Northwestern | Kellogg

CMS-EMS

Center for Mathematical Studies in Economics and Management Sciences

Discussion Paper #1597

"Searching for "Arms": Experimentation with Endogenous Consideration Sets"

Daniel Fershtman
Tel Aviv University

Alessandro Pavan
Northwestern University

May 2023

Searching for "Arms": Experimentation with Endogenous Consideration Sets*

Daniel Fershtman† Alessandro Pavan‡

May 2023

Abstract

A decision-maker alternates between exploring alternatives in the consideration set and expanding the latter. When the expansion technology is stationary, or improving, alternatives are replaced at each expansion. When, instead, it deteriorates, alternatives are revisited, and each expansion is treated as the last one. The model endogenizes the set of boxes in Weitzman's (1979) Pandora's problem. When applied to online consumer search, it (a) endogenizes click-through rates and values per click in sponsored search auctions, (b) explains the phenomenon of non-sequential-non-cascading clicking, and (c) illustrates why the generalized second-price auction may lead to inefficient assignments even under its ascending-clock implementation.

Keywords: Search, Experimentation, Learning, Endogenous Consideration Sets

*The paper supersedes previous versions circulated under the titles "Searching for Arms" and "Sequential Learning with Endogenous Consideration Sets." We are grateful to various participants at conferences and workshops where the paper was presented, including Bar-Ilan University, University of Bonn, Cornell University, Emory University, Haifa University, Hebrew University, HEC Paris, University of Miami, Paris Dauphine, University of Pennsylvania, University of Pittsburgh, Princeton University, Simon Fraser University, Stanford GSB, Tel-Aviv University, Toulouse School of Economics, Yale University, Warwick Economic Theory Workshop, 3rd Columbia Conference on Economic Theory (CCET), and the 30th Stony Brook International Conference on Game Theory. Dirk Bergemann, Eddie Dekel, David Dillenberger, Laura Doval, Kfir Eliaz, Stephan Lauermann, Benny Moldovanu, Xiaosheng Mu, Philip Reny, Eran Shmaya, Andy Skrzypacz, Rani Spiegler, Bruno Strulovici, Asher Wolinsky, and Jidong Zhou provided very useful comments and suggestions. We also thank Matteo Camboni and Tuval Danenberg for outstanding research assistance. Fershtman gratefully acknowledges funding from ISF grant #1202/20. The usual disclaimer applies.

†Eitan Berglas Scheool of Economics, Tel Aviv University. Email: danielfer@tauex.tau.ac.il

‡Department of Economics, Northwestern University. Email: alepavan@northwestern.edu

1 Introduction

Classic models of sequential experimentation or learning involve a decision-maker (hereafter, DM) exploring a *fixed* set of alternatives with unknown returns. Yet, a ubiquitous feature of many dynamic decision problems is that the set of alternatives a DM can explore is expanded over time, in response to the information gathered by exploring the alternatives already in the consideration set (hereafter, CS).

In this paper, we study the tradeoff between the exploration of alternatives already in the CS and the expansion of the latter through search for additional alternatives. A key difference between exploration and expansion is the direct vs indirect nature of the two activities. When an alternative is in the CS, the DM can "point to it," that is, she can choose to explore that particular alternative instead of others. When, instead, an alternative is outside the CS, the DM cannot point to it, meaning that she cannot choose to explore that specific alternative instead of others.[1] This inability may reflect natural randomness in the search process, which may bring to the CS alternatives different from those the DM was looking for. Alternatively, search may bring more than a single alternative and such batching may have implications for the decision to expand the CS in the first place. Finally, the DM may have limited knowledge about the alternatives outside the CS, and/or her ability to bring new alternatives to it, and may revise her beliefs about the "search technology" based on the results of past searches.

To study the tradeoff between exploration of alternatives already in the CS and expansion of the latter, we consider a generalization of the classic multi-armed bandit problem in which the set of "arms" is *endogenous*. Exploring an alternative already in the CS (pulling an arm) yields a flow payoff and generates information (for example, about the distribution from which the flow payoff is drawn). Searching for new "arms" (that is, choosing to expand the CS) is costly and brings a random set of new alternatives (i.e., of arms).

The solution to the above problem takes the form of an "index" policy. Each alternative in the CS is assigned a history-dependent number that is a function only of the state of that alternative. This number (the arm's "index") is the same as in Gittins and Jones' (1974) original work on bandit problems with an exogenous set of arms. Search (that is, the decision to expand the CS) is also assigned an index, which depends only on the state of the search technology. Crucially, the search index does not depend on the information generated by the exploration of any of the alternatives already in the CS. It also differs from the value the DM attaches to the expansion of the CS but is linked to the indexes of the new alternatives the DM expects to find through current and future searches. The optimal policy consists in selecting at any period the alternative for which the index is the highest.

[1] Likewise, the DM cannot choose to bring a specific alternative from outside of the CS into the CS: If she could, there would be no distinction between exploring alternatives inside and outside the CS, making the latter irrelevant.

Our environment can be viewed as a special case of the branching problem in the operations-research literature, where certain arms, after being activated, branch into new ones and then disappear (see, e.g., Weiss, 1988, Weber, 1992, and Keller and Oldale, 2003). In our problem, the decision to expand the CS corresponds to the activation of a branching arm that yields negative rewards (in the form of search costs) and brings a stochastic set of new arms according to a distribution that depends on the results of past searches.

Our proof of indexability is based on a novel recursive characterization of the search index which facilitates its computation and permits us to uncover various properties of the dynamics of exploration and CS-expansion that are relevant for economic applications. At any point in time, the decision to search for new alternatives depends on the current CS only through (a) the value of the alternative with the highest index, and (b) the state of the search technology. This property holds despite the fact that the opportunity cost of searching for new alternatives (which is linked to the value of continuing with the current CS) depends on the entire composition of the current CS. Similarly, conditional on forgoing search in a given period, the decision of which alternative to explore in the current CS is independent of the search technology, despite the fact that search may bring alternatives that are more similar to certain alternatives currently in the CS than others.[2] If the search technology is stationary, or improving, in a sense made precise below, then alternatives in the CS at the time of its expansion never receive attention in the future, and hence are effectively discarded once the CS is expanded. Each search is then equivalent to replacement of the current CS with a new one. When, instead, the search technology deteriorates over time (e.g., because the DM becomes pessimistic about the possibility of finding attractive new alternatives), the alternatives in the current CS are put on hold and may be revisited after the CS is expanded. Furthermore, in this case, the decision to expand the CS is made as if there will be no further expansions after the current one.

The analysis can be applied to a broad class of experimentation and/or sequential learning problems of interest in economics. In the paper, we first show how the model permits one to endogenize the set of boxes in Weitzman's (1979) "Pandora's problem" where the index of search has a natural interpretation in terms of the prizes of the various types of boxes that the DM expects to bring to the CS. We then apply the results to the problem of a consumer searching online for new products. The model permits us to endogenize both the probability with which any given ad is clicked, and a firm's value per click (with the latter varying with the position the ad occupies) that result from a consumer's search. The results also provide an explanation for clicking patterns that have been observed in empirical work but that are

[2] These properties can be seen as a generalization of the IIA (independence of irrelevant alternatives) property of classic multi-armed bandit problems. What makes this problem different from the classic one enriched with a "meta" arm that comprises all the alternatives brought to the CS by search is that the evaluation of such a "meta" arm requires knowing how to subsequently explore the arms that search brings to the CS, which is what is investigated in the first place. Furthermore, dynamic problems with "meta" arms rarely admit an index solution. See also the discussion in footnote 22.

inconsistent with the properties of existing models used in the literature. Finally, the analysis has important implications for the efficiency of the allocations induced by the auctions typically used to allocate ad space, as discussed below.

Pandora's problem with an endogenous set of boxes. In Weitzman (1979)'s problem, the DM faces an exogenous set of boxes, each containing a prize of unknown value drawn from a known distribution. Opening a box reveals its prize, is costly (with the cost box-specific), and is necessary to collect its prize. The DM can collect only one prize among any of the opened boxes, and must choose the optimal sequence of inspections as well as when to stop. We extend this problem by allowing the DM, at any period, to choose between opening a box among those in the CS, or expanding the latter by searching for new boxes. The outcome of each search is stochastic and brings to the CS a new set of boxes whose characteristics are unknown at the time of the expansion but may depend on the composition of the current CS. Our solution generalizes Weitzman (1979)'s by introducing an appropriate reservation price for each expansion of the CS, which we characterize using our recursive description of the search index.

Online consumer search. As an illustration of the usefulness of the model and its key results in concrete economic applications, we consider the problem of a consumer alternating between (a) reading new ads (bringing the corresponding products to the CS), (b) clicking on the ads of those products already in the CS (revealing the products' value to the consumer, net of the purchasing price), and (c) finalizing the purchase with one of the visited vendors. We show how our results can be used to describe the selection of the various products by means of a comparison of the products' "discovery values." The latter generalize the "effective values" of Choi, Dai and Kim (2018)'s eventual purchase theorem by accounting for the uncertainty the consumer faces about the ads that occupy the various positions. The results also provide a structural relationship between the ads' positions and their click-through rates (CTRs)—that is, the ratio of users who click on the link directing to a vendor's webpage after reading the ad displayed at a given position.

Other models used to study online consumer search typically assume that positions' CTRs are either exogenous (see, e.g., Edelman et al., 2007) or are such that consumers click on the ads in the order in which the ads are displayed, which implies that firms advertising on a platform experience externalities only from the ads displayed at higher positions than the one their ad occupies (see, e.g., Athey and Ellison, 2011). Neither assumption seems to square well with empirical findings. For example, Jeziorski and Segal (2015) show that (a) approximately half of the users do not click on ads sequentially in the order they are displayed (*non-sequential clicking),* (b) over half of the users who click more than once, click on a higher position's ad after clicking on a lower position's ad (*non-cascading clicking*), and (c) the rate at which an ad at a given position is clicked upon depends not only on the ads displayed at higher positions but also on the ads displayed at lower positions (*externalities from lower positions*). These phenomena are consistent with the dynamics in our model.[3]

The model also explains why CTRs need not be monotone in the ads' positions, even when the purchasing probabilities are. More generally, the analysis delivers a structural relationship relating the probability each product is purchased to the primitives of the problem (consumers' realized values, ads' positions, search costs, and consumer's beliefs over the attractiveness of the products displayed at the various positions). Such relationship also provides a characterization of the firms' value per click (VPC), and shows why the latter are naturally position-specific.

These results have implications for bidding in sponsored-search auctions and for the efficiency of the allocations sustained in equilibrium. For example, we use them to show that the generalized second-price auction may lead to inefficient assignments even under the ascending-price implementation considered in the literature (see, e.g., Edelman et al., 2007 and Gomes and Sweeney (2014)). The model also explains why a firm advertising on a platform may experience a decline in its profits when the probability that it displays ads for additional products at downstream positions increases. Such a decline can happen even if the extra ad is unambiguously profitable when brought exogenously to the consumer's CS.

Other applications. The model, along with the characterization of the optimal policy and its implications for the dynamics of exploration and CS expansion, can also be applied to various other problems of interest in economics in which the endogeneity of the CS is important, such as the problem of a search committee alternating between evaluating known candidates and searching for additional ones, the administration of medical products when physicians can search for new treatments after observing disappointing results with known ones, and clinical trials, where firms alternate between testing existing products and conducting basic R&D activities that may lead to the eventual discovery of new products, with all products requiring regulatory approval before they can be brought to the market.

Outline. The rest of the paper is organized as follows. The remainder of this section discusses the paper's contribution vis-a-vis the pertinent literature. Section 2 introduces the model. Section 3 characterizes the optimal policy and identifies key properties for the dynamics of experimentation and CS expansion. Section 4 contains results for the extension of Weitzman (1979)'s Pandora's boxes problem to a setting with endogenous boxes. Section 5 contains all the results for our primary application, online consumer search. Finally, Section 6 concludes with a brief discussion of and a few possible enrichments of the baseline model. All proofs are either in the Appendix at the end of the document or in the online Supplement.

[3]The endogeneity of the CS is important. A sequential search model in which the set of products is known to the consumer from the outset (such as Weitzman's Pandora's boxes model with an exogenous set of boxes) fails to deliver any structural relationship between the positions at which the ads are displayed and the corresponding CTRs.

1.1 Related literature

To the best of our knowledge, the problem studied in the present paper (where the DM alternates between exploring "arms" already in the CS and stochastically expanding the latter) is new. As mentioned above, this problem can be viewed as a special case of the branching problem in the operations-research literature (see, e.g., Weiss, 1988, Weber, 1992, and Keller and Oldale, 2003). The recursive characterization of the index for CS-expansion is new and exploits a novel classification of the alternatives into "categories," with the latter summarizing all characteristics of the arms that are relevant for the dynamics of the technology governing the expansion of the CS. This characterization permits us to arrive at a novel representation of the DM's payoff under the index policy, which we use to establish the optimality of such a policy. Importantly, the same recursive characterization is also crucial for the characterization of the dynamics of exploration and CS expansion, and favors the computation of the index. When applied to relevant economic environments, the characterization permits us to uncover new insights.

The paper is also related to the literature on experimentation and sequential learning with an exogenous CS.[4] Most closely related are Weitzman (1979), Austen-Smith and Martinelli (2018), Fudenberg, Strack and Strzalecki (2018), Gossner, Steiner and Stewart (2021), Ke and Villas-Boas (2019), and Ke, Shen, and Villas-Boas (2016). The problem studied in these papers involves a DM acquiring costly information about a set of options before stopping and choosing one of them. Related are also Che and Mierendorff (2019) and Liang, Mu, and Syrgkanis (2022). The first paper studies the optimal sequential allocation of attention to two different signal sources biased towards alternative actions. The second paper studies the dynamic acquisition of information about an unknown Gaussian state. In all of these papers, the set of alternatives is fixed ex-ante. In our model, instead, the DM expands the CS over time in response to the information she collects about the alternatives already in it. Related are also Garfagnini and Strulovici (2016) and Carnehl and Schneider (2023). The first paper considers a setting where successive (forward-looking) agents experiment with endogenous technologies; trying "radically" new technologies reduces the cost of experimenting with similar technologies, which effectively expands the set of affordable technologies.[5] The second paper studies the time-risk tradeoff of an agent who wishes to solve a problem before a deadline and allocates her time between implementing a given method and developing (and then implementing) a new one. While, at a high level, the problems examined in these papers are related to ours in that they also consider environments in which the set of alternatives expands over time, both the models and the questions addressed are different.

In Fershtman and Pavan (2021), we study "soft" affirmative-action policies in a setting in

[4] See Bergemann and Välimäki (2008) for an overview of applications of multi-armed-bandit problems in economics.

[5] Technologies are interdependent in their environment. In particular, a radically new technology is informative about the value of similar technologies.

which the candidate pool is endogenous.[6] That paper studies the effects of changes in the search technology on the recruitment of minority candidates. The present paper, instead, establishes the optimality of an index policy, provides a recursive characterization of the index of search, and shows the value of these results by applying them to the problem of online consumer search.

Section 4 extends Weitzman (1979) to a setting with an endogenous set of boxes. Despite its many applications, relatively few extensions of Weitzman's problem have been studied in the literature. Notable exceptions include Olszewski and Weber (2015), Choi and Smith (2016), and Doval (2018). In these papers, though, the set of boxes is fixed.

The application to online consumer search in Section 5 is related to recent work by Choi, Dai and Kim (2018) who derive an "eventual-purchase theorem" relating the probability each product is selected to the primitives of the environment, in a setting with a fixed CS.[7] Our contribution is in showing how the purchasing probabilities change when accounting for the (endogenous gradual resolution of the) uncertainty the consumer faces about the products occupying the various positions. The analysis in this section is also related to the literature on sponsored search (in addition to the papers by Edelman et al., 2007, Athey and Ellison, 2011, and Gomes and Sweeney, 2014, cited above, see also Edelman and Schwarz, 2010 and the references therein). As mentioned above, our contribution is in showing how the model of sequential search with endogenous CS permits one to endogenize the CTRs and the VPCs (with the latter naturally varying with the position occupied by the ads), and generates predictions about the dynamics of clicking and purchases that are consistent with the findings in the empirical literature, as reported in Jeziorski and Segal (2015). The same application is also related to independent work by Greminger (2021). While that paper focuses on the comparison between direct and indirect search, we use the model to (a) endogenize the CTRs and VPCs, (b) show how the uncertainty about the ads occupying the different positions can lead to non-sequential and non-cascading clicking, (c) derive implications for bidding and the efficiency of the equilibrium allocations of the auctions typically considered in the literature, and (d) study the effects of additional ad space on firms' profits.[8]

Finally, the paper is part of a fast-growing literature on CS.[9] Eliaz and Spiegler (2011) study implications of different CS on firms' behavior, assuming such sets are exogenous. Manzini and Mariotti (2014) and Masatlioglu, Nakajima, and Ozbay (2012), instead, identify CS from choice behavior. Caplin, Dean, and Leahy (2019) provide necessary and sufficient conditions for

[6]See also Bardhi, Guo, and Strulovici (2021) for the effects of initial asymmetries across alternatives on the alternatives' long-run utilization, and their implications for minority hiring.

[7]See also Armstrong and Vickers (2015) and Armstrong (2017) for related results in settings with an exogenous CS.

[8]The model in that paper is a special version of the one in Subsection 5 in which payoffs are additively separable in an observable and an unobservable component.

[9]For the earlier marketing literature, see, e.g., Hauser and Wernerfelt (1990) and Roberts and Lattin (1991). For a survey of recent developments, see Honka et al (2019).

rationally-inattentive agents to focus on a subset of all available choices, thus endogenizing the CS. Simon (1955) considers a sequential search model in which alternatives are examined until a "satisfying" alternative is found. Caplin, Dean, and Martin (2011) show that the rule in Simon (1955) can be viewed as resulting from an optimal procedure when there are information costs. Our analysis complements the one in this literature by providing a dynamic micro-foundation for endogenous CS. Rather than committing to a CS up front and proceeding to evaluate its alternatives, the DM gradually expands the CS, in response to the results obtained from the exploration of the alternatives in the set.

2 Model

In each period $t = 0, 1, 2, ...$, the DM chooses between exploring one of the alternatives within her CS and expanding the CS by searching for additional alternatives. Exploring an alternative generates information about it and yields a (possibly negative) flow payoff. Expanding the CS yields a stochastic set of new alternatives, which are added to the CS and can be explored in subsequent periods.

Consideration sets. Denote by $C_t \equiv (0, ..., n_t)$ the period-t CS, with $n_t \in \mathbb{N}$. C_t comprises all alternatives $i = 0, ..., n_t$ that the DM can explore in period t, with the initial set $C_0 \equiv (0, ..., n_0)$ specified exogenously and with alternative 0 corresponding to the selection of the DM's outside option, yielding a payoff normalized to zero. Given C_t, expansion of the CS in period t (that is, search) brings a set of new alternatives $C_{t+1} \backslash C_t = (n_t + 1, ..., n_{t+1})$ which are added to the current CS and expand the latter from C_t to C_{t+1}.

Alternatives, categories, learning, and payoffs. Each alternative belongs to a fixed *category* $\xi \in \Xi$ that is observed by the DM when the alternative is brought to the CS.[10] A category contains information about an alternative's experimentation technology and payoff process. Let $\mu \in \mathbb{R}$ denote a fixed unknown parameter about the alternative that the DM is learning about, with μ drawn from a distribution Γ_ξ. When the DM explores the alternative, she observes a signal realization about μ. Let $m - 1 \in \mathbb{N}$ denote the number of past explorations of an alternative, and $\vartheta^{m-1} \equiv (\vartheta_s)_{s=0}^{m-1}$ its history of past signal realizations, with $\vartheta_0 \equiv \emptyset$. When the DM explores the alternative for the m-th time, she receives an additional signal ϑ_m about it, drawn from some distribution $G_\xi(\vartheta^{m-1}; \mu)$ and updates her beliefs about μ using Bayes' rule. Importantly, signal realizations are drawn independently across alternatives, given the alternatives' categories. The flow payoff u that the DM obtains from exploring an alternative from category ξ with parameter μ for the m-th time is drawn from a distribution $L_\xi(m; \mu)$. The latter does not depend on the times at which the alternative was explored – only on the number of times it was explored and the realizations of past explorations.[11]

[10] The set of categories, Ξ, is measurable and need not be finite.

[11] The assumption that ξ is observable implies that the distribution Γ_ξ from which μ is drawn is known to

The search (expansion) technology. When DM searches for the k-th time, she incurs a cost c_k and discovers alternatives of different categories. Let $E_k = (n_k(\xi) : \xi \in \Xi)$ denote the complete description of the alternatives identified through the k-th search, with $n_k(\xi) \in \mathbb{N}$ representing the number of category-ξ alternatives discovered. Let $(c_k, E_k)_{k=0}^{m-1}$ denote the history of the past $m-1$ search outcomes. Given $(c_k, E_k)_{k=0}^{m-1}$, the m-th search outcome (c_m, E_m) is drawn from a distribution $J((c_k, E_k)_{k=0}^{m-1})$ that is independent of calendar time, with $(c_0, E_0) \equiv \emptyset$. The dependence of J on the history of past search outcomes allows us to capture, for example, learning about the effectiveness of search, as well as changes in the DM's ability to find new alternatives (e.g., learning by doing and/or fatigue).

The classification of alternatives into categories allows us to keep track of all relevant information about the evolution of the search technology. In particular, it allows the outcome of each search to depend on the composition of the CS while still permitting an index characterization of the optimal policy. In an environment with an exogenous CS, categories play no role and one can simply let each alternative belong to its own category. With an endogenous CS, instead, categories permit us to identify common information among the alternatives in the CS that is responsible for the outcomes of future searches.

Objective. A *policy* χ for the decision problem described above is a rule specifying, for each period t, whether to experiment with one of the alternatives in the CS C_t or expand the latter through search. A policy χ is *optimal* if, after each period t, it maximizes the expected discounted sum $\mathbb{E}^{\chi}\left[\sum_{s=t}^{\infty} \delta^s U_s | \mathcal{S}_t\right]$ of the flow payoffs, where $\delta \in (0,1)$ denotes the discount factor, U_s denotes the flow period-s payoff (with the latter equal to the search cost in case search is conducted in period s), $\mathcal{S}_t$ denotes the state of the problem in period t (the latter specifies, for each alternative in the CS, the history of signals, along with the history of all past search outcomes; see Section 3 for the formal definition) and $\mathbb{E}^{\chi}\left[\cdot | \mathcal{S}_t\right]$ denotes the expectation under the endogenous process for the flow payoffs obtained by starting from the state $\mathcal{S}_t$ and following the policy χ at each period $s \geq t$. To guarantee that the process of the expected payoffs is well behaved, we assume that, for any t, any $\mathcal{S}_t$ and any χ, $\delta^t \mathbb{E}^{\chi}\left[\sum_{s=t}^{\infty} \delta^s U_s | \mathcal{S}_t\right] \to 0$ as $t \to \infty$.[12]

Remark. The model above describes an infinite-horizon experimentation problem (augmented by search) in which payoffs are accumulated alongside learning. However, flow payoffs and learning need not be intertwined. In Sections 4-5, we consider settings in which the DM sequentially decides between learning about alternatives in the CS and expanding the CS, until a final choice is made among the alternatives in the CS, ending the decision problem. In the online Supplement, we discuss how the results extend to a broader family of problems where the

the DM after the alternative's category ξ is learned (which occurs at the time the alternative is brought to the CS). Note, however, that the distribution $G_\xi(\vartheta^{m-1}; \mu)$ from which the m-th signal ϑ_m is drawn, as well as the distribution $L_\xi(m; \mu)$ from which the m-th reward is drawn, are not fully known to the DM because they depend on μ, which is unknown to the DM.

[12]This property is immediately satisfied if payoffs and costs are uniformly bounded; its role is to guarantee that the solution to the Bellman equation of the above dynamic program coincides with the true value function.

DM needs to irreversibly stop learning in order to be able to a accumulate rewards.

3 Optimal Policy and Key Implications

To facilitate the characterization of the optimal policy, we start by introducing the following notation. Denote by θ a generic sequence of signal realizations about an alternative; that is, θ is given by $\vartheta^m \equiv (\vartheta_s)_{s=1}^m$ for some m. Denote by $\omega^P = (\xi, \theta)$ an *alternative's state*, and by Ω^P the set of all possible states of an alternative.[13] While the category ξ is fixed, the history θ of past signal realizations changes over time as the result of the information that the DM accumulates about the alternative through past explorations. Similarly, the *state of the search technology* is given by the history of past search outcomes, that is, $\omega^S = (c_k, E_k)_{k=0}^{m-1}$ for some m. Denote the set of the possible states of search by Ω^S.

The *state of the decision problem* is given by the pair $\mathcal{S} \equiv (\omega^S, \mathcal{S}^P)$, where $\mathcal{S}^P$ is the *state of the current CS*; formally, $\mathcal{S}^P : \Omega^P \to \mathbb{N}$ is a counting function that specifies for each possible state of an alternative $\omega^P \in \Omega^P$, the number of alternatives in the CS in that state. Let $\Omega \equiv \Omega^P \cup \Omega^S$.[14] Denote by $\mathcal{S}_t$ the state of the decision problem at the beginning of period t. This representation of the decision problem keeps track of all relevant information in a parsimonious way and, as will become clear below, greatly facilitates the analysis.

Remark. The time-varying component θ of each alternative's state $\omega^P = (\xi, \theta)$ admits interpretations other than the signals about a fixed unknown parameter μ. In particular, all of our results apply to a broader class of problems where θ evolves as the result of "shocks" that need not reflect the accumulation of information. For example, such shocks may reflect endogenous variations in preferences, as in certain habit-formation or learning-by-doing models. Furthermore, because no assumptions are made on the distributions $L_\xi(m; \mu)$ and $G_\xi(\vartheta^{m-1}; \mu)$ from which the payoffs and the signals are drawn, the analysis accommodates for cases where payoffs themselves carry information, as well as cases where information arrives without any accompanying rewards.

3.1 Optimal Policy

We now characterize the optimal policy and discuss its implications for the dynamics of experimentation and CS expansion. Recall that a policy χ for the decision problem above specifies, for each period t and each period-t state $\mathcal{S}_t$, whether to experiment with one of the alternatives in the CS or expand the latter through search. Clearly, because the entire decision problem is

[13]The initial state of each alternative from category ξ, before the DM explores it, is $(\xi, \emptyset)$. The superscript P in ω^P is meant to highlight the fact that this is the state of a "physical" alternative in the CS, not the state of the search technology, or the overall state of the decision problem, defined below.

[14]Note that $\Omega^P \cap \Omega^S = \emptyset$.

time-homogeneous (independent of calendar time), so is the optimal policy.[15]

For each state ω^P of an alternative, let[16]

$$\mathcal{I}^P(\omega^P) \equiv \sup_{\tau>0} \frac{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s u_s | \omega^P\right]}{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s | \omega^P\right]}, \tag{1}$$

denote the "index" of each alternative in the CS in state ω^P, where τ denotes a stopping time (that is, a rule prescribing when to stop, as a function of the observed signal realizations), and where u_s denotes the flow payoff from the alternative's s-th exploration. The definition in (1) is equivalent to the one in Gittins and Jones (1974). As is well known, the optimal stopping rule in the definition of the index is the first period (after the one at which the index is computed) at which the index falls weakly below the value at the time the index was computed (see, e.g., Mandelbaum, 1986).

Given each state $\mathcal{S} = (\omega^S, \mathcal{S}^P)$ of the decision problem, denote the maximal index among the alternatives within the CS by $\mathcal{I}^*(\mathcal{S}^P)$.[17]

We now define an index for search (i.e., expansion of the CS). This index depends on the state $\omega^P(\xi, \theta)$ only through the alternative's category ξ, with the latter contributing to the state ω^S of the search technology.[18] Analogously to the indexes defined above, the index for search is defined as the maximal expected average discounted net payoff, per unit of expected discounted time, obtained between the current period and an optimal stopping time. Contrary to the standard indexes, however, the maximization is not just over the stopping time, but also over the rule governing the selection among the new alternatives brought to the CS by the current and further searches. Denote by τ a stopping time, and by π a rule prescribing, for any period s between the current one and the stopping time τ, either the selection of one of the new alternatives brought to the CS by search or further search. Importantly, π selects only among search and alternatives that are not already in the CS when the decision to search is made.[19]

Formally, given the state of the search technology ω^S, the index for search is defined by

$$\mathcal{I}^S(\omega^S) \equiv \sup_{\pi,\tau} \frac{\mathbb{E}^\pi\left[\sum_{s=0}^{\tau-1} \delta^s U_s | \omega^S\right]}{\mathbb{E}^\pi\left[\sum_{s=0}^{\tau-1} \delta^s | \omega^S\right]}, \tag{2}$$

where U_s denotes the flow payoff from the s-th decision taken under the rule π (with this decision

[15]That is, for any two periods t and t' such that $\mathcal{S}_t = \mathcal{S}_{t'}$, the decisions specified by the optimal policy for the two periods are the same.

[16]The expectations in (1) are under the process obtained by selecting the given alternative in all periods.

[17]Formally, $\mathcal{I}^*(\mathcal{S}^P) \equiv \max_{\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}^P(\hat{\omega}^P) > 0\}} \mathcal{I}(\omega^P)$.

[18]That is, the index depends on the state of each alternative in the CS only through the information that the latter state contains for the state ω^S of the search technology.

[19]Suppose the index for search is computed in period t when the state of the search technology is ω^S. Then, for each period $t < s < \tau$, π selects between further search and the selection of alternatives in the CS at period s that were not in the CS in period t.

taking the form of further search – in which case U_s is the stochastic cost of search – or exploration of one of the alternatives brought to the CS by searches following the one for which the index is computed, in which case U_s is the stochastic payoff associated with the exploration of the alternative), and where the expectations are under the process generated by the rule π.

Definition 1 (**Index policy**). *The **index policy** χ^* selects at each period t the option with the greatest index given the overall state $\mathcal{S}_t = (\omega^S, \mathcal{S}^P)$ of the decision problem: search if $\mathcal{I}^S(\omega^S) \geq \mathcal{I}^*(\mathcal{S}^P)$, and an arbitrary alternative with index $\mathcal{I}^*(\mathcal{S}^P)$ if $\mathcal{I}^S(\omega^S) < \mathcal{I}^*(\mathcal{S}^P)$.*[20]

Ties between alternatives are broken arbitrarily. In order to maintain consistency throughout the analysis, we assume that, when $\mathcal{I}^S(\omega^S) = \mathcal{I}^*(\mathcal{S}^P)$, search is carried out. To characterize the optimal policy, we first introduce the following notation. Let $\kappa(v) \in \mathbb{N} \cup \{\infty\}$ denote the first time at which, when the DM follows the index policy χ^*, (a) the search technology reaches a state in which its index is no greater than v, and (b) *all* alternatives in the CS – regardless of when they were introduced into it – have an index no greater than v. That is, $\kappa(v)$ is the minimal number of periods until all indexes are weakly below v ($\kappa(v) = \infty$ if this event never occurs).[21]

Let $\mathcal{V}^*(\mathcal{S}_0) \equiv (1-\delta) \sup_\chi \mathbb{E}^\chi \left[\sum_{t=0}^\infty \delta^t U_t | \mathcal{S}_0\right]$ denote the supremum expected per-period payoff the DM can attain across all feasible policies χ, given the initial state $\mathcal{S}_0$.

Theorem 1 (**Optimal policy**).

1. *The policy χ^* is optimal in the sequential experimentation problem with endogenous CS.*

2. *The index for search, as defined in (2), satisfies the following recursive representation. For any $\omega^S \in \Omega^S$,*

$$\mathcal{I}^S(\omega^S) = \frac{\mathbb{E}^{\chi^*}\left[\sum_{s=0}^{\tau^*-1} \delta^s U_s | \omega^S\right]}{\mathbb{E}^{\chi^*}\left[\sum_{s=0}^{\tau^*-1} \delta^s | \omega^S\right]}, \tag{3}$$

where τ^* is the first time (strictly after the one at which the index is computed) at which $\mathcal{I}^S$ and all the indexes of the new alternatives brought to the CS by current and subsequent searches fall weakly below the value $\mathcal{I}^S(\omega^S)$ of the search index when search was launched, and where the expectations are under the process induced by the index policy χ^*.

3. *The DM's expected (per-period) payoff under the index policy χ^* is equal to*

$$\int_0^\infty \left(1 - \mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)} | \mathcal{S}_0\right]\right) dv. \tag{4}$$

[20] Recall that $\mathcal{I}^*(\mathcal{S}^P)$ is the largest index among the alternatives in the CS.

[21] Note that between the current period and the first period at which all indexes are weakly below v, if the DM searches, new alternatives are added to the CS, in which case the evolution of their indexes is also taken into account in the calculation of $\kappa(v)$.

As in the classic multi-armed bandit problem with exogenous CS, independence across alternatives is the key assumption behind the optimality of the index policy. That is, the payoffs (and the signals) from the various alternatives are drawn independently across the alternatives, given the latter's categories, and the set of new alternatives brought to the CS at each expansion only depends on the state of the current CS through the number of alternatives from each category in the CS. Under such assumptions, the theorem establishes a generalization of the Gittins-index Theorem, according to which selecting in each period the alternative, or search, with the highest index is optimal.[22] Part (ii) characterizes the stopping time in the index of search. Such recursive representation facilitates an explicit characterization of the index in applications, and permits us to identify various properties of the dynamics of experimentation and CS expansion that are useful for comparative statics and for our proof of indexability. Finally, part (iii) offers a convenient representation of the DM's payoff under the optimal rule that can be used, among other things, to determine the DM's willingness to pay for changes in the search technology with limited knowledge about the details of the environment (see also the discussion in the next subsection).

3.2 Implications for Exploration and Expansion Dynamics

We now highlight several properties of the dynamics of exploration and CS expansion, under the optimal policy. To do so, we first describe properties the search technology may satisfy.

Definition 2 (**Search technology**). *(i) A search technology is* ***stationary*** *if, given any two states of the search technology* $\omega^S = (c_j, E_j)_{j=0}^m$ *and* $\hat{\omega}^S = (\hat{c}_j, \hat{E}_j)_{j=0}^{\hat{m}}$, $J(\omega^S) = J(\hat{\omega}^S)$. *(ii) A search technology is* ***deteriorating*** *if, given any state* $\omega^S = (c_j, E_j)_{j=0}^m$ *and subsequent state* $\hat{\omega}^S = \big((c_j, E_j)_{j=0}^m, (c_j, E_j)_{j=m+1}^{m+s}\big)$, $m, s \in \mathbb{N}$, *the distribution* $J(\omega^S)$ *first-order stochastically dominates the distribution* $J(\hat{\omega}^S)$. *(iii) A search technology is* ***improving*** *if, for any state* ω^S *and subsequent state* $\hat{\omega}^S$, *as defined in part (ii),* $J(\hat{\omega}^S)$ *first-order stochastically dominates* $J(\omega^S)$.[23]

[22]The reason why indexability of the optimal policy is not obvious is that search is a "meta" arm bringing alternatives that one needs to process optimally once brought to the CS. While our results imply that search can effectively be treated as a meta arm with its own index, the result is not a priori warranted. Indeed, problems in which alternatives correspond to meta arms, i.e., to sub-problems with their own sub-decisions (sometimes referred to as super-processes), typically do not admit an index solution, even if each sub-problem is independent from the others, and even if one knows the solution to each independent sub-problem. In the same vein, dependence, or correlation, between alternatives typically precludes indexability. This is so even if each subset of dependent alternatives evolves independently of all other subsets, and even if one knows how to optimally choose among the dependent alternatives in each subset in isolation. We provide an example illustrating these difficulties in the online Supplement.

[23]That is, the search technology is deteriorating if, regardless of the outcome of past searches, for any k and any upper set $A \subset \mathbb{R} \times \mathbb{N}^{|\Xi|}$ (that is, any set $A \subset \mathbb{R} \times \mathbb{N}^{|\Xi|}$ such that for each $a_1, a_2 \in \mathbb{R} \times \mathbb{N}^{|\Xi|}$ with $a_2 \geq a_1$, $a_2 \in A$ if $a_1 \in A$), one has that $\Pr((-c_{k+1}, E_{k+1}) \in A) \leq \Pr((-c_k, E_k) \in A)$. This definition is quite strong. In more specific environments, such as those in Sections 4 and 5 where there is an order on the set of categories Ξ, weaker definitions are consistent with the results below.

The next result uses the recursive characterization in Theorem 1 to identify various properties of the dynamics of exploration and expansion of the CS, which are useful in applied work.

Proposition 1 (Properties of the optimal exploration vs CS expansion dynamics).

1. ***Invariance of expansion to CS composition***: *At any period, the decision to expand the CS is invariant to the composition of the CS, conditional on the value $\mathcal{I}^*(\mathcal{S}^P)$ of the alternative with the highest index, and the state ω^S of the search technology.*

2. ***Independence of Irrelevant Alternatives***: *At any period t, for any pair of alternatives $i, j \in C_t$ with $i \neq j$, the choice between exploring alternative i or exploring alternative j is invariant to the period-t state ω^S of the search technology.*

3. ***Possible irrelevance of improvements in search technology***: *An improvement in the search technology increasing the probability of finding alternatives of positive expected value (vis-a-vis the outside option) need not affect the decision to expand the CS even at histories at which, prior to the improvement, the DM is indifferent between expanding the CS and exploring one of the alternatives already in it.*

4. ***Stationary value function***: *If the search technology is stationary, for any two states $\mathcal{S}$, $\mathcal{S}'$ at which the DM expands the CS, $\mathcal{V}^*(\mathcal{S}) = \mathcal{V}^*(\mathcal{S}')$.*

5. ***Stationary replacement***: *If the search technology is stationary or improving and search is carried out at period t, without loss of optimality, the DM never comes back to any alternative in the CS at period t.*

6. ***Single search ahead***: *If the search technology is stationary or deteriorating, at any history, the decision to expand the CS is the same as in a fictitious environment in which the DM expects she will have only one further opportunity to search.*

7. ***Pricing formula***: *Consider two states $\mathcal{S}_0 = (\mathcal{S}^P, \omega^S)$ and $\hat{\mathcal{S}}_0 = (\mathcal{S}^P, \hat{\omega}^S)$ that differ only in terms of the state of the search technology. The DM's willingness-to-pay to change the state of the search technology from ω^S to $\hat{\omega}^S$ is equal to*

$$\mathcal{P}^*(\mathcal{S}^P, \omega^S, \hat{\omega}^S) = \int_0^\infty \left(\mathbb{E}\left[\delta^{\kappa(v)} | \mathcal{S}^P, \hat{\omega}^S\right] - \mathbb{E}\left[\delta^{\kappa(v)} | \mathcal{S}^P, \omega^S\right]\right) dv.$$

Part 1 of the proposition is an implication of Theorem 1. The result is not trivial, however, because the *opportunity cost* of expanding the CS (i.e., the value of continuing with the current CS) may well depend on the entire composition of the CS, beyond the information contained in $\mathcal{I}^*(\mathcal{S}^P)$ and ω^S.

Part 2 also follows from Theorem 1. Starting with each period t, the *relative* amount of time the DM spends on each pair of alternatives in the period-t CS is invariant to what the DM expects to find by expanding the CS. This is true despite the fact that further expansions of the CS may bring alternatives that are more similar to one alternative than the other.

Part 3 follows from the fact that improvements in the search technology need not imply an increase in the index of search. This is because, as shown in part (ii) of Theorem 1, the optimal stopping time in the index of search is the first time at which the index of search and the indexes of all alternatives brought to the CS by the current and future searches fall weakly below the value of the search index at the time the current search is launched. As a result, any improvement in the search technology affecting only those alternatives whose index at the time of arrival is below the value of the search index at the time search is launched does not affect the value of the search index, and hence the decision to expand the CS.

Part 4 of the proposition says that the continuation value when search is launched is invariant to the state of the CS. This follows from the fact that, without loss of optimality, the DM never comes back to any alternative in the CS after search is launched. The same property holds in case of improving search technologies. For Part 5, note that since the state of an alternative changes only when the DM selects it, if, in period t, $\mathcal{I}^S(\omega^S) \geq \mathcal{I}^*(\mathcal{S}^P)$, under a stationary or improving search technology, the same inequality remains true in all subsequent periods. Hence, in this case, search corresponds to disposal of all alternatives in the current CS. Each time the DM searches, she starts fresh.

Part 6 follows from the recursive characterization of the stopping time in the index of search, as per part (ii) of Theorem 1. Recall that this time coincides with the first time at which the index of any physical alternative brought to the CS by the current or future searches, and the index of search itself, drop below the value of the search index at the time the current search is launched. If the search technology is stationary, or deteriorating, the index of search falls (weakly) below its current value immediately after search is launched. Hence, $\mathcal{I}^S(\omega^S)$ is invariant to the outcome of any search following the current one, conditional on ω^S.

The final part of the proposition follows from part (iii) in Theorem 1, and can be used to price changes in the search technology, with limited knowledge about the details of the environment. To see this, suppose that the econometrician, the analyst, or a search engine, have enough data about the average time it takes for an agent with an exogenous outside option equal to $v \in \mathbb{R}_+$ to exit and take the outside option, under different search technologies. Then by integrating over the relevant values of the outside option one can compute the maximal price $\mathcal{P}^*(\mathcal{S}^P, \omega^S, \hat{\omega}^S)$ that the DM is willing to pay to change the search technology from ω^S to $\hat{\omega}^S$.

4 Pandora's Problem with Endogenous Set of Boxes

Consider the following variant of Weitzman's (1979) "Pandora's boxes problem," in which the set of boxes is endogenously expanded over time. Each alternative is a "box" and belongs to a category $\xi \in \Xi$. To each category corresponds a pair (F^ξ, λ^ξ), where F^ξ is the distribution from which the box's prize v is drawn and λ^ξ is the cost of inspecting (i.e., of opening) the box. As in Weitzman's (1979) original setting, each box's prize v is drawn independently (conditional on the boxes' categories) and revealed upon the first inspection.

At each period, the DM can either (a) search for additional boxes to add to the CS, (b) open one of the boxes in the CS to learn its prize, or (c) stop and either recall the prize of one of the previously opened boxes, or take the outside option (with a value normalized to 0), with either one of the last two choices ending the decision problem. For simplicity (but also motivated by the application to online consumer search in the next section), assume that each search $m \in \mathbb{N}$ brings exactly one box, whose category ξ is drawn from Ξ according to a distribution $\rho(m) \in \Delta(\Xi)$, which may depend on the number of past searches $m-1$ but is invariant to the realizations of such past searches. The draw from each $\rho(m)$ is independent of the draw from each $\rho(l)$, $l \neq m$.

We assume that $\Xi \subset \mathbb{N}$, with higher ξ denoting superior boxes, in the sense that, for any $\xi', \xi'' \in \Xi$ with $\xi'' > \xi'$, $F^{\xi''} \succeq_{FOSD} F^{\xi'}$ and $\lambda^{\xi''} \leq \lambda^{\xi'}$ (with one of the two relationships strict). Let $\underline{\xi} \equiv \inf \Xi$ and $\overline{\xi} \equiv \sup \Xi$. The cost of the m-th expansion of the CS is $c(m)$, where $c(\cdot)$ is a positive and increasing function. In addition, we assume that, for all m, $\rho(m) \succeq_{FOSD} \rho(m+1)$; that is, the distribution $\rho(m) \in \Delta(\Xi)$ from which the category of the m-th box is drawn first-order-stochastically dominates, weakly, the distribution $\rho(m+1) \in \Delta(\Xi)$ from which the category of the $(m+1)$-th box is drawn. The combination of the assumption that $c(m)$ is weakly increasing in m and the distribution $\rho(m) \in \Delta(\Xi)$ from which the boxes are drawn "decreases" with m in a FOSD sense implies that the index of search $\mathcal{I}^S(m)$ defined below is decreasing in m and can be characterized using the same properties as when the search technology deteriorates in the sense of Definition 2 (as per part 6 of Proposition 1).

We denote by $\rho^\xi(m)$ the probability that the m-th search brings a ξ-box, with $\sum_{\xi \in \Xi} \rho^\xi(m) = 1$ for all m.[24] As in the baseline model, the DM discounts the future according to δ.

The setting described above is one in which the decision to walk away with the prize of an opened box, or the outside option, brings an end to the DM's problem. The framework described in Section 2, instead, has an infinite horizon, and the DM chooses indefinitely among the alternatives. Despite this difference, we show that the solution to this problem takes the form of an index policy akin to the one in Definition 1. Proposition 2 below characterizes the optimal policy, prescribing when to search for an additional box, the order in which existing boxes should be opened, and when to stop and either recall an opened box or the outside option.

[24] All the results extend to the case where Ξ is infinite.

The proof (in the Appendix) maps the Pandora's boxes problem with an endogenous set of boxes into an auxiliary problem that fits into the setting of Section 2, and then uses Theorem 1 and Proposition 1 to identify the properties of the optimal policy in Proposition 2.

Proposition 2 (**Pandora's-boxes with an endogenous set of boxes**). *For any ξ-box that has not been opened yet (i.e., for which $\omega^P = (\xi, \emptyset)$ for some $\xi \in \Xi$) the reservation price $\mathcal{I}^P(\xi, \emptyset)$ is given by the solution to:*

$$\mathcal{I}^P(\xi,\emptyset) = \frac{-\lambda^\xi + \delta \int_{\frac{\mathcal{I}^P(\xi,\emptyset)}{1-\delta}}^{\infty} v dF^\xi(v)}{1 + \frac{\delta}{1-\delta}\left(1 - F^\xi\left(\frac{\mathcal{I}^P(\xi,\emptyset)}{1-\delta}\right)\right)}. \tag{5}$$

For any $l \in \mathbb{R}$, let $\Xi(l) \equiv \left\{\xi \in \Xi : \ \mathcal{I}^P(\xi,\emptyset) > l\right\}$ denote the set of boxes whose reservation price exceeds l. For any m, the reservation price of search $\mathcal{I}^S(m)$ is given by the solution to:[25]

$$\mathcal{I}^S(m) = \frac{-c(m) + \delta \sum_{\xi\in\Xi(\mathcal{I}^S(m))} \rho^\xi(m)\left(-\lambda^\xi + \delta \int_{\frac{\mathcal{I}^S(m)}{1-\delta}}^{\infty} v \mathrm{d}F^\xi(v)\right)}{1 + \sum_{\xi\in\Xi(\mathcal{I}^S(m))} \rho^\xi(m)\left[\delta + \frac{\delta^2}{1-\delta}\left(1 - F^\xi\left(\frac{\mathcal{I}^S(m)}{1-\delta}\right)\right)\right]}. \tag{6}$$

The solution to Pandora's-boxes problem with an endogenous CS takes the following form:

1. *If the highest reservation price among all unopened boxes in the CS is greater than the reservation price $\mathcal{I}^S(m)$ of search, and is greater than the flow value $(1-\delta)v$ of each opened box and the outside-option, the DM opens one of the boxes with the highest reservation price.*

2. *If the reservation price of search $\mathcal{I}^S(m)$ is higher than the reservation price $\mathcal{I}^P(\xi,\emptyset)$ of any unopened box and of the flow value $(1-\delta)v$ of each opened box and the outside-option, the DM searches.*

3. *If neither of the above two situations applies, the DM stops. He then takes the prize of one of the opened boxes whose flow value $(1-\delta)v$ is the highest among the opened boxes if the latter value exceeds the outside-option, and takes the outside-option otherwise.*

As in Weitzman's problem, the reservation prices $\mathcal{I}^P(\xi,\emptyset)$ of the boxes that have not been opened yet have the following interpretation.[26] Suppose there are only two alternatives. One is an unopened ξ-box and the other is a hypothetical box, whose prize is an annuity yielding K in each period, where K is known. The reservation price is the value of K for which the DM

[25] Because all the relevant information about the state of the search technology is summarized in the number of past searches, we abuse notation and let $\mathcal{I}^S(m)$ denote the index for the m-th search.

[26] Weitzman defines the reservation price $\hat{\mathcal{I}}^P(\omega^P)$ for $\omega^P = (\xi,\emptyset)$ as the solution to $\lambda^\xi = \delta\int_{\hat{\mathcal{I}}^P(\omega^P)}^{\infty}(v - \hat{\mathcal{I}}^P(\omega^P))\mathrm{d}F^\xi(v) - (1-\delta)\hat{\mathcal{I}}^P(\omega^P)$, which yields
$\hat{\mathcal{I}}^P(\omega^P) = [-\lambda^\xi + \delta\int_{\hat{\mathcal{I}}^P(\omega^P)}^{\infty} v\mathrm{d}F^\xi(v)]/[1 - F^\xi(\hat{\mathcal{I}}^P(\omega^P))]$. The reservation prices in (5) are thus equal to those in Weitzman (1979) multiplied by $(1-\delta)$, that is, $\mathcal{I}^P(\omega^P) = (1-\delta)\hat{\mathcal{I}}^P(\omega^P)$.

is indifferent between taking the hypothetical box (yielding a continuation payoff of $K/(1-\delta)$) and inspecting the ξ-box while maintaining the option to recall the hypothetical box once the prize v of the ξ-box is discovered.

The reservation price $\mathcal{I}^S(m)$ of search extends this interpretation as follows. Suppose there are two options: the hypothetical box with known value K described above, and the option of expanding the CS. The reservation price of search is the value K for which the DM is indifferent between taking the hypothetical box right away, and expanding the CS, maintaining the option to take the hypothetical box either (a) once the category ξ of the newly discovered box is discovered and $\mathcal{I}^P(\xi, \emptyset) \leq K$, or (b), in case $\mathcal{I}^P(\xi, \emptyset) > K$, after the prize v of the newly discovered ξ-box is learned and $v \leq K/(1-\delta)$.

5 Application: Online Consumer Search

In the past two decades, internet advertising – and in particular, search advertising – has become one of the most prominent channels through which consumers learn about, and purchase, goods and services. Understanding the behavior of consumers in these markets is important in order to better understand online advertising, its structure, and design.

Search advertising, which accounts for a large fraction of Internet advertising revenues due to its effectiveness, describes a scenario where sponsored links are displayed alongside the results of consumers' search queries online.[27] Search advertising has received significant attention, with a particular focus on the behavior of firms in the auctions used in these markets, such as the Generalized Second-Price (GSP) auction, and others. However, the behavior of consumers in these markets – how they read and click ads, and which products they purchase – remains largely under-explored. In particular, existing models of search advertising have often made restrictive assumptions about users' behavior, which do not appear to square well with empirical studies.

In their analysis of consumer demand for search advertising, Jeziorski and Segal (2015) document three properties of users' behavior that, while ubiquitous, are inconsistent with existing models of search advertising:

1. *Non-sequential clicking*: Nearly half of users do not click on ads sequentially in the order of the positions in which ads are presented.

2. *Non-cascading clicking*: Over half of users who click more than once click on a higher position after having clicked on a lower position.

3. *Externalities from lower positions*: The rate at which an ad at a given position is clicked depends on which ads are displayed *below* it.

[27]The effectiveness of search advertising has been attributed to the fact that consumers' search inputs are informative about the products they are interested in, which allows targeting through relevant ads.

The Pandora's boxes problem with an endogenous set of boxes of the previous section can shed light on consumer behavior in these markets, and on the role that ad positions play given such a behavior. In particular, it provides a theoretical explanation for the patterns documented in Jeziorski and Segal (2015).

5.1 Consumer's online search and eventual purchases

Consider the problem of a consumer searching online for a product to purchase. In this environment, the consumer brings a product to her CS by reading the product's ad. Because the consumer does not know which products are advertised at the various positions before reading the corresponding ads, assume the consumer reads the ads in the order they are displayed by the platform. After reading a new ad, the consumer brings the corresponding product to her CS. At that moment the consumer decides whether to read the next ad or click on one of the products whose ad the consumer has read already. After clicking on a product's link, the consumer is directed to the vendor's website where she learns her value for the vendor's product (net of the product's price). The consumer then decides whether or not to finalize the purchase. The purchase of a product brings to an end the consumer's search process. Note that, while the consumer naturally reads the ads in the order in which they are displayed, she clicks on the links of the products whose ads have been read in the order of her choice.

Hence, in this problem, reading the next ad displayed by the platform corresponds to expanding the set of boxes in the version of Pandora's problem with an endogenous set of boxes of the previous version. Clicking on a product's link corresponds to opening a box, and purchasing a product from a visited vendor corresponds to selecting an opened box.

Consistently with the analysis in the previous section, suppose that each category $\xi \in \Xi$ corresponds to a different ad's type (equivalently, a different type of box), with each type indexing a different (absolutely continuous) distribution F^{ξ} from which the consumer's value v for the corresponding ad's product is drawn, and a different inspection cost λ^{ξ} to learn such a value.[28]

Each position $m \in \mathbb{N}$ is occupied by the ad of one and only one firm, with the same firm possibly advertising at multiple positions. Reading the m-th ad reveals to the consumer the ad's type of the firm advertising on the m-th position. We denote the ad's type of the firm occupying the m-th position by $\xi(m) \in \Xi$. The consumer believes that $\xi(m)$ is drawn from Ξ according to a distribution $\rho(m) \in \Delta(\Xi)$ that may depend on m but is invariant in the ads' types of those firms occupying the upstream positions $l < m$. For example, the consumer may expect lower positions to be occupied, on average, by lower-quality ads (that is, by products that are more costly to learn about, i.e., higher λ^{ξ}, and deliver, on average, lower values, i.e., "smaller" F^{ξ}, in the sense of FOSD), but does not change her beliefs based on the ads' types $\xi(l)$, $l < m$, encountered at

[28] The assumption that each F^{ξ} is absolutely continuous is made in order to avoid the need to keep track of possible indifferences in the consumer's optimal behavior which affect the formulas but not the qualitative results.

upstream positions.

Let $c(m)$ denote the cost of reading the m-th ad. We then have that the index for the decision to read the m-th ad is equal to $\mathcal{I}^S(m)$, with $\mathcal{I}^S(m)$ as in (6), whereas the index for the decision to click on the m-th ad, after discovering the ad's type $\xi(m)$, is equal to $\mathcal{I}_m \equiv \mathcal{I}^P(\xi(m), \emptyset)$, with $\mathcal{I}^P(\xi(m), \emptyset)$ as in (5).

One can use the model to endogenize the probability with which the consumer reads the ads, clicks on them, and finalizes her purchases. Furthermore, one can derive a structural relationship between the various positions and their click-through-rates (CTRs), accounting for the uncertainty that the consumer faces about the ads displayed at the various positions – a feature that the model with exogenous CSs does not capture.

In a similar setting, but with an exogenous CS, Choi, Dai and Kim (2018) – and, independently, Armstrong (2017) – derive a static condition characterizing eventual purchasing decisions based on a comparison of "effective values." Proposition 3 below extends their characterization to search problems with an endogenous CS. Let v_m denote the value to the consumer for the product sold by the firm advertising at the m-th position. For all $m \geq 1$, let $w_m \equiv \min\{\mathcal{I}_m, v_m(1-\delta)\}$ be the "effective value" of the product advertised at the m-th position (for brevity, product m) when the product is already in the consumer's CS, as in Choi, Dai and Kim (2018), and $d_m \equiv \min\{w_m, \mathcal{I}^S(m)\}$ the product's "discovery value," when the product must be brought to the CS before it can be explored (that is, before the consumer learns the product's ad type $\xi(m)$). Let product $m = 0$ correspond to the consumer's outside option, with $w_0 = d_0 = 0$. Note that w_m and d_m are learned by the consumer only after reading the m-th ad (which reveals its type $\xi(m)$) and clicking on it which reveal its value v_m.

Proposition 3 (**Eventual purchases**). *The consumer purchases product m if, for all $l \in \mathbb{N} \cup \{0\}$, $l \neq m$, $d_l < d_m$ (and only if $d_l \leq d_m$, for all $l \neq m$).*

As in Choi, Dai and Kim (2018), purchasing decisions are determined by a static comparison of the products' values, as in canonical discrete-choice models. Contrary to Choi, Dai and Kim (2018), however, such values account for the uncertainty the consumer faces over the products occupying the various positions (equivalently, over each product's ad type $\xi(m)$ prior to reading the ad displayed in the m-th position). Allowing for such an uncertainty is important. When all products are already in the consumer's CS, positions do not play any specific role and there is no reason why downstream positions should be expected to receive fewer clicks than upstream ones.

In contrast, when the consumer faces uncertainty about the ads occupying the various positions and chooses how to alternate between reading new ads and clicking on the links of those ads she read already, the model delivers useful structural relationships linking the positions' CTRs to the primitives of the problem. In particular, Proposition 3 implies that, when the reading cost $c(\cdot)$ is non-decreasing and the probability of finding "attractive" ads declines (weakly) with the positions, all other things equal, the further down a product is on the list, the lower the ex-ante probability the product is purchased (and hence its ex-ante demand), a property often assumed, but not micro-founded, in the models considered in the pertinent literature.

The result in Proposition 3 follows from the fact that the optimal policy is an index rule, along with the fact that the search index $\mathcal{I}^S(m)$ declines with m. Heuristically, if a consumer reads the m-th ad, it must be that the reservation prices $\mathcal{I}_l$ of all products $l < m$ already in her CS, as well as the discovered values $v_l(1-\delta)$ of those products $l < m$ that have been inspected already, are no greater than $\mathcal{I}^S(m)$. When the search index $\mathcal{I}^S(\cdot)$ is non-increasing, $\mathcal{I}^S(m+1) \leq \mathcal{I}^S(m)$. Hence, if after reading the m-th ad, $\mathcal{I}_m \geq \mathcal{I}^S(m)$, the consumer necessarily clicks on the m-th ad, thus learning product m's value v_m. Once v_m is learned, if $(1-\delta)v_m \geq \mathcal{I}_m$, the consumer stops the search and purchases product m. The above properties imply the result in the proposition. The formal proof is in the Appendix.

5.2 Non-sequential, non-cascading clicking

The question of interest is then what does it take for the index of search to decline with the position. One can use the recursive characterization of the index of search of Theorem 1 to answer the question. For simplicity, let $\Xi = \mathbb{N}$ and label the ads' types according to their attractiveness, with higher ξ denoting "higher" distributions F^{ξ} and lower inspection costs λ^{ξ}, that is, for any $\xi', \xi'' \in \Xi$ with $\xi'' > \xi'$, $F^{\xi''} \succeq_{FOSD} F^{\xi'}$ and $\lambda^{\xi''} \leq \lambda^{\xi'}$ (with one of the two relationships strict). Let $\underline{\xi} \equiv \inf \Xi$ and $\bar{\xi} \equiv \sup \Xi$. Suppose the cost of reading $c(m)$ is non decreasing in m and the consumer expects lower positions to be occupied by less attractive ads, in the sense that, for all m, the distribution $\rho(m) \in \Delta(\Xi)$ over Ξ first-order-stochastically dominates, weakly, the corresponding distribution $\rho(m+1) \in \Delta(\Xi)$. Then $\mathcal{I}^S(m+1) \leq \mathcal{I}^S(m)$ for all m. Furthermore, for all m, $\mathcal{I}^S(m) \leq \mathcal{I}^P(\bar{\xi}, \emptyset)$, that is, the index of search is smaller than the index of any ad whose type is the most attractive one. We then have the following result:

Proposition 4 (Clicking behavior). *Suppose that cost of reading $c(m)$ is non decreasing in m and that the consumer expects lower positions to be occupied by less attractive ads, in the sense that, for all m, $\rho(m) \succeq_{FOSD} \rho(m+1)$. The consumer's search for the optimal product is consistent with non-cascading and non-sequential clicking, and generates externalities from lower positions.*

The clicking dynamics under the index policy of Proposition 2 are thus consistent with the consumer behavior documented empirically in Jeziorski and Segal (2015). Importantly, these dynamics do not obtain under the search model of Athey and Ellison (2011). They can emerge in models of consumer search with an exogenous CS à la Weitzman's (1979) if one assumes a specific relationship between the positions and the ads' types. In such a model, the consumer's beliefs over the ads displayed at the various positions are degenerate, as the consumer knows (exogenously) which ad occupies each position. Hence, there is no reason why firms would want to bid more for higher positions, in contrast with what is documented in the literature on bidding for sponsored search.

Our model of consumer search with an endogenous CS can also generate dynamics under which the probability of non-cascading and non-sequential clicking is non-monotone in the positions. To see this, suppose that $\Xi = \{\underline{\xi}, \overline{\xi}\}$ and that the conditions in Proposition 4 hold. Because $\mathcal{I}^S(m)$ is decreasing in m and $\mathcal{I}^S(m) \leq \mathcal{I}^P(\overline{\xi}, \emptyset)$ for all m, when $\mathcal{I}^S(0) > \mathcal{I}^P(\underline{\xi}, \emptyset)$, there exists a position m^* such that: (a) for any $m < m^*$, the consumer clicks on the m-th ad immediately after reading it if and only if it is of type $\overline{\xi}$, whereas (b) for any $m > m^*$, the consumer clicks on the m-th ad immediately after reading it, regardless of the ad's type. When, for any $m < m^*$, the probability that the m-th ad is of type $\overline{\xi}$ is strictly decreasing in m, we then have that the probability of non-sequential and non-cascading clicking is single-peaked (increasing in m for $m < m^*$, and equal to zero for $m > m^*$). In turn, it can be shown that the probability that the m-th ad is occupied by a firm of type $\overline{\xi}$ is indeed decreasing in m when firms' profits for selling to the consumer are drawn from a distribution that is related to the ads' types by MLRP and the assignment of the ads is governed by an auction that induces monotone bidding.

5.3 Click-through-rates

The results in the previous subsection can also be used to characterize the positions' click-through rates (hereafter, CTRs), i.e., the fraction of ads at each position that, once read, are clicked upon. Formally, for each position m, the corresponding CTR is equal to

$$CTR(m) \equiv \Pr\left(m\text{'s ad is clicked}\right).$$

Depending on the problem of interest, the information used to compute the above probability may contain the type of the firms advertising at the different positions (as when firms know the attractiveness of each others' ads at the bidding stage and the above probability is computed by the firms given the induced allocation) or only the knowledge of the rules used by the search engine to assign the ads to the various positions (as when the probability is computed by a platform that does not know the attractiveness of the firms' ads, or by a firm that also lacks such information).

The next proposition relates the CTRs to the effective and discovery values introduced above.

Proposition 5 (Click-through rates). *For each position $m \geq 1$, the click-through-rate is given* by[29]

$$CTR(m) = Pr\left(\mathcal{I}^S(m) \geq max_{l<m}\{w_l\} \cap \mathcal{I}_m \geq \{max_{l<m}\{w_l\}, max_{l>m}\{d_l\}\}\right).$$

In order for the ad in position m to be read, it must be that $\mathcal{I}^S(m) \geq \max_{l<m}\{w_l\}$, for otherwise the consumer selects one of the products advertised in one of the preceding positions before reading the ad displayed in the m-th position. Once product m is read, in order for it to be clicked upon, it must be that its index $\mathcal{I}_m$ exceeds the effective value of each product brought to the consumer's CS prior to m, but also the discovery value of all products advertised further down the list, for otherwise the consumer selects one of the other products before clicking on m. Note that the property that $\mathcal{I}^S(l)$ is weakly decreasing in l is important here. It implies that, if for some position $l > m$, $d_l > \mathcal{I}_m$, then for all $j = m+1, ..., l$, $\mathcal{I}^S(j) > \mathcal{I}_m$, meaning that the consumer will necessarily read the ad of any product displayed between position m and position l before clicking on m. If for any of such product the discovery value exceeds $\mathcal{I}_m$, the consumer purchases one of these products before clicking on m, and hence never clicks on the m-th product.

Take the perspective of an observer (e.g., a platform, a firm, or a savvy consumer) knowing the rules used to assign the ads to the positions but not the attractiveness of the ads. While the probability each ad is read is decreasing in m, $\Pr(\mathcal{I}_m \geq \max_{l>m}\{d_l\})$ need not be decreasing in m. Hence, from the observer's perspective, CTRs need not be monotone in positions, consistently with what has been noticed in the empirical literature.

5.4 Implications for equilibrium bidding in sponsored-search auctions

We now show how the results above can be put to work to identify important properties of bidding in sponsored-search auctions. For simplicity, suppose there are only two firms, with each firm advertising a single product (the results below extend to auctions with more than two positions and more than two bidders). The two firms compete for placing their ads on a platform offering two different positions. The platform uses the ascending-clock version of the GSP auction of Edelman et al. (2007) to allocate the two positions. The firm dropping out first is allocated the second position and pays nothing, whereas the other firm is allocated the first position and pays the price at which the other firm drops out, per click. As in the analysis above, each firm's product can be of multiple types, with each type $\xi \in \Xi$ parametrizing the attractiveness of the firm's product/ad (formally, the distribution F^ξ from which the consumer's value is drawn) and the consumer's exploration cost λ^ξ. Let z denote the profit each firm derives from selling its product and assume that each z is drawn from $[\underline{z}, \bar{z}]$ according to a distribution F_z, with the

[29]For simplicity, the formula in the proposition assumes that, in case of indifference, the consumer favors position m (both when it comes to reading and clicking it). This is what justifies the weak inequalities in the formula. The proof in the Appendix discusses how alternative ways of breaking the indifferences must be accounted for if one were to compute bounds for such probabilities.

draws independent across firms. Assume that the payoff the consumer expects from learning her value for each firm's product (net of the exploration cost λ) exceeds her outside option, and that this is true for all possible types ξ.

The search model introduced above delivers a structural characterization of the CTRs, the purchasing probabilities, and the firms' values per click (VPC), for each possible profile of firms' types (ξ_1, ξ_2) and each possible assignment of the positions to the firms. It also delivers a characterization of the same variables from the perspective of an observer who does not know the firms' types. Formally, and consistently with the notation above, an assignment is a vector $(\xi(1), \xi(2))$ where the first entry denotes the type of firm occupying the top position, and the second entry the type of firm occupying the second position. For each position $m = 1, 2$ and each assignment $(\xi(1), \xi(2))$, let $P(m; \xi(1), \xi(2))$ denote the probability that the consumer purchases the product advertised in position m under the assignment $(\xi(1), \xi(2))$. Clearly, the consumer does not know the assignment $(\xi(1), \xi(2))$ at the beginning of the search process and learns it by reading the various ads.

Let $CTR(m; \xi(1), \xi(2))$ denote the probability that the consumer clicks on the ad displayed in the m-th position under the assignment $(\xi(1), \xi(2))$. Finally, for any position m, assignment $(\xi(1), \xi(2))$, and unit profit z, let

$$VPC(m; \xi(1), \xi(2), z) \equiv z \frac{P(m; \xi(1), \xi(2))}{CTR(m; \xi(1), \xi(2))} \tag{7}$$

denote the *value-per-click* (VPC) that a firm with unit profit z assigns to occupying the m-th position under the assignment $(\xi(1), \xi(2))$. Note that, contrary to Edelman et al. (2007) and Athey and Ellison (2011), the values per click here are not only heterogeneous across firms but also position-specific, reflecting the property that the probability the consumer finalizes a trade after clicking on a firm's ad depends on the position at which the ad is displayed, a property also documented by the empirical literature. In the online supplement, we provide a parametric example where all the above variables can be computed in closed form.

Suppose that the two firms commonly know the attractiveness of their ads/products, possibly as a result of past experiences with consumers who searched similar products on the same platform. We then have the following result:

Proposition 6 (Firms' bidding behavior). *Consider the sponsored-search model described above and assume that firms do not follow weakly dominated strategies. For any profile of ad types (ξ_1, ξ_2), there exists a threshold $b(z; \xi_1, \xi_2)$ such that each firm with unit profit z drops out at price $b(z; \xi_1, \xi_2)$ irrespective of whether its ad is more or less attractive than the rival's.*

To gather some intuition, consider a state (ξ_1, ξ_2) in which firm 1's ad is more attractive than firm 2's (in the sense that $\lambda^{\xi_1} < \lambda^{\xi_2}$, meaning that the cost to the consumer to learn her value for firm 1's product is lower than for firm 2's product, and $F^{\xi_1} \succ_{FOSD} F^{\xi_2}$, meaning that product 1 delivers, on average, more utility than product 2).

Consider the interesting case in which the consumer finds it optimal to click on the ad encountered at the first position before reading the ad in the second position, irrespective of whether the ad in the first position is firm 1's or firm 2's. That is, the index $\mathcal{I}_1 \equiv \mathcal{I}^P(\xi(1), \emptyset)$ for clicking on the ad displayed in the first position is higher than the index $\mathcal{I}^S(2)$ for reading the ad in the second position, irrespective of the type $\xi(1)$ of ad encountered in the top position. When this property does not hold, firms are indifferent between advertising in the first and second position given that the consumer always clicks first the ad of the most attractive firm irrespectively of the position at which the ad is displayed. Firm 1 then finds it optimal to drop out at a price b implicitly defined by

$$[VPC(1; (\xi_1, \xi_2, z) - b]\, CTR(1; \xi_1, \xi_2) = VPC(2; \xi_2, \xi_1, z) CTR(2; \xi_2, \xi_1). \tag{8}$$

The above indifference condition reflects the fact that both the CTRs and the VPCs are not only position-specific but also ad-specific. Equivalently, using the relation between VPCs, CTRs and selling probabilities P in Condition (7) above, we have that the price at which firm 1 drops out is equal to

$$z[P(1; \xi_1, \xi_2) - P(2; \xi_2, \xi_1)].$$

Because the first position is clicked with probability one, the firm optimally drops out when the price reaches a value equal to the extra profit the firm expects from placing its ad on the first position instead of the second one. Note that the term in square brackets is the difference between the probability the firm assigns to selling its product when listed in the top position (that is, under the assignment (ξ_1, ξ_2)) and when listed in the second position (that is, under the assignment (ξ_2, ξ_1)).

Likewise, firm 2, given its markup z, drops out when the price reaches the value b implicitly defined by

$$[VPC(1; \xi_2, \xi_1, z) - b]\, CTR(1; \xi_2, \xi_1) = VPC(2; \xi_1, \xi_2, z) \cdot CTR(2; \xi_1, \xi_2). \tag{9}$$

Note that Condition (9) differs from Condition (8) because the two firms expect different CTRs and have different VPCs for the two positions. Equivalently, using again the relationship in (7), we have that the price at which firm 2 drops out is equal to

$$z[P(1; \xi_2, \xi_1) - P(2; \xi_1, \xi_2)].$$

Clearly, the probability $P(1; \xi_2, \xi_1)$ that firm 2 assigns to selling its good when advertising in the top position is different from the probability $P(1; \xi_1, \xi_2)$ that firm 1 assigns to selling its product when occupying the same position, reflecting the difference in the distributions from which the consumer's values are drawn, the exploration costs, and the probability the consumer clicks on the second ad when encountering an ad of type ξ_1 or of type ξ_2 in the first position. However, the differential in the probability of selling when occupying the first and second position is the same for the two firms; that is,

$$P(1;\xi_2,\xi_1) - P(2;\xi_1,\xi_2) = P(1;\xi_1,\xi_2) - P(2;\xi_2,\xi_1). \tag{10}$$

As a consequence, in equilibrium, the price at which the two firms drops out is the same, despite the fact that one firm is more attractive than the other. Because, in each state, the two firms follow identical bidding strategies, a consumer who understands the rules of the auction should hold beliefs about the type of firms displaying in the second position that are invariant in the type of firm encountered in the first position, consistently with what assumed in the rest of this section.

The property in Proposition 6 has important implications for the efficiency of the equilibrium allocations under the ascending-clock implementation of the GSP auction considered in the literature. To see this, consider any welfare objective that assigns strictly positive weight to consumer surplus. We then have the following result:

Corollary 1 (**Inefficiency of equilibrium ad-allocation**). *Assume that firms do not follow weakly dominated strategies. In each state in which the attractiveness of the firms' products is not homogeneous across firms, the positions are assigned inefficiently with strictly positive probability.*

Note that efficiency requires that, in each state in which the attractiveness of the two firms' products is different across firms (that is, $\xi_1 \neq \xi_2$), whenever the difference $|z_1 - z_2|$ between the two firms' profits is small, the firm whose product is the most attractive be assigned the top position. This is because the top position is clicked more often. Hence, when the first position is occupied by the most attractive firm, the chances the consumer purchases the product she values the most (formally, for which her ex-post net value v is the highest) are higher than when the top position is occupied by the least attractive firm. Hence, no matter the weight the planner assigns to consumer surplus in the welfare objective function, as long as the latter is strictly positive, the auction allocates the positions inefficiently with positive probability. The result is a direct consequence of the fact that, in each state, the two firms follow symmetric bidding strategies, which implies that the assignment of the two positions is based entirely on the firms' unit profits z and not their attractiveness.

The result in the corollary applies also to settings with more than two firms and more than two positions. To see this, it suffices to note that the situation described above continues to represent a valid description of the problem each firm faces when there are only two firms left in the auction.[30]

The two key properties responsible for the result are (1) that firms possess some information about their attractiveness at the bidding stage, and (2) that the differential in the selling probabilities is equalized across the remaining firms, which is always the case when the remaining firms expect the buyer to purchase one of their products with certainty.

5.5 Detrimental effect of additional ad space

The model introduced above can also be used to investigate the effects of additional ad space on firms' profits. Typically, a firm receiving additional ad space expects larger profits. This, however, is not guaranteed when consumers' CS are endogenous. To see this, consider the following situation. There are three types of ads, i.e., $\xi \in \Xi = \{A, B, C\}$. The consumer's initial CS contains three products, each from a different firm and each of a different type. By searching, the consumer is presented with a fourth product whose ad's type is drawn from Ξ according to $\rho \in \Delta(\Xi)$. As above, the consumer believes that an ad of type ξ, when clicked upon, yields the consumer a net value v drawn from a distribution F^{ξ}, independently across products.[31] The cost to the consumer of learning the value of a product whose ad's type is ξ is λ^{ξ}. The consumer has unit demand and each firm makes the same profit from selling one of its products.

Suppose that each firm's ads are all of the same type and that the ads of different firms are of different type (in other words, in this example, ξ also indexes the identity of each firm). As we show in the online Supplement, an increase in the probability that the new search brings an additional product of type ξ (equivalently, an additional product from firm ξ) may reduce the index of search, inducing the consumer to visit the website of one of firm ξ's competitors before searching for the new product. When strong enough, such an effect may reduce the probability that one of firm ξ's products is eventually purchased, and hence firm ξ's profits. See the online Supplement for the details.

6 Conclusions and Extensions

We introduce a model of experimentation in which the decision maker alternates between exploring alternatives already in the consideration set and searching for new ones to explore in the future. Each search brings stochastically a new set of alternatives of different types that is added to the current consideration set. The consideration set is thus constructed gradually over time in response to the information the decision maker collects. We characterize the optimal policy and

[30] The selling probabilities $P(m; \xi(1), \xi(2))$, click-through rates $CTR(m; \xi(1), \xi(2)))$, and values-per-click $VPC(1; (\xi_1, \xi_2, z)$, for positions $m = 1, 2$ should then be interpreted as given the assignment $(\xi(m))_{m>2}$ of the lower positions observed by the two remaining firms.

[31] If the extra product the consumer is presented when searching is from firm ξ, the value the consumer derives from such a product is also drawn from F^{ξ}, independently from the value derived from the three products already in the CS.

study how the tradeoff between the exploration of existing alternatives and the expansion of the consideration set depends on the search technology. The evolution of this tradeoff is driven by a comparison of independent indexes, where the index for search is computed in recursive form, accounting for future optimal decisions.

The analysis may be of interest to dynamic problems in which the decision maker is unable to consider all feasible alternatives from the outset, either because of limited attention, or because of the sequential provision of information by interested third parties such as online platforms and search engines.

The results accommodate several extensions that may be relevant for applications.

Multiple expansion possibilities. In certain problems of interest, the decision to search also involves an intensive margin, as when the DM chooses "how much" to invest in search. As we show in the online Supplement, in general, such problems do not admit an index solution because of the correlation in the search outcomes. Instead, the analysis readily extends to an environment in which there are multiple search possibilities with independent outcomes, by allowing for multiple "search arms".

No discounting. All results above assume that $\delta < 1$. However, they extend to $\delta = 1$ (i.e., no discounting). As noted in Olszewski and Weber (2015), bandit problems in which $\delta = 1$ can be thought of as problems with non-discounted "target processes" where arms reaching a certain (target) state stop delivering payoffs. A well-known result for such problems is that the finiteness they impose allows one to take the limit as $\delta \to 1$ (e.g., Dumitriu, Tetali, and Winkler, 2003).

Irreversible Choice. In many decision problems, in addition to learning about existing options and searching for new ones, the DM can irreversibly commit to one of the alternatives, bringing to an end the exploration process. In general, such problems do not admit an index solution. In the online Supplement, we derive a sufficient condition under which the optimality of an index rule extends to such problems. We assume the DM must explore each alternative of category ξ at least $M_\xi \geq 0$ times before she can irreversibly commit to it (for example, a consumer must visit a vendor's webpage at least once to finalize a transaction with that vendor, as in the consumer search problem of Section 5). The condition guarantees that, once an alternative reaches a state in which the DM can irreversibly commit to it, its "retirement value" (that is, the value of irreversibly committing to it) either drops below the value of the outside option, or improves, weakly, with the number of future explorations. This property is related to a similar condition in Glazebrook (1979), who establishes the optimality of an index policy in a class of bandit problems with stoppable processes. Our proof, however, is different and accounts for the fact that the set of alternatives evolves endogenously over time.

Relative length of expansion. In order to allow for frictions in the search for new alternatives, we assume that, whenever the DM searches, she cannot explore any of the alternatives in the CS, with search occupying the same amount of time as the exploration of any of the alternatives in the CS. All the results extend to a setting in which both the time that each search occupies and the time that each exploration takes vary stochastically with the state.[32] Furthermore, because the time that each exploration takes can be arbitrary, by resccaaling the payoffs and adjusting the discount factor appropriately, one can make the length of time during which the exploration of the existing alternatives is paused because of search arbitrarily small. The results therefore also apply to problems in which search and learning occur "almost" in parallel.

References

Armstrong, M. (2017). "Ordered consumer search," *Journal of the European Economic Association, 15*(5), 989-1024.

Armstrong, M. and J. Vickers (2015). "Which demand systems can be generated by discrete choice?" *Journal of Economic Theory*, 158, 293-307.

Athey, S. and G. Ellison (2011). "Position auctions with consumer search." *The Quarterly Journal of Economics* 126.3, 1213-1270.

Austen-Smith, D. and C. Martinelli (2018). "Optimal exploration." Working paper.

Bardhi, A., Y. Guo, and B. Strulovici (2021). "Early-Career Discrimination: Spiraling or Self-Correcting?" mimeo Duke and Northwestern Universities.

Bergemann, Dirk and J. Välimäki (2008). "Bandit problems". In: *The New Palgrave Dictionary of Economics*. Ed. by Steven N. Durlauf and Lawrence E. Blume. Basingstoke: Palgrave Macmillan.

Caplin, A., Dean, M. and D. Martin (2011). "Search and satisficing." *American Economic Review*, 101(7), 2899-2922.

Caplin, A., Dean, M. and J. Leahy (2019). "Rational inattention, optimal consideration sets, and stochastic choice," *The Review of Economic Studies*, 86(3), 1061-1094.

Che, Y. K. and K. Mierendorff (2019). "Optimal dynamic allocation of attention." *American Economic Review*, 109(8), 2993-3029.

Choi, M., Dai, A. Y. and K. Kim (2018). "Consumer search and price competition." *Econometrica*, 86(4), 1257-1281.

Choi, M. and L. Smith (2016). "Optimal sequential search among alternatives." mimeo University of Wisconsin.

Doval, L. (2018). "Whether or not to open Pandora's box." *Journal of Economic Theory* 175, 127–158.

Dumitriu, I., Tetali, P. and P. Winkler (2003). "On playing golf with two balls," *SIAM Journal on Discrete Mathematics*, 16(4), 604-615.

[32]More generally, all of the results can be extended to a semi-Markov environment, where time is not slotted.

Edelman, B., Ostrovsky, M., and M. Schwarz (2007). "Internet advertising and the generalized second-price auction," *American Economic Review*, 97 (1), 242–259.

Edelman, B., and M. Schwarz (2010). "Optimal auction design and equilibrium selection in sponsored search auctions." *American Economic Review* 100(2), 597–602.

Eliaz, K. and R. Spiegler (2011). "Consideration sets and competitive marketing," *The Review of Economic Studies*, 78(1), 235-262.

Fershtman, D. and A. Pavan (2021). "Soft Affirmative Action and Minority Recruitment," *American Economic Review: Insights*, 3(1), 1-18.

Fudenberg, D., Strack P. and T. Strzalecki (2018). "Speed, accuracy, and the optimal timing of choices," *American Economic Review* 108 (12), 3651–3684.

Garfagnini, U. and B. Strulovici (2016). "Social experimentation with interdependent and expanding technologies," *Review of Economic Studies*, 83(4), 1579-1613.

Gittins, J. C. (1979). "Bandit processes and dynamic allocation indexes," *Journal of the Royal Statistical Society. Series B (Methodological)* 41 (2), 148–177.

Gittins, J. and D. Jones (1974). "A dynamic allocation index for the sequential design of experiments. In J. Gani (Ed.)," *Progress in Statistics*, pp. 241-266. Amsterdam, NL: North-Holland.

Glazebrook, K. D. (1979). "Stoppable families of alternative bandit processes." *Journal of Applied Probability* 16.4: 843-854.

Gomes, R. and K. Sweeney (2014). "Bayes–Nash equilibria of the generalized second-price auction," *Games and Economic Behavior* 86, 421–437.

Gossner, O., Steiner, J. and C. Stewart (2021). "Attention, please!," *Econometrica*, 89(4), 1717-1751.

Greminger, R. (2021). "Optimal search and discovery," mimeo, Tilburg University.

Hauser, J. R. and B. Wernerfelt (1990). "An evaluation cost model of consideration sets," *Journal of Consumer Research*, 16(4), 393-408.

Honka, E., Hortacsu, A. and M. Wildebeest (2019). "Empirical search and consideration sets," *Handbook of the Economics of Marketing*, Jean-Pierre Dube and Peter Rossi (ed.), Elsevier, 193-257.

Jeziorski, P. and I. Segal (2015). "What makes them click: Empirical analysis of consumer demand for search advertising," *American Economic Journal: Microeconomics*, 7(3), 24-53.

Ke, T. T., Z.-J. M. Shen and J. M. Villas-Boas (2016). "Search for information on multiple products," *Management Science* 62 (12), 3576-3603.

Ke, T. T. and J. M. Villas-Boas (2019). "Optimal learning before choice," *Journal of Economic Theory* 180, 383-437.

Keller, G. and A. Oldale (2003). "Branching bandits: a sequential search process with correlated pay-offs," *Journal of Economic Theory*, 113(2), 302-315.

Liang, A., Mu, X., and V. Syrgkanis (2022). "Dynamically aggregating diverse information," *Econometrica*, 90(1), 47-90.

Mandelbaum, A. (1986). "Discrete multi-armed bandits and multi-parameter processes," *Probability Theory and Related Fields 71*(1), 129–147.

Manzini, P. and M. Mariotti (2014). "Stochastic choice and consideration sets," *Econometrica*, 82(3), 1153-1176.

Masatlioglu, Y., Nakajima, D. and E. Y. Ozbay (2012). "Revealed attention," *American Economic Review*, 102(5), 2183-2205.

Olszewski, W. and R. Weber (2015). "A more general Pandora rule?," *Journal of Economic Theory* 160, 429–437.

Roberts, J. H. and J. M. Lattin (1991). "Development and testing of a model of consideration set composition," *Journal of Marketing Research*, 28(4), 429-440.

Carnehl, C. and J. Schneider (2023). "On risk and time pressure: When to think and when to do," *Journal of the European Economic Association,* 21(1), 1-47.

Simon, H. A. (1955). "A behavioral model of rational choice," *The Quarterly Journal of Economics*, 69(1), 99-118.

Villar, S. S., Bowden, J., and J. Wason (2015). "Multi-armed bandit models for the optimal design of clinical trials: benefits and challenges." *Statistical Science: A Review Journal of the Institute of Mathematical Statistics, 30*(2), 199.

Weber, R., (1992). "On the Gittins Index for Multiarmed Bandits," *The Annals of Applied Probability*, Vol. 2(4), 1024-1033.

Weiss, G. (1988). "Branching bandit processes," *Probability in the Engineering and Informational Sciences*, 2(3), 269-278.

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

7 Appendix

Proof of Theorem 1. Below we first establish the result in part (ii) and then use the recursive representation of the search index in (3) to show that, when the DM follows an index policy, her expected (per-period) payoff satisfies the representation in (4), thus establishing part (iii). In the online Supplement, we also show how the representation of the DM's payoff in (4), along with the recursive representation of the search index in part (ii) and an appropriate description of the state space that exploits the classification of the alternatives into categories, permits us to establish part (i), i.e., the optimality of the index policy, by means of a novel proof that shows that the DM's payoff under such a policy satisfies the Bellman equation for the dynamic program under consideration.

Part (ii). Let $\hat{\tau}$ be the optimal stopping time in the definition of $\mathcal{I}^S(\omega^S)$. Note that, at $\hat{\tau}$, the index of each alternative brought to the CS by the search under consideration (initiated in state ω^S), as well as the index of search itself, must be weakly smaller than $\mathcal{I}^S(\omega^S)$. Otherwise, by continuing to search, or by selecting one of the alternatives brought to the CS by the search under consideration for which the index is larger than $\mathcal{I}^S(\omega^S)$ and stopping optimally from that moment onward, the DM would attain an average payoff per unit of average discounted time

$$\frac{\mathbb{E}^\pi\left[\sum_{s=0}^{\tau-1}\delta^s U_s|\omega^S\right]}{\mathbb{E}^\pi\left[\sum_{s=0}^{\tau-1}\delta^s|\omega^S\right]}$$

strictly greater than $\mathcal{I}^S(\omega^S)$, contradicting the optimality of $\hat{\tau}$ in the definition of $\mathcal{I}^S(\omega^S)$.[33] This implies that $\hat{\tau}$ is weakly greater than τ^*, where the latter is the first time at which the index of search and the index of each alternative brought to the CS by the search under consideration are weakly below $\mathcal{I}^S(\omega^S)$. Moreover, since at τ^* the index of search and of each alternative brought to the CS by the search under consideration are weakly below $\mathcal{I}^S(\omega^S)$, if $\hat{\tau} > \tau^*$, the average payoff per unit of average discounted time between τ^* and $\hat{\tau}$ must be equal to $\mathcal{I}^S(\omega^S)$. Hence, under the optimal selection rule in the definition of $\mathcal{I}^S(\omega^S)$, the average payoff per unit of average discounted time from 0 to τ^* must also be equal to $\mathcal{I}^S(\omega^S)$. This implies that the optimal stopping time in the definition of $\mathcal{I}^S(\omega^S)$ can be taken to be τ^*. Because the index policy χ^* selects in each period between 0 and τ^* the alternative for which the average payoff per unit of average discounted time is the largest (including search), we have that the optimal selection rule π in the definition of $\mathcal{I}^S(\omega^S)$ must coincide with the index policy χ^*. That $\mathcal{I}^S(\omega^S)$ satisfies the representation in part (ii) then follows from the arguments above.

Part (iii). We construct the following stochastic process based on the values of the indexes, and the expansion of the CS through search, *under the index policy* χ^*. Starting with the initial state $\mathcal{S}_0 = (\mathcal{S}_0^P, \omega_0^S)$, let $v^0 \equiv \max\{\mathcal{I}^*(\mathcal{S}_0^P), \mathcal{I}^S(\omega_0^S)\}$. Let $t(v^0)$ be the first time at which, when the DM follows the policy χ^*, all indexes are *strictly* below v^0, with $t(v^0) = \infty$ if this event never occurs. Note that $t(v^0)$ differs from $\kappa(v^0)$, as $\kappa(v^0) = 0$ is the first time at which all indexes are *weakly* below v^0. Next let $v^1 \equiv \max\{\mathcal{I}^*(\mathcal{S}_{t(v^0)}^P), \mathcal{I}^S(\omega_{t(v^0)}^S)\}$ be the value of the largest index at $t(v^0)$, where $\mathcal{S}_{t(v^0)} = (\mathcal{S}_{t(v^0)}^P, \omega_{t(v^0)}^S)$ is the state of the decision problem in period $t(v^0)$. Note that, by construction, $t(v^0) = \kappa(v^1)$. Furthermore, when $t(v^0) < \infty$, if $v^0 > \mathcal{I}^S(\omega_0^S)$, then $\omega_{t(v^0)}^S = \omega_0^S$. We can proceed in this manner to obtain a strictly decreasing sequence of values $(v^i)_{i\geq 0}$, with corresponding stochastic times $(\kappa(v^i))_{i\geq 0}$. Note that the values v^i are all non-negative, as the DM's outside option is normalized to zero.

Next, for any $i = 0, 1, 2, ...$, let $\eta^i \equiv \sum_{s=\kappa(v^i)}^{\kappa(v^{i+1})-1}\delta^{s-\kappa(v^i)}U_s$ denote the discounted sum of the net payoffs between periods $\kappa(v^i)$ and $\kappa(v^{i+1}) - 1$, when the DM follows the index policy, and let $(\eta^i)_{i\geq 0}$ denote the corresponding sequence of discounted accumulated net payoffs, with $\eta^i = 0$ if

[33] Since infinity is allowed as a value of the stopping time, the supremum in the definitions of $\mathcal{I}^S$ (and $\mathcal{I}^P$) is attained, that is, an optimal stopping time exists (the arguments are similar to those in Mandelbaum, 1986, and hence omitted).

$\kappa(v^i) = \infty$.

Denote by $\mathcal{V}(\mathcal{S}_0)$ the expected (per-period) net payoff under the index policy χ^*, given the initial state of the problem $\mathcal{S}_0$. That is, $\mathcal{V}(\mathcal{S}_0) = (1-\delta)\mathbb{E}^{\chi^*}\left[\sum_{t=0}^{\infty} \delta^t U_t | \mathcal{S}_0\right]$. By definition of the processes $(\kappa(v^i))_{i\geq 0}$ and $(\eta^i)_{i\geq 0}$, $\mathcal{V}(\mathcal{S}_0) = (1-\delta)\mathbb{E}^{\chi^*}\left[\sum_{i=0}^{\infty} \delta^{\kappa(v^i)} \eta^i | \mathcal{S}_0\right]$. Next, using the definition of the indexes (1) and (2), observe that

$$v^i = \frac{(1-\delta)\mathbb{E}^{\chi^*}\left[\eta^i | \mathcal{S}_{\kappa(v^i)}\right]}{\mathbb{E}^{\chi^*}\left[1 - \delta^{\kappa(v^{i+1}) - \kappa(v^i)} | \mathcal{S}_{\kappa(v^i)}\right]}. \tag{11}$$

To see why (11) holds, recall that, at period $\kappa(v^i)$, given the state of the decision problem $\mathcal{S}_{\kappa(v^i)}$, the value of the highest index is v^i. Now suppose that the alternative corresponding to v^i is a physical alternative and that all other physical alternatives' indexes, as well as the index of search, are strictly below v^i. Recall that the optimal stopping time τ in the definition of the index of the physical alternative corresponding to v^i in (1) is the first period (strictly above $\kappa(v^i)$) at which the alternative's index falls below v^i. While it is convenient to take this fall to be weak, it is well known that one can equivalently take the fall to be strict. That is, stopping at the first period at which the index reaches a value equal to or smaller than the value at the time the index was computed is optimal, but so is stopping at the first period at which the index reaches a value strictly below the one at the time the index was computed. Now recall that $t(v^i)$ is the first time at which all indexes are strictly below v^i. Because the CS in period $\kappa(v^i)$ contains only one alternative with index equal to v^i (the physical one under consideration), $t(v^i)$ also coincides with the first period at which the index of the specific alternative under consideration drops strictly below v^i. Recall that v^{i+1} is the largest index at period $t(v^i)$ and that $t(v^i) = \kappa(v^{i+1})$. The definition of the index in (1), along with the optimality of stopping at the first time the index drops strictly below its initial value, and the definition of η^i, then imply that

$$v^i = \frac{\mathbb{E}^{\chi^*}\left[\sum_{s=\kappa(v^i)}^{\kappa(v^{i+1})-1} \delta^{s-\kappa(v^i)} U_s | \mathcal{S}_{\kappa(v^i)}\right]}{\mathbb{E}^{\chi^*}\left[\sum_{s=\kappa(v^i)}^{\kappa(v^{i+1})-1} \delta^{s-\kappa(v^i)} | \mathcal{S}_{\kappa(v^i)}\right]} = \frac{\mathbb{E}^{\chi^*}\left[\eta^i | \mathcal{S}_{\kappa(v^i)}\right]}{\mathbb{E}^{\chi^*}\left[\frac{1-\delta^{\kappa(v^{i+1})-\kappa(v^i)}}{1-\delta} | \mathcal{S}_{\kappa(v^i)}\right]}$$

which corresponds to the formula in (11).

Next, suppose that the alternative with the highest index at period $\kappa(v^i)$ is search, and that all physical alternatives in the CS in period $\kappa(v^i)$ have an index strictly smaller than v^i. As shown above, the optimal stopping time in the definition of the index of search in (2) is the first period (strictly above $\kappa(v^i)$) at which the index of search and of all the alternatives introduced through search, fall weakly below v^i. Equivalently, as discussed above, the optimal stopping time can also be taken to be the first period at which the index of search and of all the alternatives introduced through search fall strictly below v^i. Because all physical alternatives in the CS at period $\kappa(v^i)$ have an index strictly below v^i, such a period coincides with $t(v^i)$, that is, with the first period at which the index of search and of all alternatives in the CS are strictly below v^i. Using the above property of the optimal stopping time in the definition of the search index in (2), along with the fact that $t(v^i) = \kappa(v^{i+1})$ and the definition of η^i, we then have that the search index evaluated at period $\kappa(v^i)$ also satisfies the condition in (11).

Finally, suppose that, at period $\kappa(v^i)$, there are multiple options ("physical" alternatives and/or search) with index v^i. Then observe that the average sum $\mathbb{E}^{\chi^*}\left[\sum_{s=\kappa(v^i)}^{\kappa(v^{i+1})-1} \delta^{s-\kappa(v^i)} U_s | \mathcal{S}_{\kappa(v^i)}\right]$ of the discounted net payoffs from utilizing all options whose period-$\kappa(v^i)$ index is equal to v^i till the first period $t(v^i) = \kappa(v^{i+1})$ at which the indexes of all options are strictly below v^i, normalized by the average per unit discounted time $\mathbb{E}^{\chi^*}\left[\sum_{s=\kappa(v^i)}^{\kappa(v^{i+1})-1} \delta^{s-\kappa(v^i)} | \mathcal{S}_{\kappa(v^i)}\right]$ is the same as the average sum $\mathbb{E}^{\chi^*}\left[\sum_{s=\kappa(v^i)}^{T-1} \delta^{s-\kappa(v^i)} U_s | \mathcal{S}_{\kappa(v^i)}\right]$ of the discounted net payoffs from utilizing each individual option with index (at period $\kappa(v^i)$) equal to v^i till the first time T at which that option's index (and, in case the option is search, also the indexes of all alternatives brought to the CS by the search initiated at $\kappa(v^i)$) fall strictly below v^i, normalized by the average discounted time $\mathbb{E}^{\chi^*}\left[\sum_{s=\kappa(v^i)}^{T-1} \delta^{s-\kappa(v^i)} | \mathcal{S}_{\kappa(v^i)}\right]$. This follows from the independence of the processes. Hence, Condition (11) also holds when, at $\kappa(v^i)$, there are multiple options with index v^i.

Multiplying both sides of (11) by $\delta^{\kappa(v^i)}$, rearranging terms, and using the fact that $\delta^{\kappa(v^i)}$ is known at $\kappa(v^i)$, we have that

$$(1-\delta)\mathbb{E}^{\chi^*}\left[\delta^{\kappa(v^i)} \eta^i | \mathcal{S}_{\kappa(v^i)}\right] = v^i \mathbb{E}^{\chi^*}\left[\delta^{\kappa(v^i)} - \delta^{\kappa(v^{i+1})} | \mathcal{S}_{\kappa(v^i)}\right].$$

Taking expectations of both sides of the previous equality given the initial state $\mathcal{S}_0$, and using the law of iterated expectations, we have that

$$(1-\delta)\mathbb{E}^{\chi^*}\left[\delta^{\kappa(v^i)} \eta^i | \mathcal{S}_0\right] = \mathbb{E}^{\chi^*}\left[v^i\left(\delta^{\kappa(v^i)} - \delta^{\kappa(v^{i+1})}\right) | \mathcal{S}_0\right].$$

If follows that

$$\mathcal{V}(\mathcal{S}_0) = \mathbb{E}^{\chi^*}\left[\sum_{i=0}^{\infty} v^i\left(\delta^{\kappa(v^i)} - \delta^{\kappa(v^{i+1})}\right) | \mathcal{S}_0\right]. \tag{12}$$

Next, note that $\delta^{\kappa(v^i)} = 0$ whenever $\kappa(v^i) = \infty$, and that, for any $i = 0, 1, ..., \kappa(v) = \kappa(v^{i+1})$ for all $v^{i+1} < v < v^i$. It follows that (12) is equivalent to

$$\mathcal{V}(\mathcal{S}_0) = \mathbb{E}^{\chi^*}\left[\int_0^\infty v \mathrm{d}\delta^{\kappa(v)} | \mathcal{S}_0\right] = \mathbb{E}^{\chi^*}\left[\int_0^\infty \left(1 - \delta^{\kappa(v)}\right) \mathrm{d}v | \mathcal{S}_0\right] = \int_0^\infty \left(1 - \mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)} | \mathcal{S}_0\right]\right) \mathrm{d}v. \tag{13}$$

The construction of the integral function (13) is illustrated in Figure 1.

Proof of Proposition 2. Consider a relaxed problem in which the DM gets a flow payoff equal to $(1-\delta)v$ each time she selects an opened box with value v, and can revert her decision at any period. The solution to such a problem is the index policy of Theorem 1 and has the property

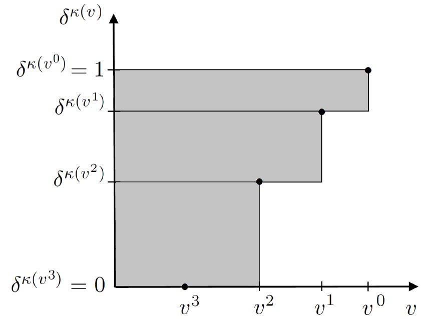

Figure 1: An illustration of the function $\delta^{\kappa(v)}$ and the region $\sum_{i=0}^{\infty} v^i \left(\delta^{\kappa(v^i)} - \delta^{\kappa(v^{i+1})}\right) = \int_0^{\infty} v \mathrm{d}\delta^{\kappa(v)}$, for a particular path with $\kappa(v^3) = \infty$.

that, once an opened box is selected, it continues to be selected in all subsequent periods. The index policy for such a problem is thus feasible (and hence optimal) also in the primitive problem.

To see that the index of a ξ-box that has not been opened yet is given by (5), note that the index of an opened box is equal to $(1-\delta)v$. Because the optimal stopping time τ^* in the definition of the index $\mathcal{I}^P(\omega^P)$ in (1) is the first time at which the value of the index drops below its value $\mathcal{I}^P(\omega^P)$ at the time the index is computed, we then have that $\tau^* = 1$ if $(1-\delta)v \leq \mathcal{I}^P(\omega^P)$ and $\tau^* = \infty$ otherwise.

Turning to the index for search, the combination of the assumption that $c(m)$ is weakly increasing in m with the assumption that the distribution $\rho(m) \in \Delta(\Xi)$ from which the boxes are drawn "decreases" with m in a FOSD sense implies that the optimal stopping-time τ^* in (2) is equal to (a) $\tau^* = \infty$ if the box identified at the m-th search has a reservation price $\mathcal{I}^P(\omega^P) > \mathcal{I}^S(m)$ and its realized flow payoff satisfies $v(1-\delta) > \mathcal{I}^S(m)$, (b) $\tau^* = 1$ if $\mathcal{I}^P(\omega^P) \leq \mathcal{I}^S(m)$, and (c) $\tau^* = 2$ if $\mathcal{I}^P(\omega^P) > \mathcal{I}^S(m)$ and $v(1-\delta) \leq \mathcal{I}^S(m)$. ■

Proof of Proposition 3. Since product 0 corresponds to the outside option, a product is always purchased. Let $l \neq m$ be such that $d_l < d_m$. We show that product l will not be purchased.

Case 1: $l > m$ (i.e., l is read after m is read). First, suppose that $d_l = \mathcal{I}^S(l)$. Because $\mathcal{I}^S(l) \leq \mathcal{I}^S(m)$ and because $\min\{\mathcal{I}_m, (1-\delta)v_m\} \geq d_m > \mathcal{I}^S(l)$, under the index policy of Theorem 1, product l is read only after product m is clicked upon. Once m is clicked, however, because $(1-\delta)v_m > \mathcal{I}^S(l)$, l is never read. Hence, l will not be purchased. Next suppose that $d_l = \mathcal{I}_l$. Then, $\min\{\mathcal{I}_m, (1-\delta)v_m\} \geq d_m > \mathcal{I}_l$. Thus, product l is clicked only after m is clicked. But again, once m is clicked, because $(1-\delta)v_m > \mathcal{I}_l$, l is never clicked, implying that l is not purchased. Finally, suppose $d_l = (1-\delta)v_l$. Then, because $\min\{\mathcal{I}_m, (1-\delta)v_m\} \geq d_m > (1-\delta)v_l$, m must be clicked before l is purchased. Because $v_m > v_l$, l is not purchased after m's value is learned.

Case 2: $l < m$ (i.e., l is read before m is read). Because

$$\mathcal{I}^S(m) \geq d_m > d_l \equiv \min\{\mathcal{I}_l, (1-\delta)v_l, \mathcal{I}^S(l)\},$$

and because $\mathcal{I}^S(m) \leq \mathcal{I}^S(l)$, it must be that $d_l = \min\{\mathcal{I}_l, (1-\delta)v_l\}$ and hence

$$\min\{\mathcal{I}_l, (1-\delta)v_l\} < d_m \leq \min\{\mathcal{I}_m, (1-\delta)v_m\}. \tag{14}$$

Furthermore, because the search technology is non-improving, $\mathcal{I}^S(l+1) \geq ... \geq \mathcal{I}^S(m-1) \geq \mathcal{I}^S(m)$. Along with the fact that $d_l = \min\{\mathcal{I}_l, (1-\delta)v_l\} < d_m \leq \mathcal{I}^S(m)$, this implies that $\min\{\mathcal{I}_l, (1-\delta)v_l\} < \mathcal{I}^S(k)$ for all $(l+1) \leq k \leq m$. This last property in turn implies that either clicking on l, or purchasing l, is dominated by reading any product k, with $(l+1) \leq k \leq m$. If m is read, then (14) implies that l will not be purchased (the arguments are similar to those for case 1). If, instead, m is not read, it must be that another product $k \neq l, m$ is purchased. In either case, product l is not purchased. ■

Proof of Proposition 5. The proof is in two steps. Step 1 shows that $\mathcal{I}^S(m) \geq \max_{l<m}\{w_l\}$ is necessary for product m to be read and that $\mathcal{I}^S(m) > \max_{l<m}\{w_l\}$ implies that product m is necessarily read. Step 2 shows that product m is clicked only if

$$\mathcal{I}^S(m) \geq \max_{l<m}\{w_l\} \ \text{ and } \ \mathcal{I}_m \geq \max\left\{\max_{l>m}\{d_l\}, \max_{l<m}\{w_l\}\right\} \tag{15}$$

and that, when both of the above inequalities are strict, product m is necessarily clicked. The result in the proposition then follows directly from the above properties.

Step 1. To see that $\mathcal{I}^S(m) \geq \max_{l<m}\{w_l\}$ is necessary for product m to be read, suppose that, for some $l < m$, $w_l > \mathcal{I}^S(m)$. That is, both the index corresponding to clicking on product l, $\mathcal{I}_l$, and the one corresponding to purchasing product l, $(1-\delta)v_l$, are strictly greater than $\mathcal{I}^S(m)$. Because product l is read before product m is read, by Theorem 1, m is never read.

Next, we show that, when $\mathcal{I}^S(m) > \max_{l<m}\{w_l\}$, product m is always read. To see this, note that since the search cost $c(\cdot)$ is increasing, $\mathcal{I}^S(1) \geq ... \geq \mathcal{I}^S(m-1) \geq \mathcal{I}^S(m)$. Therefore, $\mathcal{I}^S(m) > \max_{l<m}\{w_l\}$ implies that, for any $1 \leq l \leq m$, $\mathcal{I}^S(l) > w_{l-1} = \min\{\mathcal{I}_{l-1}, (1-\delta)v_{l-1}\}$. Hence, by Theorem 1, for any $1 \leq l \leq m$, it cannot be that product $l-1$ is purchased before product l is read. Repeatedly applying this argument for all $1 \leq l \leq m$ implies product m must be read before any product $l < m$ is purchased.

Step 2. To see that both inequalities in (15) must hold for product m to be clicked, first observe that we already established in Step 1 that the first inequality in (15) is necessary for product m to be read. Thus assume that such inequality holds. To see that the second inequality in (15) must also hold, suppose that $\mathcal{I}_m < \max\left\{\max_{l>m}\{d_l\}, \max_{l<m}\{w_l\}\right\}$. Then either there exists a product $l < m$ such that $w_l > \mathcal{I}_m$, or a product $l > m$ such that $d_l > \mathcal{I}_m$, or both. Suppose there is a product $l < m$ such that $w_l > \mathcal{I}_m$. Then product m cannot be clicked, because product l is necessarily read before m and, because both $\mathcal{I}_l$ and $(1-\delta)v_l$ are strictly greater than $\mathcal{I}_m$, product l is purchased before m is clicked. Next, suppose that there exists a product $l > m$ such that $d_l = \min\{\mathcal{I}^S(l), \mathcal{I}_l, (1-\delta)v_l\} > \mathcal{I}_m$. By the monotonicity of the search indexes, $\mathcal{I}^S(m) \geq \mathcal{I}^S(m+1) \geq ... \geq \mathcal{I}^S(l)$. That $\mathcal{I}^S(l) > \mathcal{I}_m$, then implies that $\mathcal{I}^S(k) > \mathcal{I}_m$ for any $k = m, m+1, ..., l$. In turn, this last property implies that clicking on m is dominated by reading product k, for any $k = m+1, ..., l$. If product l is read, because both $\mathcal{I}_l$ and $(1-\delta)v_l$ are strictly greater than $\mathcal{I}_m$, product m is not clicked. If, instead, product l is not read, it must be that another product $k \neq l, m$, with $k \in \{m+1, ..., l-1\}$, is purchased. In either case, product m is not clicked. Hence, the two inequalities in (15) are necessary for product m to be clicked.

Next, we show that when the two inequalities in (15) are strict, product m is necessarily clicked. We already established in Step 1 that, when the first inequality in (15) is strict, product m is read. Now suppose that the second inequality is also strict. That $\mathcal{I}_m > \max_{l<m}\{w_l\}$ implies that, for each product $l < m$, either $\mathcal{I}_l$ or $(1-\delta)v_l$ are strictly smaller than $\mathcal{I}_m$. Because product m is read, by Theorem 1, it cannot be that any product $l < m$ is purchased before product m is clicked. Similarly, that $\mathcal{I}_m > \max_{l>m}\{d_l\}$ implies that, for each $l > m$, either $\mathcal{I}^S(l)$, or $\mathcal{I}_l$, or $(1-\delta)v_l$ are strictly smaller than $\mathcal{I}_m$, which again guarantees that no product $l > m$ can be purchased before product m is clicked. Since one of the products is necessarily purchased (product 0 representing the outside option), it must be that product m is clicked. Hence, we conclude that when the two inequalities in (15) are both strict, product m is necessarily clicked. ■

Proof of Proposition 6. Consider any state of the world in which the two firms' attractiveness differs, that is, $\xi_1 \neq \xi_2$, and, without loss of generality, assume that firm 1's product is the most attractive one. As explained in the main text, when the index for clicking $\mathcal{I}^P(\xi_2, \emptyset)$ of the least attractive firm is smaller than the index for reading the second ad, each firm drops out instantaneously because it expects the consumer to click on the ad of the most attractive firm first, irrespective of the position at which such an ad is displayed. That the two firms drop out at the same price then follows from the analysis in the main text after the proposition by observing that, given any assignment $(\xi(1), \xi(2))$, $P(2; \xi(1), \xi(2)) = 1 - P(1; \xi(1), \xi(2))$, which in turn is due to the assumption that the payoff the consumer expects from discovering her value for each firm's product exceeds her outside option.

That firms' bidding strategies are symmetric in states in which their attractiveness coincides is obvious (and hence the proof is omitted). ■