (1-F(p_1))p_1)$. As $(p_1 - p_2) > 0$ implies $\bar{\beta}(p_1 - p_2) > \underline{\alpha}\left(\left(1 - F(p_2)\right)p_2 - \left(1 - F(p_1)\right)p_1\right)$, or equivalently $\pi_2(p_1) > \pi_2(p_2)$, we have a contradiction.

It will be necessary to recall some notions from directional derivatives. Let $\mathbf{z} \in \mathbb{R}^2$ be a vector. The linear subspace in direction $\mathbf{z}$ is the vector space

$$L_z := \{\mathbf{x} \in \mathbb{R}^2 \;:\; \mathbf{x} = \lambda\mathbf{z} \text{ for some } \lambda \in \mathbb{R}\}.$$

The affine subspace at $\mathbf{x}$ in parallel to L_z is the set $L_z(\mathbf{x}) = \mathbf{x} + L_z$, i.e. a translation of the subspace in direction $\mathbf{z}$ by $\mathbf{x}$. Recall that the directional derivative along $\mathbf{z}$ at $\mathbf{x}$, $D_z f(\mathbf{x})$ is the derivative of the restriction of f to the one-dimensional affine subspace $L_z(\mathbf{x})$. Moreover, for a function $f \in \mathcal{C}^2(\mathbb{R}^2)$ the directional derivative along $\mathbf{z}$ can be represented as $D_z f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{z}$, where ∇f is the usual gradient. With this terminology in place, I call a function *concave in direction* $\mathbf{z}$ if its restriction to any one-dimensional affine subspace in direction $\mathbf{z}$ is a concave function. More precisely, f defined on a convex subset C is concave along $\mathbf{z}$ if for all $\mathbf{x}_0 \in C$, $\mathbf{x}_1, \mathbf{x}_2 \in L_z(\mathbf{x}_0) \cap C$ and $\lambda \in [0,1]$ it holds that

$$f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \geq \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2).$$

Note in particular that $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in L_z(\mathbf{x})$ since this is an affine subspace. Evidently a function is concave along $\mathbf{z}$ if and only if $(\mathbf{z}')^T H_f(\mathbf{x})\mathbf{z}' \leq 0$ for all $\mathbf{z}' \in L_z(\mathbf{x})$, which is equivalent to $\lambda^2(\mathbf{z})^T H_f(\mathbf{x})\mathbf{z} \leq 0$ for all $\lambda \in \mathbb{R}$, where $H_f(\mathbf{x})$ is the Hessian of $f(x)$. This can be easily seen from a Taylor expansion.

Finally observe that also Jensen's inequality holds for any family of points in

$L_z(\mathbf{x})$ whenever f is concave along $\mathbf{z}$.

Lemma B.1 (Jensen's Inequality) *Suppose $Z \subset \mathbb{R}^2$ is a closed and convex set and let $G \in \mathcal{C}^2$ be concave along all directions $\mathbf{z}$ in Z. Let $N \in \mathbb{N}$ and $\mathbf{x}_1, \ldots, \mathbf{x}_k$ be N points in $\mathbb{R}^2$. If $\mathbf{x}_k - \mathbf{x}_j \in Z$ for all $N \geq k > j \geq 1$, then*

$$G\left(\sum_{k=1}^{N} \phi_k \mathbf{x}_k\right) \geq \sum_{k=1}^{N} \phi_k G(\mathbf{x}_k), \tag{B.1}$$

is valid whenever $\phi_k \in [0,1]$ are such that $\sum_{k=1}^{N} \phi_k = 1$.

PROOF. The proof is by induction. For $N = 2$, the statement holds by the remark above. Now let $N \geq 3$, fix $\mathbf{x}_1, \ldots, \mathbf{x}_N$ and Z, and suppose the statement holds for the first $N-1$ points. Consider $\rho_k \in (0,1)$ such that $\sum_{k=1}^{N} \rho_k = 1$. Using the induction hypothesis (B.1) for $N-1$, the inequality

$$(1-\rho_N) G\left(\frac{1}{1-\rho_N} \sum_{k=1}^{N-1} \rho_k \mathbf{x_k}\right) + \rho_N G(\mathbf{x_N}) \geq \sum_{k=1}^{N-1} \rho_k G(\mathbf{x_k}) + \rho_N G(\mathbf{x_N}) \tag{B.2}$$

is easily verified to hold. Since $\mathbf{x_k} - \mathbf{x_j}$ is in Z for all $0 \leq j < k \leq N$, also

$$\mathbf{z}'(\rho_1, ..., \rho_N) := \mathbf{x_N} - \sum_{k=1}^{N-1} \frac{\rho_k}{1-\rho_N} \mathbf{x}_k = \sum_{k=1}^{N-1} \frac{\rho_k}{1-\rho_N} \big(\mathbf{x_N} - \mathbf{x_k}\big)$$

is in Z by convexity. Therefore $G(\mathbf{x})$ is concave in direction $\mathbf{z}'(\rho_1, ..., \rho_N)$ and by definition the left-hand side of (B.2) satisfies

$$G\left((1-\rho_N) \sum_{k=1}^{N-1} \frac{\rho_k}{1-\rho_N} \mathbf{x_N} + \rho_n \mathbf{x_N}\right) \geq (1-\rho_N) G\left(\sum_{k=1}^{N-1} \frac{\rho_k}{1-\rho_N} \mathbf{x_k}\right) + \rho_N G(\mathbf{x_N}).$$

Upon collecting terms and using (B.2) the desired inequality (B.1) follows.

Lemma B.2 *Let $\{(\mu_k, p_k^+)\}_{k \leq N}$ denote the set of equilibrium informedness-price points under tracking where the index k represents a seller's position in the* *search process. Then, there exists a threshold $\hat{s} > 0$ such that:*

$$\omega\Bigg(\sum_{k=1}^{N} \phi_k(N)\mu_k, \sum_{k=1}^{N} \phi_k(N)p_k^+\Bigg) \geq \sum_{k=1}^{N} \phi_k(N)\omega(\mu_k, p_k^+) \text{ for all } s \leq \hat{s}.$$

PROOF. I show that ω is concave along the relevant directions. To this end define

$$Z := \text{conv}\{(z_\mu, z_p) = (\mu_k - \mu_j, p_k^+ - p_j^+) \ : \text{ for any } j \neq k, \ k > j\}.$$

Note that all vectors in Z are bounded below by a vector (ϵ, ϵ) for some $\epsilon > 0$. Now consider the Hessian of $\omega(\mu, p)$:

$$H_\omega(\mu, p) = \begin{bmatrix} a(\mu, p) & c(\mu, p) \\ c(\mu, p) & 0 \end{bmatrix}$$

with

$$\begin{aligned} a(\mu, p) &= -(1-\mu)m\left[f(p) + pf'(p)\right] - \mu m\left[f\left(p + \frac{s}{m}\right) + pf'\left(p + \frac{s}{m}\right)\right] \\ c(\mu, p) &= mp\left[f(p) - f\left(p + \frac{s}{m}\right)\right]. \end{aligned}$$

Because the Hessian is not negative semi-definite, ω is not globally concave. Yet, by the preceding discussion ω is concave in any direction $\mathbf{z} \in Z$ if and only if

$$\lambda^2\Big(a(\mu, p)z_p^2 + 2c(\mu, p)z_\mu z_p\Big) \leq 0 \quad \text{if and only if} \quad \lambda^2\Big(a(\mu, p) + 2c(\mu, p)\frac{z_\mu}{z_p}\Big) \leq 0$$

for all λ and $\mathbf{z} = (z_\mu, z_p) \in Z$. Note that

$$\lim_{s \to 0}\left[a(\mu, p) + 2c(\mu, p)\frac{z_\mu}{z_p}\right] = \Big(-m\big(f(p) + pf'(p)\big) + 0\Big) < 0, \qquad \text{(B.3)}$$

where the last inequality follows from the assumption that $\alpha(p) < 1$. By the continuity of a and c, and since $z_\mu/z_p \leq z_\mu/\epsilon$ is bounded for all $\mathbf{z} \in Z$, there exists $\hat{s} > 0$ such that $a(\mu, p)(z_p')^2 + 2c(\mu, p)z_\mu' z_p' < 0$ holds for all $s \leq \hat{s}$ and $(z_\mu', z_p') \in Z$. Consequently ω is concave in all directions in Z. Because $\sum_{k=1}^{N} \phi_k(N) = 1$, the claim now follows from applying Lemma B.1.

Lemma B.3 *Let* $\{(\mu_k, p_k^+)\}_{k\leq N}$ *denote the set of equilibrium informedness-price points under tracking such that the index* k *represents a seller's position in the search process. Then,*

$$\pi\Bigg(\sum_{k=1}^{N}\phi_k(N)\mu_k, \sum_{k=1}^{N}\phi_k(N)p_k^+\Bigg) \geq \sum_{k=1}^{N}\phi_k(N)\pi(\mu_k, p_k^+).$$

PROOF. I show that π is concave along the relevant directions. Define

$$Z := \text{conv}\{(z_\mu, z_p) = (\mu_k - \mu_j, p_k^+ - p_j^+) \ : \ \text{for any } j \neq k,\ k > j\}$$

and note that all vectors in Z are coordinate-wise non-negative. Now consider the Hessian of $\pi(\mu, p)$:

$$H_\pi(\mu, p) = \begin{bmatrix} r(\mu, p) & t(\mu, p) \\ t(\mu, p) & 0 \end{bmatrix}$$

with

$$r(\mu, p) = -(1-\mu)\Big(2f(p) + f'(p)p\Big)m - \mu\Big(2f(p + s/m) + f'(p + s/m)p\Big)m$$
$$t(\mu, p) = -\Big(\big[1 - F(p)\big] - f(p)p\Big)m + \Big(\big[1 - F(p + s/m)\big] - f(p + s/m)p\Big)m.$$

Again, $H_\pi(\mu, p)$ is not negative semi-definite and, thus, $\pi(\mu, p)$ is not concave in general. However, by the preceding discussion π is concave in any direction $\mathbf{z} \in Z$ if and only if

$$\lambda^2\left(r(\mu, p)z_p^2 + 2t(\mu, p)z_\mu z_p\right) \leq 0 \tag{B.4}$$

for all λ and $\mathbf{z} = (z_\mu, z_p) \in Z$. Note that by the assumption of $\alpha(p) < 1$, Lemma A.12 implies that $r(\mu, p) < 0$ for all $\mu < 1$.

Moreover $t(\mu, p) \leq 0$. To see this note that $t(\mu, p, s)$ is continuous and differentiable in s and that $t(\mu, p, s) = 0$ if $s = 0$. Hence, $t(\mu, p, s) \leq 0$ for all $s > 0$ if $\partial_s t(\mu, p, s) \leq 0$. Taking the derivative with respect to s yields

$\partial_s t(\mu, p, s) = -f(p + s/m) - pf'(p + s/m)$. Thus, $\partial_s t(\mu, p, s) \leq 0$ if and only if

$$-p\frac{f'(p + s/m)}{f(p + s/m)} \leq 1. \tag{B.5}$$

Inequality (B.5) is equivalent to (A.45) in the proof of Lemma A.12, except that the LHS equals 1 here. By the same argument as in Lemma A.12, it can thus be verified that $\alpha(p) < 1$ implies (B.5). Because $z_p z_\mu \geq 0$ for all $(z_\mu, z_p) \in Z$, inequality (B.4) is valid for all $\mathbf{z} \in Z$. The claim now follows from applying the version Jensen's inequality provided in Lemma B.1 as $\sum_{k=1}^{N} \phi_k(N) = 1$.