Sequential Learning with Endogenous Consideration Sets*

Daniel Fershtman[†] Alessandro Pavan[‡]

June 2020

Abstract

We study the problem of a decision maker alternating between exploring existing alternatives and searching for new ones. We show that the decision to search depends on the composition of the consideration set only through the information the latter contains about the probability of finding new alternatives. When the search technology is stationary, or improves over time, search is equivalent to replacement. With deteriorating technologies, instead, alternatives are revisited after search is launched and each expansion is treated as if it were the last one. The analysis yields a formula for pricing new alternatives and/or the option to expand the consideration set in the future. We also show how to accommodate for certain irreversible choices that admit as special case a generalization of Weitzman's (1979) "Pandora's boxes" problem in which the set of boxes is endogenous and each alternative may require multiple explorations before it can be selected.

*The paper supersedes a previous version circulated under the title "Searching for Arms". For useful comments and suggestions we thank Dirk Bergemann, Eddie Dekel, Laura Doval, Kfir Eliaz, Stephan Lauermann, Benny Moldovanu, Philip Reny, Eran Shmaya, Rani Spiegler, Bruno Strulovici, Asher Wolinsky, and seminar participants at University of Bonn, Cornell University, Hebrew University of Jerusalem, HEC Paris, Paris Dauphine, Tel Aviv University, Toulouse School of Economics, Yale University, Warwick Economic Theory Workshop, 3rd Columbia Conference on Economic Theory (CCET), and 30th Stony Brook International Conference on Game Theory. We also thank Matteo Camboni for excellent research assistance. Daniel Fershtman gratefully acknowledges funding by the German Research Foundation (DFG) through CRC TR 224 (Project B1). The usual disclaimer applies.

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

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

1 Introduction

An important implication of limited attention is that a decision maker (hereafter, DM) can consider only a subset of the alternatives available to her. This scarcity gives rise to the concept of "consideration sets," (hereafter, CS) which has recently been applied to a range of areas of study.[1]

Classic models of sequential learning often involve a DM exploring a fixed set of options with unknown returns. Yet, a ubiquitous feature of dynamic decision problems is that the set of options in a DM's CS is not predetermined. Rather, the DM may choose to expand her CS by seeking out additional options as part of her learning process, in response to information she has gathered. This paper studies the tradeoff between learning about alternatives within the DM's CS and expansion of the latter through search for additional options.

Consider, for example, a consumer's online search for a product. Time is costly, and the consumer faces a potentially huge amount of options, displayed sequentially (and, from the perspective of the consumer, stochastically) across multiple pages and/or different websites. The information displayed in each ad is limited. By clicking on any of the ads, the consumer is directed to the vendor's webpage from which she can gather further information and, when desirable, finalize a purchase. Rather than clicking on one of the ads on the existing page (or on one of the pages previously visited), the consumer can also move on to the next page of results, or switch to a different website. The decision to do so entails some time and/or cognitive cost, and depends on the information the consumer has gathered about the products thus far, as well as the relevance and number of suitable new options she expects to find by expanding her CS. Assuming the consumer behaves "optimally," under what conditions does she seek additional new options to explore? If she does move beyond the first page of results, or switches to a different website, is she likely to go back to results she has already explored? How do such decisions depend on the gradual resolution of uncertainty and on the properties of the search technology (e.g., on how search engines distribute ads across multiple pages). These questions are relevant for both advertisers on a platform and for the platform's design of its search environment (e.g., how many ads to display on each page).[2]

[1] The concept of consideration sets also has a long tradition in the marketing literature.

[2] Consumers typically limit their attention to a relatively small number of websites when buying products over the internet (De los Santos, Hortaçsu, and Wildenbeest, 2012). According to Epstein and Robertson (2015), a "recent analysis of ∼300 million clicks on one search engine found that 91.5% of those clicks were on the first page of search results" although not necessarily in the order by which the

Similarly, parents choosing a prospective school district may trade-off further evaluations of schools they are already aware of (or that are in their vicinity) with the expansion of their CS by searching for new options in locations they may not know of ex-ante. R&D often involves pursuing a number of alternative technologies whose ability to produce the desirable breakthrough is unknown, but also spending time and/or effort to search for new alternatives to subsequently explore. The tradeoff between the two activities (exploration and expansion of the CS) evolves over time based on the development of existing projects and the ability to find new alternatives.

To study this tradeoff, we introduce a model of sequential learning among an endogenous set of alternatives. In each period, the DM can either focus attention on a single alternative within her current CS or choose to expand it. We refer to such expansion as "search" for new alternatives, as we assume it is costly and its outcome may be stochastic. Focusing attention on an alternative generates a signal that is informative about the alternative's value (and independent of other alternatives) and may yield a payoff (positive or negative). The decision to expand the CS (i.e., to search) triggers the discovery of a stochastic set of new alternatives as a function of the state of the search technology. This state may evolve over time based on the past outcomes of search and on the information the DM possesses about the process governing the search results. For example, the state of the search technology may be stationary, yielding i.i.d. sets of new options from a known distribution. Alternatively, it may evolve over time in a non-stationary manner reflecting the DM's beliefs about the set of alternatives outside her CS.

Our environment and characterization of the DM's optimal policy extend Weitzman's (1979) classic problem, and its solution based on independent "reservation prices," to an environment in which the CS can be sequentially expanded and is endogenous to the DM's problem. Alternatives are assigned independent reservation prices, which are used to determine the order in which attention is allocated among them. These reservation prices are equivalent to those which determine the optimal policy in the absence of the option to expand the CS. Expansion of the CS is also assigned a reservation price, which depends on the information the DM has about the "state" of the existing search technology, but also on the exploration and future expansion policy the DM plans to follow once she starts the search. At any given time, the DM's decision under the optimal policy is determined by the action – focusing on one of the alternatives within the CS

ads were displayed on the same page (see, e.g., Jeziorski and Segal, 2015). Millward Brown examined how likely customers are to view products beyond the first page of search results on Amazon, and found that approximately 70% did not click past the first page of results (see clavisinsight.com 2015).

or expanding the latter – whose reservation price is the greatest.

Rather than committing to a CS up front and then proceeding to evaluate the alternatives in the optimal order, the DM chooses when to expand the CS based on the results of her past explorations and past search (i.e., expansion) outcomes. We show that, despite the fact that the size and composition of the CS may affect the DM's beliefs about the results of future searches and in particular the likelihood of finding alternatives similar to those already in the CS, in expectation, the relative attention the DM allocates to any pair of alternatives within the CS does not vary with the results of past search outcomes.

Importantly, the reservation price the DM assigns to the option to expand the CS does not coincide with the *value* of the expansion. That is, the decision to expand is typically sensitive to only partial information about the prospective alternatives the DM expects to add to the CS, even if such information is relevant for the expected continuation payoff under the optimal policy. The reservation price corresponding to the expansion of the CS is linked to the reservation prices of the new alternatives expected to be introduced in the future. We show that if the search technology is stationary (or "improving" in an appropriate sense made precise below) then alternatives in the CS at the time of the expansion never receive attention in the future, and hence are effectively discarded once the set is expanded. That is, search is equivalent to replacement. When, instead, the search technology "deteriorates" over time, alternatives are put on hold and returned to at a later stage, after the CS has been expanded. Furthermore, in this case, the DM proceeds as if each decision to expand were the last one (that is, the reservation price of expansion takes into account only information about alternatives that are expected to be added to the CS as the result of the current search, as if further expansions were not feasible).

More generally, we show that the decision to expand the CS depends on the composition of the set only through the information that the latter contains about the probability of finding new alternatives of different types. This result holds despite the fact that the alternatives in the set may share similarities with those that are expected to be found through the expansion and despite the fact that the opportunity cost of searching for new alternatives (which is linked to the value of continuing to explore the current set) depends on the entire composition of the CS. Likewise, at any point in time, the relative attention allocated to any pair of alternatives in the CS is invariant to the search technology, and in particular is independent of the probability that search will bring new alternatives similar to those already under consideration. Finally, improvements in the search technology yielding an increase in the probability that search brings alternatives of positive expected value (vis-a-vis the outside option) need not affect the decision to search, even at histories at which, prior to the improvement, the DM is indifferent between searching and continuing with the current CS.

The results are not specific to sequential learning. They apply to a broad class of dynamic experimentation problems with an endogenous set of alternatives. We start by considering the case in which the DM can revert her decision at all periods, as in the multi-armed bandit literature. We prove that the optimal policy takes the form of an index rule with a special index for search. The result is related to the optimality of index policies in the branching-arm literature (e.g., Weiss, 1988, Weber, 1992, and Keller and Oldale, 2003). Our contribution is in identifying conditions under which indexability also applies to the sequential learning problem with endogenous CSs under examination. To the best of our knowledge, our proof of indexability is new and uses a recursive representation of the index of search which also yields a novel representation of the DM's payoff under the optimal policy (with or without the possibility of expansion). The representation can be used to price the access to new alternatives as well as the option to expand the set in the future.[3]

Next, we consider the case where, in addition to focusing attention on existing alternatives and searching for new ones, the DM can also *irreversibly* commit to one of the alternatives in the CS, putting an end to the exploration. In general, the irreversibility of choice is known to preclude a tractable solution. We identify a condition under which the optimal policy remains indexable, which admits as a special case a generalization of Weitzman's (1979) original problem to a setting in which (a) the CS (i.e., the set of boxes) is endogenous, (b) learning the value of an alternative in the CS may require multiple explorations, and (c) the DM may derive a positive payoff from exploring an alternative (possibly higher than the value she derives from irreversibly committing to it) for an arbitrary large (and possibly infinite) number of periods.[4]

[3]The reason why indexability is not obvious is that search is a "meta-arm" bringing alternatives with correlated returns that one needs to process optimally. Problems in which alternatives correspond to "meta arms", i.e., to sub-decision problems with their own sub-decisions, 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 an index solution. 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 such difficulties in the Online Appendix.

[4]The condition is based on a certain "better-later-than-sooner" property guaranteeing 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,

In Section 2, we kick off by considering the simplest extension of Weitzman's (1979) model to a setting with endogenous CSs. As in Weitzman's model, we assume that it takes a single exploration to learn an alternative's value and that exploring an alternative without committing to it comes at a cost. We characterize the "prizes" that guide the DM's sequential alternation between the exploration of the alternatives already in the CS and the expansion of the latter. We also derive an "eventual-purchase theorem" in the spirit of Choi, Dai and Kim (2018) (see also Armstrong and Vickers, 2015 and Armstrong, 2017) that relates the probability that each alternative is eventually selected to the primitives of the search problem (realized values and search technology) and discuss how the endogeneity of the CS affects the selection probabilities.

The possibility to expand the CS may change quite radically some of the comparative statics of the canonical model. To illustrate this possibility in concrete terms, we consider a stylized market in which three firms advertise on a platform. Each firm has two different products, and a representative consumer seeks to purchase at most one of the firms' products. As in Weitzman (1979), the consumer must inspect a product to finalize the purchase. We consider two cases: one with a fixed CS, and one with an endogenous set. In the fixed-consideration-set case, there are four advertising slots, with equal visibility. Each firm is endowed with one slot, but one of the firms, chosen by the platform before the consumer's search begins, is given a second slot to advertise her other product. As in Weitzman (1979), the consumer sees all four ads (i.e., the identity of the firm that was awarded the second slot) before starting the exploration. She then sequentially decides between inspecting a product (by clicking on a firm's ad) and stopping and then either choosing an inspected product or her outside option (this version of the problem is thus identical to the one in Weitzman's problem). In the second environment, instead, ads are displayed on two different pages. Each firm advertises on the first page, and one of the firms also advertises on the second page. The identity of this firm is unknown to the consumer who may have correct beliefs about the probability the slot is assigned to each firm but nonetheless does not know the realization of the relevant risk. This environment thus corresponds to a sequential learning problem with an an endogenous CS. The consumer may expand her CS by visiting the second page but this entails a cost (the magnitude of which may be small and possibly due only to the postponing of the exploration of one of the alternatives already in the CS). The consumer's optimal

or improves over time. This "better-later-than-sooner" 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 approach is, however, quite different and accommodates for the possibility that the set of alternatives is endogenous.

inspection strategy is an index policy with a special index for the decision to explore the second page. When the CS is exogenous (and contains all items), each firm benefits from an increase in the probability she receives the additional slot that permits it to display its second product. This is not the case when the CS is endogenous: a firm may be strictly worse off when the probability the slot on the second page is assigned to it increases.

The rest of the paper is organized as follows. The remainder of this section discusses the related literature. Section 2 considers the extension of Weitzman's model to a setting with an endogenous set of boxes and contains the example discussed above. Section 3 presents the general model, contains all the main results, and discusses how the optimal policy in Weitzman (1979) must be amended to accommodate not only for endogenous CSs but also for gradual resolution of uncertainty. Section 4 discusses several simple extensions, whereas Section 5 concludes. Most of the proofs are relegated to the Appendix.

1.1 Relation to the literature

The paper is related to the fast-growing literature on CSs. A number of papers in the marketing literature study the formation of CSs (e.g., Hauser and Wernerfelt, 1990; Roberts and Lattin, 1991). Eliaz and Spiegler (2011) study implications of different CSs on firms' behavior, assuming such sets are exogenous. Masatlioglu, Nakajima, and Ozbay (2012) and Manzini and Mariotti (2014), instead, identify CSs from choice behavior. Caplin, Dean, and Leahy (2018) provide necessary and sufficient conditions for rationally inattentive agents to focus on a subset of all available choices, thus endogenizing the CSs. Simon (1955) considers a sequential search model, in which alternatives are examined until a "satisfying" alternative is found.[5] Our analysis complements the one in this literature by providing a dynamic micro-foundation for endogenous CSs. Rather than committing to a CS up front and proceeding to evaluate the alternatives in it, the DM expands the CS over time, in response to the results obtained from the exploration of the alternatives in the set.

The paper is also related to the literature on sequential learning in settings in which the DM explores one alternative at a time. Most closely related are Ke, Shen, and Villas-Boas (2016), Austen-Smith and Martinelli (2018), Ke and Villas-Boas (2019),

[5] 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.

and Gossner, Steiner and Stewart (2019).[6] Related is also Che and Mierendorff (2019) who study the optimal sequential allocation of attention to two different signal sources biased towards alternative actions. In all the above papers, the set of alternatives is fixed ex-ante. In our model, instead, the DM chooses when to expand the CS based on the information she has collected about the alternatives already in it.

As mentioned above, our results cover as a special case an extension of Weitzman's (1979) classic problem in which the set of boxes is endogenous and where exploration leads to a gradual resolution of uncertainty. 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. In independent work, Greminger (2020) considers a version of Weitzman's (1979) problem in which the DM can bring new alternatives to the CS over time. His problem is a special case of ours, both in terms of the search technology and of the DM's payoff.

Related is also Garfagnini and Strulovici (2016), who study how successive (forward-looking) agents experiment with endogenous technologies. Trying a "radically" new technology reduces the cost of experimenting with similar technologies, which effectively expands the space of affordable technologies.[7] While both their work and ours consider environments in which the set of alternatives/technologies is expanded over time, the two models, as well as the analysis and questions addressed, are fundamentally different. Schneider and Wolf (2019) study the time-risk tradeoff of an agent who wishes to solve a problem before a given deadline, and allocates her time between implementing a given method and developing (and then implementing) a new one. Related is also Fershtman and Pavan (2020), which considers the effects of "soft" affirmative action on minority recruitment in a setting in which the candidate pool is endogenous.[8]

2 Pandora's problem with endogenous boxes

Consider the following extension of Weitzman's (1979) "Pandora's problem" in which a DM optimally constructs her CS over time, accounting for the costs of expanding it.

A DM must make a single choice among alternatives. An alternative is characterized

[6] Ke and Villas-Boas (2019) study optimal search before choice in a setting where the optimal policy is not indexable and identify various properties of the optimal policy.

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

[8] The problem in that paper is a special version of the one considered in the present paper.

by a pair (F, λ), where F denotes the distribution over the alternative's unknown value, u, and where λ denotes the cost of inspecting the alternative (here, as in Weitzman's setting, we consider the simple case where the value of each alternative is revealed immediately upon its first inspection; this assumption is relaxed in the general model in Section 3). Initially, the DM is aware of only a subset of alternatives – this is her initial CS – and has an outside option, normalized to zero.

The DM's CS is endogenous to her decision problem, and adding alternatives into the CS is increasingly costly. More precisely, at each period $t = 0, 1, ...$, the DM can either (a) search for an additional alternative to add to her CS, (b) inspect an alternative to learn its value, or (c) stop and either recall an observed payoff u from one of the inspected alternatives or take her outside option (in which case the decision problem ends). Search expands the CS, introducing a new alternative whose characteristics (F, λ) are drawn from a set $\mathcal{A}$ according to some known distribution $\mathcal{F}$ (we consider here the simple case where search yields a single new alternative in each period and where the draws from $\mathcal{A}$ are independent; such assumptions are relaxed in the general model in Section 3). The cost of adding an alternative to the CS is equal to $c(m)$, where m is the number of past searches (i.e., expansions of the CS), and c is a positive, increasing function. Besides the direct costs of inspecting and adding new alternatives, the DM discounts the future according to δ.

In such a setting, how should the DM optimally (and dynamically) balance the trade-off between bringing new alternatives to the CS to inspect in the future and inspecting her current options? The following Proposition describes the optimal rule. Let

$$\mathcal{I}(F, \lambda) = \frac{-\lambda + \delta \int_{\mathcal{I}(F,\lambda)}^{\infty} u \mathrm{d}F(u)}{1 + \frac{\delta}{1-\delta}\Pr(u > \mathcal{I}(F, \lambda)|F, \lambda)}. \tag{1}$$

denote the "reservation price" (equivalently, the index) of a box of type (F, λ), as defined in Weitzman (1979) and, for any $l \in \mathbb{R}$, let $\mathcal{A}(l) \equiv \{(F, \lambda) \in \mathcal{A} : \ \mathcal{I}(F, \lambda) > l\}$ denote the set of box types whose reservation price exceeds l.[9] Let

$$\mathcal{I}^{\mathcal{S}}(m) = -c(m) + \frac{\delta \int_{\mathcal{A}(\mathcal{I}^{\mathcal{S}}(m))} \left(-\lambda + \delta \int_{\mathcal{I}^{\mathcal{S}}(m)}^{\infty} u dF(u)\right) \mathrm{d}\mathcal{F}(F, \lambda)}{1 + \int_{\mathcal{A}(\mathcal{I}^{\mathcal{S}}(m))} \left(\delta + \int_{\mathcal{I}^{\mathcal{S}}(m)}^{\infty} \frac{\delta^2}{1-\delta} dF(u)\right) \mathrm{d}\mathcal{F}(F, \lambda)}. \tag{2}$$

[9]Weitzman defines the reservation prices as the solution to $\lambda = \delta \int_{\mathcal{I}(F,\lambda)}^{\infty} (u - \mathcal{I}(F, \lambda)) \mathrm{d}F(u) - (1 - \delta)\mathcal{I}(F, \lambda)$, which yields $\mathcal{I}(F, \lambda) = \left[-\lambda + \delta \int_{\mathcal{I}(F,\lambda)}^{\infty} u \mathrm{d}F(u)\right] / [1 - \delta + \delta \Pr(u > \mathcal{I}(F, \lambda)|F, \lambda)]$. The reservation prices in (1) are the same, but multiplied by $(1 - \delta)$ to facilitate a comparison with the more general model introduced in the next section.

denote the reservation price of the option to expand the CS (equivalently, the "search index").[10]

Proposition 1. *The optimal policy is the following: (i) Expand the CS if $\mathcal{I}^S(m)$ is positive and is greater than the reservation price $\mathcal{I}(F, \lambda)$ of all uninspected boxes in the CS and the value u of all inspected boxes. (ii) Inspect any of the boxes in the CS whose reservation price is the highest, provided that such price is positive and is greater than $\mathcal{I}^S(m)$ and the value of all inspected boxes. (iii) Stop and choose any of the inspected boxes whose observed value is the highest provided that such value is positive and higher than $\mathcal{I}^S(m)$ and the reservation price of any uninspected box in the CS. (iv) Stop and take the outside option if $\mathcal{I}^S(m)$, the reservation price of all uninspected boxes, and the value of all inspected boxes are all negative.*

The reservation prices $\mathcal{I}(F, \lambda)$ have the following interpretation. Suppose there are only two alternatives. One is the alternative i characterized by (F, λ), and the other is a hypothetical alternative, j, with a known value u_j. The reservation price is the value of u_j, multiplied by $(1-\delta)$, for which the DM is indifferent between taking j and inspecting the alternative i while maintaining the option to recall j once the value u_i is discovered.

The reservation price of search extends this interpretation as follows. Suppose there are two options: the hypothetical alternative j, and the option of expanding the CS. The reservation price of search is the value u_j of the fictitious alternative j for which the DM is indifferent between taking j right away, and expanding the CS, maintaining the option to take j either (a) once the characteristics of the new alternative (F, λ) are discovered and $u_j \geq \mathcal{I}(F, \lambda)$, or (b), in case $u_j < \mathcal{I}(F, \lambda)$, after the value u_i of the new alternative is learned and $u_i \leq u_j$.

Rather than committing to a CS up front and then exploring its alternatives, the DM optimally chooses the time at which to expand the CS, based on the information the DM has learned over time. Note that $\mathcal{I}^S$, which we interpret as the reservation price corresponding to the expansion of the CS, is independent of any information about the composition of the DM's current CS beyond the number of times it has been expanded, and is easy to calculate (see, e.g., Subsection 2.2). Importantly, as anticipated in the Introduction, $\mathcal{I}^S$ need not coincide with the DM's expected value of expanding the CS.

An immediate implication of Proposition 1 is that, despite the fact that search may bring alternatives that are more similar in nature to certain alternatives in the set than

[10] The characterization of the search index follows from the arguments in the proofs of Theorems 1 and 2.

others, the relative likelihood of selecting any pair of alternatives in the set remains constant over time.

The problem described above, and its solution, are a special case of the more general problem studied in Section 3. Before turning to the general model, we describe below two applications of this simple problem.

2.1 Consumer search and eventual purchases

Despite the importance of sponsored search in modern business activities, the few models of online consumer search that have been developed remain quite restrictive.[11] For example, the literature has typically assumed that consumers click ads sequentially in the order they are displayed, and that click-through-rates depend on positions but not on the ads displayed at the various positions. Such assumptions do not appear to square well with empirical findings.[12] An important feature of online consumer search is that the consumer is presented with a potentially huge amount of search results, which are displayed in a sequence across multiple pages. Most of the options are initially unobservable, and require the consumer to incur the (time, or mental) cost of scrolling through multiple pages. Clearly, consumers do not read all of the results. Instead, the set of search results a consumer reads (and considers clicking on) is partial, and fashioned by the way ad links are assigned by search engines to different pages.

As a first step toward a better understanding of consumer search in such markets, we apply Proposition 1 as follows. When a consumer enters a query on a search engine, a first list of alternatives is presented (the ones displayed on the first page). Reading the text displayed on a page is costly, and adds the alternatives displayed on the page to the consumer's CS. The consumer may also click on one of the alternatives in the CS (that is, on one of the links displayed on one of the pages she has read already) in which case she is directed to a vendor's website (also at a cost). Once the consumer visits a vendor's website, she learns her value for the vendor's product or service (think of such values as net of prices, so that the discovery of the prices is incorporated into the learning process). At any point in time, the consumer can then stop and purchase a product among those offered by those vendors she visited. To recap, at each point in

[11] Sponsored search advertising accounts for a large fraction of Internet advertising revenues (see, e.g., Edelman, Ostrovsky, and Schwarz, 2007).

[12] For example, in an empirical analysis of consumer search in online advertising markets using data from Microsoft Live, Jeziorski and Segal (2015) find that almost half of the users who click on a link do not click in sequential order of positions, and that click-through-rates do depend on the identity of competing ads.

time, the consumer can either read the information displayed in response to the search query (i.e., search), click on one of the texts/links she has read already to be directed to the corresponding vendor's webpage (i.e., learn about a specific alternative's value), or purchase a product, in which case the decision problem ends.[13]

Suppose the results are indexed by natural numbers in increasing order $i = 1, 2, \ldots$. When (search) result i is read, and hence added to the consumer's CS, the consumer receives information (F_i, λ_i) about the result under consideration. Such information is drawn from a set $\mathcal{A}$ according to a distribution $\mathcal{F}$ that reflects the consumer's beliefs over the way the search results are displayed by the platform. The information F_i denotes the distribution over the consumer's ultimate value u_i for the corresponding product, whereas λ_i represents the cost the consumer assigns to visiting the vendor's webpage and learning her value u_i for the vendor's product (heterogeneity in both F_i and λ_i may reflect the consumer's prior experiences, the vendors' reputations, as well as the consumer's expectations about the ease of navigating within the vendors' websites).

Denote the index of product i, once added to the CS, by $\mathcal{I}_i \equiv \mathcal{I}(F_i, \lambda_i)$. The cost of reading a result is $c(m)$, where c is increasing and m is the number of past searches. Let $\mathcal{I}_i^{\mathcal{S}} \equiv \mathcal{I}^{\mathcal{S}}(i-1)$, with the function $\mathcal{I}^{\mathcal{S}}$ as defined in (2). In this formulation, we are assuming that the consumer reads the results in the order with which they are presented. Under such an assumption, the order by which the results are displayed matters, but how the search engine bundles different results on different pages is inconsequential (see the next subsection for a case where the bundling plays a role). Importantly, the consumer need not click on results in the same order as she reads them.

The consumer's initial outside option, normalized to zero, is captured by product $i = 0$ (purchasing product 0 is therefore interpreted as taking the outside option).

This formulation corresponds to a special case of the model in the previous subsection. With the consumer's optimal behavior described by Proposition 1, the model can then be used to endogenize the probability with which the consumer selects the various products. Choi, Dai and Kim (2018) (and, independently, Armstrong, 2017) derive a static condition characterizing eventual purchase decisions based on a comparison of "effective values," in a model where consumers face a fixed set of alternatives, and therefore the optimal policy is the one characterized by Weitzman. Building on Proposition 1, Proposition 2 below shows how to extend the characterization of the eventual purchasing

[13]That is, in this environment, search coincides with reading, whereas inspecting an alternative coincides with clicking on its link. Note that this formulation implicitly assumes that the consumer does not click on a result without first reading it, and that a purchase cannot be made without first visiting a vendor's website. Both assumptions seem quite natural in this context.

probabilities to the case of endogenous CSs.

Proposition 2. *For all $i \geq 1$, let $w_i = \min\{\mathcal{I}_i, u_i, \mathcal{I}_i^{\mathcal{S}}\}$* denote the "effective value" of product i. *The consumer purchases product i if, for all $j \neq i$, $w_j < w_i$ (and only if $w_i \geq w_j$, for all $j \neq i$).*

Proposition 2 provides a micro-foundation for why (and the extent to which) higher positions imply higher eventual purchase probabilities, a property typically assumed exogenously in existing models. As in Choi, Dai and Kim (2018), consumers' eventual purchase decisions are determined by a static comparison of effective values, as in canonical discrete-choice models. However, contrary to Choi, Dai and Kim (2018), such effective values now account for the endogenous order by which the various alternatives are found, and hence can be used to study the effects of varying such order (e.g., by platforms).

The result follows from the following monotonicity property of the optimal policy in Proposition 1: If a consumer reads result i (equivalently, engages into the i-th search), then the reservation prices and discovered values of all options previously encountered must be no greater than $\mathcal{I}_i^{\mathcal{S}}$. Hence, if after reading the i-th result, $\mathcal{I}_i \geq \mathcal{I}_i^{\mathcal{S}}$, the consumer proceeds to inspect product i. Similarly, if i is inspected and $u_i \geq \mathcal{I}_i$, the consumer will proceed to purchase product i. Note that an immediate implication is that, if the cost of reading $c(\cdot)$ is strictly increasing, all other things equal, the further a result i is down the list, the more likely it is that its effective value coincides with $\mathcal{I}_i^{\mathcal{S}}$. Therefore, we have the following result:

Corollary 1. *Suppose that the cost of reading $c(\cdot)$ is strictly increasing. Then, conditional on being read, results further down the list are more likely to be clicked on and purchased immediately.*

A key difference with respect to an environment with an exogenous set of alternatives is that reading additional results is a substitute for inspecting those that have been read already. This property leads to very different dynamics. For example, suppose there is a fixed set of alternatives (and hence search is à la Weitzman), and consider the effects of a reduction in inspection costs or an increase in products' values that raise the reservation prices of all products uniformly. When the set of alternatives is exogenous, the eventual purchase probabilities are unaffected by such a change. When, instead, alternatives are brought endogenously to the CS, the aforementioned change shifts the balance between expansion of the CS and inspection of its alternatives, and therefore has implications for the eventual purchase decisions. The following subsection provides an illustration of such new effects and their implications.

2.2 Consumer search and multi-product competition

Consider the following stylized model of multi-product competition. There are three firms advertising on a platform. Each firm $i \in \{1, 2, 3\}$ has two different products with similar characteristics (meaning that (F, λ) is the same for both products; the value the buyer attaches to the two products may be different though), and a representative consumer seeks to purchase at most one of these product. Formally, each firm i's product is characterized by the pair (v_i, p_i), where v_i represents the product's value to the consumer in case it is a good match for her tastes, and p_i is the probability with which the product is a good match for the consumer. If a product is not a good match, its value is 0.[14] Hence, in this problem, $u_i \in \{0, v_i\}$ and $Pr(u_i = v_i) = p_i$. The consumer learns v_i and p_i by reading the product's ad, but must inspect the product (e.g., by clicking on the ad to be directed to the vendor's website) to learn whether $u_i = v_i$ or $u_i = 0$, and in order to finalize the purchase (as in the rest of the search literature, we assume that the consumer cannot purchase the product without visiting the vendor's website). Whether or not a product is a good match for the consumer is independent across products. For example, firms may be hotel chains advertising their hotels in different locations, on a platform such as Kayak.

We consider two environments. In the first, the consumer's CS is fixed in advance, while in the second it is endogenously determined as part of the consumer's optimal policy. As it will become clear from the discussion below, the two environments have very different implications when it comes to the profitability of securing more ad space.

Fixed CS. There are four slots for advertising products, each equally visible. Each firm receives a single slot, and the remaining slot, to be used for advertising an additional product, is assigned to one of the firms randomly (such randomness may reflect uncertainty the firms face about the platform's objectives when choosing the assignment of the slots). Specifically, suppose the probability with which firm i gets a second slot is $\gamma_i \in [0, 1]$, with $\gamma_1 + \gamma_2 + \gamma_3 = 1$. In this fixed-consideration-set environment, the identity of the firm that receives the additional slot is determined ex-ante, i.e., prior to the consumer visiting the platform. Once the consumer visits the platform, she sees all four products (that is, all four products are in the consumer's CS at the beginning of the

[14] Once again, v is the consumer's value, net of the product's price.

exploration/inspection process). The consumer sequentially decides which product to inspect and when to stop, at which point she either chooses an inspected product or her outside option (normalized to zero). To keep things simple, suppose the consumer incurs no cost for inspecting a product other than the time cost of postponing the purchase of the final product. The consumer discounts time geometrically with a discount factor δ. In other words, the consumer faces a standard problem à la Weitzman (1979), with four products: (v_1, p_1), (v_2, p_2), (v_3, p_3), and (v_j^s, p_j^s), where $j \in \{1, 2, 3\}$ is the identity of the firm selected by the platform to display the second ad and where (v_j^s, p_j^s) are the characteristics of the selected firm's second product. To make things simple, assume that $(v_j^s, p_j^s) = (v_j, p_j)$ for all $j \in \{1, 2, 3\}$, meaning that the two products that each firm displays are identical in the eyes of the consumer prior to visiting the firm's webpage and learning whether each product is a good match or not.

It is easy to verify that, in this environment, the reservation prices are equal to

$$\mathcal{I}(v, p) = \frac{(1-\delta)\delta p v}{1-\delta+\delta p}. \tag{3}$$

As proved in Weitzman (1979), the optimal policy is to inspect products in descending order of their reservation prices, stopping when the remaining reservation prices are all smaller than the maximal realized value among the inspected products.

Firms are interested in maximizing the probability with which one of their products is selected. For simplicity, here we assume that each firm makes equal profits on each of its two products. Clearly, in this environment, any firm i benefits from an increase in the probability γ_i it is given a second slot.

Endogenous CS. Now suppose search results are displayed on two separate pages. All three firms advertise on the first page, but one of them, selected at random by the platform, is also offered the possibility to advertise on the second page.[15] Thus, in this case, there are three products in the consumer's initial CS, one for each firm. The consumer has the option of expanding her CS by visiting the second page. If the consumer does so, the identity of the firm selected to display the additional ad is revealed to the consumer. As indicated above, the probability with which each firm i gets to display the additional ad is $\gamma_i \in [0, 1]$. By visiting the second page, the consumer thus adds a new product to her CS, with each product (v_i, p_i) selected with probability γ_i.

[15]This stylized setting easily extends to one with more than two pages and/or with a stochastic number of slots per page.

Again, there is no direct cost for either inspecting a product or for expanding the CS. The cost of each decision is simply the opportunity cost of waiting a period to make a different decision, with each period discounted according to δ.[16]

In this case, the consumer's optimal policy is given by Proposition 1, and is based on the comparison of Weitzman's reservation prices (1) with the search index (2). As in the case of exogenous CS, the reservation price of each product is given by (3).

Without loss of generality, suppose firms are ordered in decreasing order of reservation prices $\mathcal{I}_i$. It is easy to verify from (2) that, in this case, the reservation price of expanding the CS (equivalently, the search index) is equal to

$$\mathcal{I}^S = \delta^2 \max_{k\in\{1,2,3\}} \left\{ \frac{\sum_{i=1}^{k} \gamma_i p_i v_i}{1+\sum_{i=1}^{k} \gamma_i \delta \left(1+\frac{p_i \delta}{1-\delta}\right)} \right\}. \tag{4}$$

Importantly, note that $\mathcal{I}^S$ is different from the consumer's expected payoff from expanding her CS. Whether or not $\mathcal{I}^S$ takes into account the benefits of inspecting a particular type of product depends on the relationship between the reservation price of that product and $\mathcal{I}^S$ itself. In particular,

$$\mathcal{I}^S = \begin{cases} \frac{\delta^2 \gamma_1 p_1 v_1}{1+\gamma_1 \delta\left(1+\frac{p_i\delta}{1-\delta}\right)} & \text{if } \mathcal{I}_2 < \frac{\delta^2 \gamma_1 p_1 v_1}{1+\gamma_1 \delta\left(1+\frac{p_i\delta}{1-\delta}\right)} < \mathcal{I}_1 \\ \frac{\sum_{i=1}^{2} \gamma_i p_i v_i}{1+\sum_{i=1}^{2} \gamma_i \delta\left(1+\frac{p_i\delta}{1-\delta}\right)} & \text{if } \mathcal{I}_3 < \frac{\sum_{i=1}^{2} \gamma_i p_i v_i}{1+\sum_{i=1}^{2} \gamma_i \delta\left(1+\frac{p_i\delta}{1-\delta}\right)} < \mathcal{I}_2 \\ \frac{\sum_{i=1}^{3} \gamma_i p_i v_i}{1+\sum_{i=1}^{3} \gamma_i \delta\left(1+\frac{p_i\delta}{1-\delta}\right)} & \text{if } \frac{\sum_{i=1}^{3} \gamma_i p_i v_i}{1+\sum_{i=1}^{3} \gamma_i \delta\left(1+\frac{p_i\delta}{1-\delta}\right)} < \mathcal{I}_3 \end{cases}. \tag{5}$$

Suppose again that firms are interested in maximizing the probability with which a product of theirs is selected, given the consumer's optimal policy. Perhaps surprisingly, when the latter's CS is endogenous, a firm may suffer from an increase in the probability it is given a second slot, even if this implies that its competitors are less likely to display their ads, and even if all products have a positive expected value in the eyes of the consumer.

For concreteness, let $\delta = 0.9$ and suppose the three types of products are: $(v_1, p_1) = (10, \frac{1}{10})$, $(v_2, p_2) = (3, \frac{1}{3})$, and $(v_3, p_3) = (2, \frac{1}{2})$. Note that the lotteries corresponding to each of the products have the same mean value, but are mean preserving spreads

[16] The assumption that expanding the CS takes the same amount of time as inspecting a product is completely innocuous and made only for simplicity. In Section 4, we discuss how the results can be amended to accommodate for the possibility that the time it takes to evaluate an alternative and to expand the CS may differ.

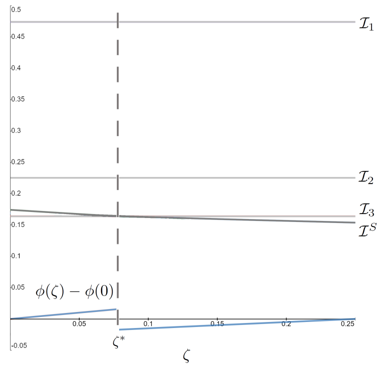

Figure 1: The change $\phi(\zeta) - \phi(0)$ in the probability with which a product of firm 2 is selected, as a function of ζ (in blue). The horizontal gray lines represent the indices $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$, and the dark gray curve represents $\mathcal{I}^S$ as a function of ζ.

of one another; hence $\mathcal{I}_1 > \mathcal{I}_2 > \mathcal{I}_3$. Initially, $\gamma_1 = \gamma_2 = \frac{1}{4}$, and $\gamma_3 = \frac{1}{2}$. Using the above characterization of the indices, it is easily verified that $\mathcal{I}_1 = 0.473$, $\mathcal{I}_2 = 0.225$, $\mathcal{I}_3 = 0.163$, and

$$\mathcal{I}^S = \frac{\gamma_1 p_1 v_1 + \gamma_2 p_2 v_2}{1 + \gamma_1 \delta \left(1 + \frac{p_1 \delta}{1-\delta}\right) + \gamma_2 \delta \left(1 + \frac{p_2 \delta}{1-\delta}\right)} = 0.174.$$

Also note that $\mathcal{I}_3 < \mathcal{I}^S < \mathcal{I}_2$. As a result, $\mathcal{I}^S$ does not take into account the benefits from inspecting firm 3's additional product, in case firm 3 is the one that is selected to display on the second page.

Now suppose the search engine increases the probability firm 2 is selected on the second page, at the expense of firm 1. Precisely, suppose that γ_2 is increased by $\zeta \in [0, 0.25]$ while γ_1 is reduced by the same amount. Let $\phi(\zeta)$ denote the probability that one of firm 2's products is ultimately chosen when firm 2 is given the second ad with probability $\gamma_2 + \zeta$. Figure 1 depicts the change $\phi(\zeta) - \phi(0)$ in the probability that one of firm 2's products is selected as a function of ζ, where $\phi(0) = (1-p_1)(p_2 + (1-p_2)\gamma_2 p_2) = 0.35$. The horizontal gray lines correspond to the indices $\mathcal{I}_1$, $\mathcal{I}_2$, and $\mathcal{I}_3$, whereas the dark gray curve depicts $\mathcal{I}^S$, as a function of ζ. Note that $\mathcal{I}^S$ is decreasing in ζ. This is because $\mathcal{I}_1 > \mathcal{I}_2$. Hence, an increase in ζ implies a lower reservation value for expanding the CS. $\mathcal{I}^S$ starts out above $\mathcal{I}_3$, and intersects $\mathcal{I}_3$ at ζ^* (the vertical dashed line). For $\zeta < \zeta^*$, $\mathcal{I}_3 < \mathcal{I}^S < \mathcal{I}_2$, whereas for $\zeta > \zeta^*$, $\mathcal{I}^S < \mathcal{I}_3$. The function $\mathcal{I}^S(\zeta)$ has a kink at $\zeta = \zeta^*$ (see (5)). For $\zeta \in [0, \zeta^*)$, the CS is expanded before firm 3's product is inspected, while for $\zeta \in (\zeta^*, 0.25]$ the opposite is true. The probability that one of firm 2's products is chosen is equal to $\phi(\zeta) = (1 - p_1)(p_2 + (1 - p_2)(\gamma_2 + \zeta)p_2)$ for $\zeta \in [0, \zeta^*)$ and is equal to $\phi(\zeta) = (1 - p_1)(p_2 + (1 - p_2)(1 - p_3)(\gamma_2 + \zeta)p_2)$ for $\zeta \in (\zeta^*, 0.25]$, with a downward discontinuity at $\zeta = \zeta^*$ equal to $(1-p_1)(1-p_2)p_2p_3(\gamma_2+\zeta^*)$. Furthermore, the downward drop in $\phi(\zeta)$ at $\zeta = \zeta^*$ makes $\phi(\zeta) - \phi(0)$ negative over $(\zeta^*, 0.25]$.

We summarize the implications of the above observations in the following proposition.

Proposition 3. *Suppose that the CS is endogenous. A reduction in the probability that firm 2 is given a second slot may* increase *the overall probability that firm 2 sells one of its products (and hence its profits).*

The result illustrates how novel effects emerge once the endogeneity of consumers' CS is accounted for. The reason why firm 2 suffers from an increase in the probability it is given a second ad is that this reduces the attractiveness for the consumer of expanding the CS, thus inducing her to inspect firm 3's product prior to expanding the CS. When strong enough, this new effect may imply a drop in firm 2's profits.[17]

3 Gradual resolution of uncertainty

In the model in the previous section, the resolution of uncertainty concerning the profitability of each alternative takes a single exploration, as in Weitzman's (1979) original work. In this section, we relax this assumption. We first consider a different version of the problem in which the DM alternates between the exploration of the alternatives in the CS and the expansion of the latter, without having to commit irreversibly to any alternative. This version corresponds to the one in the multi-armed bandit literature, except for the endogeneity of the set of arms. We allow for general processes governing the gradual resolution of uncertainty, establish the optimality of a certain index policy, and relate the dynamics of search and exploration to the search technology and the primitives of the model. We then enrich this problem by adding to it an irreversible choice, like in Weitzman (1979).

For concreteness, we continue to focus on sequential learning. However, it should be clear that the analysis applies more generally to a broader class of problems in which the evolution of the different alternatives may originate in shocks other than the arrival of information, and where the endogeneity of the shocks may reflect for example a preference for variety, or habit formation.

[17]Similar effects emerge in more general settings with more than two pages and more than two slots.

3.1 General framework

In each period $t = 0, ..., \infty$, the DM can either explore one of the alternatives in the CS, or expand the CS by searching for new alternatives. Denote by $C_t \equiv (0, ..., n_t)$ the period-t CS, with $n_t \in \mathbb{N}$. The DM's CS in period 0 is $C_0 \equiv (0, ..., n_0)$.

Given C_t, search brings a (stochastic) set of new alternatives $C_{t+1} \backslash C_t = (n_t + 1, ..., n_{t+1})$ that the DM can explore in subsequent periods and which are added to the current CS. Let $x_{it} \in \{0, 1\}$ denote the decision to focus on alternative i in period t (equivalently, to explore alternative i), with $x_{it} = 1$ if the DM focuses on i and $x_{it} = 0$ otherwise. Let $x_t \equiv (x_{jt})_{j=0}^{\infty}$, and denote by $x \equiv (x_s)_{s=0}^{\infty}$ a complete sequence of attention/exploration decisions. Search involves a direct cost $c_t \geq 0$. Let $y_t \in \{0, 1\}$ denote the decision to search in period t, with $y_t = 1$ if the DM searches and $y_t = 0$ otherwise. Denote by $y \equiv (y_s)_{s=0}^{\infty}$ a complete sequence of search decisions. The period-t overall decision is summarized by $d_t \equiv (x_t, y_t)$. A sequence of decisions $d \equiv (x, y)$ is *feasible* if for all $t \geq 0$, (i) $x_{jt} = 1$ only if $j \in C_t$, and (ii) $\sum_{j=0}^{\infty} x_{jt} + y_t = 1$. The DM may also have an outside option, which we normalize to zero.[18]

Selecting alternative $i \in C_t$ at period t generates a stochastic flow payoff $u_{it} \in \mathbb{R}$, the distribution of which is a function of the alternative's state. The dependence of flow payoffs on the state of the various alternatives, and the description of the search technology, are outlined below.

Each alternative has a *type* ξ, an element of an arbitrary topological space Ξ, which determines the stochastic process governing the evolution of its state. The process corresponding to each alternative of type ξ is Markov and time-homogeneous (i.e., invariant to calendar time). Slightly abusing notation, denote by $\omega^P = (\xi, \theta) \in \Omega^P = \Xi \times \Theta$ the alternative's *current state,* where θ is an element of an arbitrary set Θ. The superscript P is meant to highlight 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, which we will define below.

Depending on the application, θ may have different interpretations. In our sequential learning environment, it is natural to interpret θ as the history of signals the DM has received about the profitability of the alternative under consideration. However, it could also represent the DM's history of beliefs about that alternative, or a sufficient statistics

[18] To allow for this, simply assume alternative 0 is degenerate, with a deterministic payoff equal to zero at all periods. One can also let the search decision correspond to a particular alternative in the CS. While this poses no problem from a mathematical standpoint, the definition of CS favored in the literature (and in the popular language) suggests it is best to keep search separated from the alternatives in the CS.

of the latter (e.g., her current belief). In the case of habit-formation, but also in certain learning environments, it may represent the history of payoffs the DM derived from selecting the alternative in previous periods. Importantly, while ξ is fixed and indexes the alternative's "type," or process, θ evolves over time. Denote by σ the sigma-field associated with Ω^P, and by $H_{\omega^P} \in \Delta(\Omega^P)$ the distribution over Ω^P when the alternative's current state is ω^P. Without loss of generality, we assume that the flow payoffs the DM receives over time are governed by the same process describing the evolution of the alternative's state (that is, they can be represented by a deterministic function of the alternative's state). The first time the DM focuses on an alternative of type ξ, the latter is in state (ξ, θ_0) where, without loss of generality, θ_0 can be taken to be the same across all ξ. We do not describe explicitly the mappings from states to payoffs as the analysis does not require it.

The process governing the cost incurred due to search, the number of new alternatives introduced as the result of search, and their types, is also Markov time-homogeneous. The search technology's state is summarized by ω^S, which consists of the history

$$((c_0, E_0), (c_1, E_1), ..., (c_m, E_m))$$

of past search costs and of alternatives' types added to the CS. Here $m \in \mathbb{N}$ denotes the number of times search has been carried out in the past, and $E_k = (n_k(\xi) : \xi \in \Xi)$ is a vector representing, for each alternative's type ξ, the number of alternatives $n_k(\xi)$ of type ξ found as the result of the k'th search.[19] Denote the set of possible states of search by Ω^S. The distribution over the cost and the set of new alternatives added to the CS is denoted by H_{ω^S}.[20]

Next, we define the *state of the decision problem* as follows. For each $\omega^P \in \Omega^P$, let $\mathcal{S}^P(\omega^P) \in \mathbb{N}$ denote the number of alternatives in the CS in state ω^P. The state of the decision problem is given by the pair $\mathcal{S} \equiv (\omega^S, \mathcal{S}^P)$, where $\mathcal{S}^P : \Omega^P \to \mathbb{N}$ is a function describing, for each $\omega^P \in \Omega^P$, the number of alternatives in the CS in state ω^P. Next let $\Omega \equiv \Omega^P \cup \Omega^S$ and note that $\Omega^P \cap \Omega^S = \emptyset$. With an abuse of notation, we will sometime find it useful to denote the entire state of the decision problem as a function

[19]The first time search is carried out, its state is (c_0, E_0), where c_0 can be taken arbitrarily (it plays no role in the analysis since the cost of the first search is c_1), and E_0 is a description of the types of alternatives in C_0.

[20]Note that this formulation allows the search technology to depend in flexible ways on the results of previous searches. The key assumptions are that the search process is time-homogeneous, and that the outcome of each new search is drawn from H_{ω^S} independently from the idiosyncratic and time-varying component θ of each alternative in the CS.

$\mathcal{S} : \Omega \to \mathbb{N}$ that specifies, for each $\omega \in \Omega$, including $\omega \in \Omega^S$, the number of alternatives, including search, in state ω. We will then denote by $\mathcal{S}_t$ the state of the decision problem at the beginning of period t. Clearly, with this representation, at each period t, there is a unique $\hat{\omega}^s \in \Omega^S$ such that $\mathcal{S}_t(\omega^S) = 1$ if $\omega^s = \hat{\omega}^s$ and $\mathcal{S}_t(\omega^S) = 0$ if $\omega^s \neq \hat{\omega}^s$. The special case where the DM does not have the option to search corresponds to the case where $\mathcal{S}_t(\omega^S) = 0$ for all $\omega^S \in \Omega^S$, all t.

Defining the state of the decision problem this way allows us to keep track of all relevant information, and facilitates the analysis.

3.2 Optimal policy

A policy χ for the above decision problem is a rule governing the decisions in each period – whether to focus attention on an alternative in the CS or search for new ones – based on the available information. Given a sequence of feasible decisions $(d_t)_{t\geq 0}$, the state process $(\mathcal{S}_t)_{t\geq 0}$ generates a natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$. A policy χ is then a $\mathcal{F}_t$-measurable sequence of feasible decisions. Denote by $U_t \equiv \sum_{j=0}^{\infty} x_{jt} u_{jt} - c_t y_t$ the realized period-t net payoff. A policy χ is *optimal* if it maximizes the expected discounted sum of the net payoffs $\mathbb{E}^{\chi}\left[\sum_{t=0}^{\infty} \delta^t U_t | \mathcal{S}_0\right]$, where $\delta \in (0,1)$ denotes the discount factor. To guarantee that the process of the expected net payoffs is well behaved, we assume that for any state $\mathcal{S}$ and policy χ, $\delta^T \mathbb{E}^{\chi}\left[\sum_{s=T}^{\infty} \delta^s U_s | \mathcal{S}\right] \to 0$ as $T \to \infty$.[21]

For each $\omega^P \in \Omega^P$, let

$$\mathcal{I}^P(\omega^P) \equiv \sup_{\tau>0} \frac{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s u_s | \omega^P\right]}{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s | \omega^P\right]}, \tag{6}$$

denote the "index" of an alternative currently in state ω^P, where τ denotes a measurable stopping time.[22] The definition in (6) is equivalent to the definition in Gittins and Jones (1974) and Gittins (1979). Given the state $\mathcal{S}$, denote the maximal index among the alternatives within the DM's CS by $\mathcal{I}^*(\mathcal{S}^P) = \max_{\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}^P(\hat{\omega}^P) > 0\}} \mathcal{I}(\omega^P)$. Note that $\mathcal{I}^*(\mathcal{S})$ depends on $\mathcal{S}$ only through the state of the alternatives in $\mathcal{S}$.

We now define an index for search that depends on the state of the alternatives in the CS only through the information that the latter contains for the evolution of

[21] This property guarantees the solution to the Bellman equation corresponding to the above dynamic program coincides with the true value function; it is immediately satisfied if payoffs and costs are uniformly bounded.

[22] Specifically, τ is a stopping time with respect to the process whose filtration is obtained by focusing attention on the alternative with initial state ω^P in all periods.

the search technology. Given the representation introduced above, such information is represented by the number of alternatives of each type $\xi \in \Xi$ brought to the set by past searches. Such information is already encoded in the state ω^S. As a result, the index of search depends on the state of the entire decision problem only through ω^S. Analogously to the indices 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 expected arrival of alternatives as the result of search implies that the maximization in the definition is not just over the stopping time, but also over the rule governing the allocation of future attention among the new alternatives and further search. Denote by τ a measurable stopping time, and by π a measurable 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 in by search or further search. Importantly, π selects only among search and alternatives that were not already in the CS when search was launched.[23]

Formally, given the state of the search technology $\omega^S \in \Omega^S$, its index is defined by

$$\mathcal{I}^S(\omega^S) \equiv \sup_{\pi,\tau} \frac{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s U_s^\pi | \omega^S\right]}{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s | \omega^S\right]}, \tag{7}$$

where U_s^π denotes the stochastic net flow payoff obtained in period s, under the rule π.

Denote by χ^* the policy that selects at each period $t \geq 0$, given the overall state $\mathcal{S}_t$ of the decision problem, search if and only if $\mathcal{I}^S(\omega^S) \geq \mathcal{I}^*(\mathcal{S}^{\mathcal{P}})$, and otherwise an arbitrary alternative with index $\mathcal{I}^*(\mathcal{S}^P)$.[24] Ties between the alternatives may be broken arbitrarily. In order to maintain consistency throughout the analysis, we assume that, in the case of a tie $\mathcal{I}^S(\omega_t^S) = \mathcal{I}^*(\mathcal{S}_t^P)$, the tie is broken in favor of search.

To present the next result, 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^*, (i) the search technology reaches a state in which its index is no greater than v, and (ii) *all* alternatives – regardless of when they were introduced into the CS – have an index no greater than v. That is, $\kappa(v)$ is the minimal number of periods until all indices are weakly below v. In case this event never occurs, $\kappa(v) = \infty$. Note that between the

[23]That is, π does not select among alternatives present before the launch of search. To make things clear, suppose search is launched in period t and terminated in period $\tau > t$. Then at each period $t < s < \tau$, π selects between search and alternatives available in period s but which were not yet in the CS in period t.

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

current period and the first period at which all indices are weakly below v, if the DM searches, new alternatives are introduced, in which case the evolution of their indices must also be taken into account in the computation of $\kappa(v)$.

Let $\mathcal{V}^*(\mathcal{S}_0)$ denote the maximal expected (per-period) payoff the DM can attain, given the initial state $\mathcal{S}_0$. That is,

$$\mathcal{V}^*(\mathcal{S}_0) = (1-\delta)\sup_{\chi} \mathbb{E}^{\chi}\left[\sum_{t=0}^{\infty} \delta^t U_t | \mathcal{S}_0\right].$$

Theorem 1. *(i) The index policy χ^* is optimal in the sequential search problem with endogenous consideration sets.*

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

$$\int_0^{\infty} \left(1 - \mathbb{E}\left[\delta^{\kappa(v)} | \mathcal{S}_0\right]\right) dv. \tag{8}$$

(iii) The index of search, as defined in (7), is such that, for any $\omega^S \in \Omega^S$,

$$\mathcal{I}^S(\omega^S) = \frac{\mathbb{E}\left[\sum_{s=0}^{\tau^*-1} \delta^s U_s | \omega^S\right]}{\mathbb{E}\left[\sum_{s=0}^{\tau^*-1} \delta^s | \omega^S\right]}, \tag{9}$$

where τ^* is the first time $s \geq 1$ at which $\mathcal{I}^S$ and all the indexes of the alternatives brought in by search fall below the value $\mathcal{I}^S(\omega^S)$ of the search index when search was launched, and where the expectations are with respect to the process induced by the index rule χ^*.

Proof of Theorem 1. The proof is in three steps. Step 1 below first establishes the result in part (iii) and then uses the recursive representation of the index of search in (9) to show that, when the DM follows an index policy, her expected (per-period) payoff satisfies the representation in (8), thus establishing part (ii). Steps 2 and 3, in the Appendix, then use the representation in (8) to show that the DM's payoff under the proposed index rule satisfies the Bellman equation for the dynamic program under consideration, thus proving the optimality of the index policy.

Step 1. We start by proving part (iii). Clearly, by definition, for all ω^S,

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

To see that the opposite inequality must also hold, note that by definition of τ^* in (iii), between period $s = 0$ and period $\tau^* - 1$, any option chosen under the index policy (whether exploring an alternative or searching again) must yield an expected per-period average payoff greater than $\mathcal{I}^S(\omega^S)$. Therefore, by following the index policy until τ^*-1, the expected average payoff on the right-hand-side of (9) must be greater than $\mathcal{I}^S(\omega^S)$.

Next, consider part (ii). We first introduce some notation. Define $\mathcal{S}_t^1 \vee \mathcal{S}_t^2 \equiv (\mathcal{S}_t^1(\omega) + \mathcal{S}_t^2(\omega) : \omega \in \Omega)$ and $\mathcal{S}_t^1 \backslash \mathcal{S}_t^2 \equiv (\max\{\mathcal{S}_t^1(\omega) - \mathcal{S}_t^2(\omega), 0\} : \omega \in \Omega)$. Any feasible state of the decision problem must specify one, and only one, state of the search technology (i.e., one state $\hat{\omega}^S$ for which $\mathcal{S}_t(\hat{\omega}^S) = 1$ and such that $\mathcal{S}_t(\omega^S) = 0$ for all $\omega^S \neq \hat{\omega}^s$). However, it will be convenient to consider fictitious (infeasible) states where search is not possible, as well as fictitious states with multiple search possibilities. If the state of the decision problem is such that either (i) the CS is empty, or (ii) there is a single alternative in the CS and the latter cannot be expanded, we will denote such state by $e_t(\omega)$, where $\omega \in \Omega$ is the state of the search technology in case (i) and of the single physical alternative in case (ii). Observe that the independence of the processes governing the evolution of the alternatives (conditional on their types ξ) along with the independence of these processes from the evolution of the search technology (conditional on the information about the types of the alternatives in the CS encoded into ω^S) imply that, for any $v \in \mathbb{R}$ and states $\mathcal{S}^1$ and $\mathcal{S}^2$, $\kappa(v|\mathcal{S}^1 \vee \mathcal{S}^2) = \kappa(v|\mathcal{S}^1) + \kappa(v|\mathcal{S}^2)$. That is, the time it takes to bring all indexes below v when the state of the decision problem is $\mathcal{S}^1 \vee \mathcal{S}^2$ is the sum of the times it takes to accomplish the same thing when the state is $\mathcal{S}^1$ and $\mathcal{S}^2$, respectively.

We construct the following stochastic process based on the values of the indices, and the expansion of the CS through search, under the index policy. Starting with the initial state $\mathcal{S}_0 = (\mathcal{S}_0^P, \omega_0^S)$, let $v^0 = \max\{\mathcal{I}^*(\mathcal{S}_0^P), \mathcal{I}^S(\omega_0^S)\}$. Consider the first time t^0 in which, when the DM follows the policy χ^*, all indices are strictly below v^0, with $t^0 = \infty$ if this event never occurs. Note that t^0 differs from $\kappa(v^0)$, as t^0 is the first time at which all indices are *strictly* below v^0, whereas $\kappa(v^0) = 0$ is the first time at which all indices are *weakly* below v^0. Next let $v^1 = \max\{\mathcal{I}^*(\mathcal{S}_{t^0}^P), \mathcal{I}^S(\omega_{t^0}^S)\}$, where $\mathcal{S}_{t^0} = (\mathcal{S}_{t^0}^P, \omega_{t^0}^S)$ is the state of the system at time t^0. Note that, by construction, $t^0 = \kappa(v^1)$. Furthermore, if $t^0 < \infty$ then $v^0 > \mathcal{I}^S(\omega_0^S)$ implies $\omega_{t^0}^S = \omega_0^S$. We can proceed in this manner to obtain a stochastic, strictly decreasing, sequence of values $(v^i)_{i \geq 0}$, with corresponding stochastic times $(\kappa(v^i))_{i \geq 0}$. Next, for any $i = 0, 1, 2, ...$, let $\eta(v^i) = \sum_{s=\kappa(v^i)}^{\kappa(v^{i+1})-1} U_s$ 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 $(\eta(v^i))_{i \geq 0}$ define the corresponding sequence of discounted accumulated net payoffs, with $\eta(v^i) = 0$ if $\kappa(v^i) = \infty$.

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 $(\eta(v^i))_{i\geq 0}$,

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

Next, using the definition of the indices (6) and (7), observe that

$$v^i = \frac{(1-\delta)\mathbb{E}\left[\eta(v^i)|\mathcal{S}_{\kappa(v^i)}\right]}{\mathbb{E}\left[1-\delta^{\kappa(v^{i+1})-\kappa(v^i)}|\mathcal{S}_{\kappa(v^i)}\right]}. \tag{10}$$

To see why (10) holds, recall that, at period $\kappa(v^i)$, given the state of the system $\mathcal{S}_{\kappa(v^i)}$, the optimal stopping time in the definition of the index v^i is the first time at which the index of the physical alternative corresponding to v^i (if v^i corresponds to a physical alternative), or the index of search and of all alternatives introduced through future searches (in case v^i corresponds to the index of the search technology), drop below v^i.[25]

Rearranging, multiplying both sides of (10) by $\delta^{\kappa(v^i)}$, and using the fact that $\delta^{\kappa(v^i)}$ is known at $\kappa(v^i)$, we have that

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

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

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

If follows that

$$\mathcal{V}(\mathcal{S}_0) = \mathbb{E}\left[\sum_{i=0}^{\infty} v^i\left(\delta^{\kappa(v^i)} - \delta^{\kappa(v^{i+1})}\right)|\mathcal{S}_0\right]. \tag{11}$$

Next, note that $\delta^{\kappa(v^i)} = 0$ whenever $\kappa(v^i) = \infty$, and that, for any $i = 0, 1, ...,$

[25] Note that if, at period $\kappa(v^i)$, there are multiple options ("physical" alternatives and search) with index v^i, the average sum $\mathbb{E}\left[\eta(v^i)|\mathcal{S}_{\kappa(v^i)}\right]$ of the discounted net payoffs across all alternatives with index v^i until the indices of all such options drop below v^i (in case search is included, also those of the new arriving alternatives), per unit of average discounted time, $\mathbb{E}\left[1-\delta^{\kappa(v^{i+1})-\kappa(v^i)}|\mathcal{S}_{\kappa(v^i)}\right]/(1-\delta)$, is the same as the average sum of the discounted net payoffs of each individual option with index v^i normalized by the average discounted time until the index of that alternative falls below v^i. This follows from the independence of the processes. Hence, Condition (10) holds irrespectively of whether, at $\kappa(v^i)$, there is a single or multiple options with index v^i.

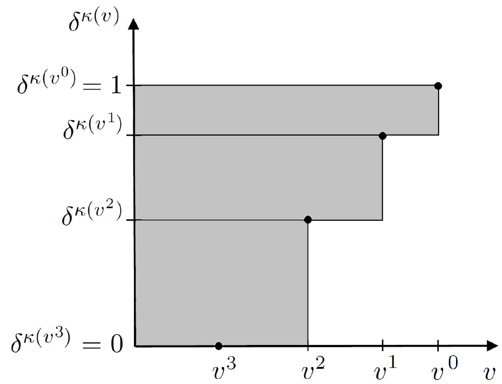

Figure 2: An illustration of the function $\delta^{\kappa(v)}$ and the region $\sum_{i=0}^{\infty} v^i \left(\delta^{\kappa(v^i)} - \delta^{\kappa(v^{i+1})}\right) = \int_0^\infty v \mathrm{d}\delta^{\kappa(v)}$, for a particular path with $\kappa(v^3) = \infty$.

$\kappa(v) = \kappa(v^{i+1})$ for all $v^{i+1} < v < v^i$. It follows that (11) is equivalent to

$$\mathcal{V}(\mathcal{S}_0) = \mathbb{E}\left[\int_0^\infty v \mathrm{d}\delta^{\kappa(v)} | \mathcal{S}_0\right] = \mathbb{E}\left[\int_0^\infty \left(1 - \delta^{\kappa(v)}\right) \mathrm{d}v | \mathcal{S}_0\right] = \int_0^\infty \left(1 - \mathbb{E}\left[\delta^{\kappa(v)} | \mathcal{S}_0\right]\right) \mathrm{d}v. \tag{12}$$

The construction of the integral function (12) is illustrated in Figure 2.

In the Appendix (Step 2) we then use the representation in (12) to characterize the difference in the DM's expected payoff between (a) following the index policy χ^* from the outset and (b) searching in the first period and then reverting to χ^* from the next period onward. To do so, we introduce an additional fictitious "retirement option," which is available at all periods and which yields a constant payoff $M < \infty$. The above characterization in turn permits us to establish that $\mathcal{V}$ solves the Bellman equation of the corresponding decision process (Step 3). Together with some standard arguments from dynamic programming, the above properties imply that χ^* is indeed an optimal policy, thus establishing part (i) in the Theorem. Some of the derivations in Steps 2 and 3 are a bit tedious and hence relegated to the Appendix. ■

3.3 Search and exploration dynamics

Equipped with the result in Theorem 1, we now highlight a few important properties of the dynamics of learning and expansion of the CS as a function of the search technology.

Corollary 2. *At any point in time, the decision to expand the CS depends on the* *composition of the CS only through the information that the latter contains about the likelihood that new searches will bring alternatives of different types ξ.*

The corollary is an immediate implication of the optimal policy being an index policy. However, note that the result is not trivial, for the opportunity cost of searching, which coincides with the value of continuing with the current CS, typically depends on the composition of the set over and above the information that the latter contains for expected search outcomes.

Corollary 3. *At any point in time, the relative likelihood of selecting any pair of alternatives in the CS is invariant to the state of the search technology.*

Corollary 3 is also an immediate implication of Theorem 1. More generally, the optimal policy being an index policy implies that the relative frequency with which the DM explores any two alternatives in the CS is invariant to what the decision maker expects to find by expanding the CS. This is true despite the fact that the expansion of the CS may bring alternatives that are more similar to some alternatives in the CS than others.

Corollary 4. *An improvement in the search technology yielding an increase in the probability that search brings an alternative of positive expected value (vis-a-vis the outside option) need not affect the decision to search even at histories at which, prior to the improvement, the DM is indifferent between searching and continuing with the current CS.*

The result 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 (9), the optimal stopping time in the search index coincides with the first time at which the index of search and the indexes of all alternatives brought in by search fall below the value of the search index at the time search was launched. As a result, any marginal 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 was launched does not affect the search index, and hence the decision to expand the CS. Note that the result hinges on the fact that the DM needs to explore the various alternatives that search brings to determine their values. When, instead, search is stochastic, but the value of the alternatives search brings is revealed to the DM upon arrival (as in the literature on undirected search), any marginal improvement in the search technology necessarily breaks the indifference in favor of search.

Definition 1. A search technology is *stationary* if $H_{\omega^S} = H$ for all $\omega^s \in \Omega^S$, *deteriorating* if $(-c_k, E_k)$ is decreasing in k in the sense of first-order stochastic dominance, and *improving* if $(-c_k, E_k)$ is increasing in k in the sense of first-order stochastic dominance.[26]

Corollary 5. *If the search technology is stationary, for any two states $\mathcal{S}$, $\mathcal{S}'$ at which the DM expands the CS, $\mathcal{V}^*(\mathcal{S}) = \mathcal{V}^*(\mathcal{S}')$.*

The corollary says that the continuation value when search is launched is invariant to the state of the CS. The result 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, as reported in the next corollary.

Corollary 6. *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.*

Since the state of an alternative changes only when the DM focuses on it, if in period t, $\mathcal{I}^S \geq \mathcal{I}^*(\mathcal{S}^P)$, under a stationary or improving search technology, the same inequality remains true in all subsequent periods. In this case, search corresponds to disposal of all alternatives within the current CS. Each time the DM launches herself into search, she starts fresh.

Corollary 7. *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.*

The result follows again from the characterization of the optimal stopping time in the recursive representation of the search index in (9). This time coincides with the first time at which the index of any physical alternative brought in by search, and the index of search itself, drop below the value of the search index at the time search was initiated. If the search technology is stationary, or deteriorating, the index of search falls (weakly) below its initial value immediately after search is launched. Hence, $\mathcal{I}^S(\omega^S)$ is independent of any information pertaining to future states of the search technology, conditional on ω^S.

[26]That is, the search technology is deteriorating if for any k and any upper set $Z \in \mathbb{R} \times \mathbb{R}^{|\Xi|}$, $\Pr((-c_{k+1}, E_{k+1}) \in Z) \leq \Pr((-c_k, E_k) \in Z)$. This definition is quite strong. In more specific environments in which there is an order on the set of types Ξ, weaker definitions are consistent with the results in the corollaries below.

Corollary 8. *Suppose the DM does not have the option to search (i.e., $\mathcal{S}_0$ is such that $\mathcal{S}_0(\omega^S) = 0$ for all $\omega^S \in \Omega^S$). Let $\hat{\mathcal{S}}_0$ denote the state that coincides with $\mathcal{S}_0$ except for the fact that $\hat{\mathcal{S}}_0(\hat{\omega}^S) = 1$ for some $\hat{\omega}^S \in \Omega^S$. The DM's willingness-to-pay to have access to a search technology in state $\hat{\omega}^S$ is equal to*

$$\mathcal{P}^*(\mathcal{S}_0; \hat{\omega}^S) = \int_0^\infty \left(\mathbb{E}\left[\delta^{\kappa(v)}|\mathcal{S}_0\right] - \mathbb{E}\left[\delta^{\kappa(v)}|\hat{\mathcal{S}}_0\right]\right) dv.$$

The result in Corollary 8 can be used to price access to a 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 exogenous outside option equal to $v \in \mathbb{R}_+$ to exit and take the outside option, both when search is available and when it is not. Then by integrating over the relevant values of the outside option one can compute $\mathcal{P}^*(\mathcal{S}_0; \hat{\omega}^S)$ and hence the maximal price that the DM is willing to pay to access the search technology.

3.4 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, thus bringing to an end the exploration process. Consider Weitzman's (1979) search problem, where uncertainty about the reward from each box is resolved immediately after opening the box, and where a box can be chosen only if it was previously opened. Under these assumptions, whether or not the choice of a box is reversible is irrelevant: since there is no additional information to be learned about the opened box, there is no reason for the DM to change her selection. Doval (2018) studies a generalization of Weitzman's problem where a box may be selected even if it has not been previously opened. She shows that the optimal solution takes the form of an index policy only under certain conditions, and studies the case in which the index policy need not be optimal.

In this section, we extend our analysis to a general model of learning, searching for new alternatives, and irreversible choice in which, at each period, the DM can (i) focus attention on one of the alternatives in the CS, (ii) expand the CS through search, or (iii) irreversibly commit to one of the alternative in the CS, possibly based on *partial* information about its value. We assume the DM must explore each alternative of type 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). The case $M_\xi = \infty$ corresponds to the model with no irreversible choice of the previous subsection. We allow for the possibility that the value of each alternative may never be fully revealed, with each exploration bringing new information, and for the possibility that the payoff from selecting an alternative in the CS without committing to it is positive in each period and evolves according to an arbitrary process.

Formally, we modify the general model of Section 3 as follows. In addition to the actions x_t and y_t defined above, we introduce an additional action, $z_{jt} \in \{0, 1\}$, representing the irreversible choice of an alternative j in the CS in period t, with $z_{jt} = 1$ if the DM irreversibly chooses alternative j, and $z_{jt} = 0$ otherwise. The period-t complete decision is then summarized by $d_t \equiv (x_t, y_t, z_t)$, with $z_t = (z_{jt})_{j=0}^{\infty}$. A sequence of decisions d is *feasible* if, for all $t \geq 0$, (i) $x_{jt} = 1$ or $z_{jt} = 1$ only if $j \in C_t$, (ii) $\sum_{j=1}^{\infty} x_{jt} + y_t + z_{jt} = 1$, (iii) $z_{jt} = 1$ if $z_{js} = 1$ for some $s < t$, and (iv) $z_{jt} = 1$ only if alternative j of type ξ has been explored at least M_ξ times. Together, the above conditions imply that, once an alternative is chosen (that is, once the DM has committed to it), there are no further decisions to be made.

Focusing on an alternative (i.e., exploring it without committing to it – formally captured by $x_{jt} = 1$) changes its state and may yield a flow payoff/cost, as in the baseline model. Irreversibly selecting alternative j of type ξ in period t (formally captured by $z_{jt} = 1$) yields a flow payoff to the DM from that moment onward, the value of which may be only imperfectly known to the DM at the time the irreversible decision is made. Denote by $R(\omega^P)$ the *expected flow value* from choosing an alternative when its current state is $\omega^P = (\xi, \theta)$. Note that since the choice is irreversible, $R(\omega^P)$ admits two equivalent interpretations: (i) If the the alternative is chosen, an immediate expected payoff equal to $R(\omega^P)/(1 - \delta)$ is obtained and there are no further payoffs; (ii) payoffs continue to accrue at all subsequent periods after the irreversible choice is made, with each expected payoff equal to $R(\omega^P)$.

Now suppose that each alternative's states can be partially ordered, based on the number of times the DM has focused on the alternative. Formally, suppose the set Θ takes the product form $\Theta = \Theta' \times \mathbb{N}$, with $m \in \mathbb{N}$ denoting the number of times the DM focused on the alternative and θ' all additional information. For any $\omega^P = (\xi, (\theta', m))$ and $\hat{\omega}^P = (\hat{\xi}, (\hat{\theta}', \hat{m}))$, we say that $\hat{\omega}^P$ "follows" ω^P if and only if $\hat{\xi} = \xi$ and $\hat{m} \geq m$. Denote this relation by $\hat{\omega}^P \succeq \omega^P$.

Condition 1. A type-ξ alternative has the *better-later-than-sooner property* if, for any $\omega^P = (\xi, (\theta', m))$, with $m \geq M_\xi$, and any $\hat{\omega}^P \succeq \omega^P$, either $R(\hat{\omega}^P) \geq R(\omega^P)$, or $R(\hat{\omega}^P), R(\omega^P) \leq 0$.

The following environments are examples of settings satisfying Condition 1.

Example 1 (**Weitzman's extended problem**). Consider the following extension of Weitzman's original problem: (i) The set of boxes is endogenous; (ii) each box of type ξ requires M_ξ explorations before the box's value is revealed; (iii) the DM can irreversibly commit (i.e., select) a box only if its value has been revealed, i.e., only after M_ξ explorations, where M_ξ can be stochastic (in this case, θ' specifies also whether or not a box can be selected and any information the DM may have about the value of M_ξ); (iv) the flow payoff from exploring a box without committing to it is equal to the cost of exploring the box (with the latter evolving stochastically with the number of past explorations) and is equal to zero for any exploration $t > M_\xi$; (v) the payoff $R(\omega^P)$ from irreversibly committing to a box whose value has been revealed (i.e., after the M_ξ'th exploration) remains constant after the M_ξ-the exploration and is equal to the box's prize. Clearly, this problem satisfies the better-later-than-sooner property.

Example 2 (**Purchase/Lease problem**). In each period, an apartment owner either chooses one of the real-estate agents she knows to lease her apartment, or searches for new agents. In addition, the owner can use one of the agents to sell the apartment. The decision to sell the apartment is irreversible. Once the apartment is sold, the owner's problem is over. The (expected) flow value u_{jt} the owner assigns to leasing the apartment through agent j of type ξ in state $\omega^P = (\xi, (\theta', m))$ is a function of the information θ' the owner has accumulated over time about agent j's ability to deal with all sorts of problems related to tenants. The (expected) value $R(\omega^P)$ the owner assigns to selling the apartment through the same agent may also depend on the agent's expertise with tenant-related problems but is primarily a function of the familiarity the agent has with the apartment, which is determined by the number of times m the agent has been hired by the owner in the past. If the agent has no or little past experience selling apartments (this information is contained in ξ), $R(\omega^P) \leq 0$. Else, for any θ' and $\hat{\theta}'$, $R(\xi, (\hat{\theta}', \hat{m})) \geq R(\xi, (\theta', m))$ if and only if $\hat{m} \geq m$. Clearly, this problem too satisfies the better-later-than-sooner property of Condition 1. Contrary to Weitzman's extended problem discussed in the previous example, the DM may derive a higher (expected) value from using an alternative without irreversibly committing to it (i.e, from leasing instead of selling through an agent) for an arbitrary long, possibly infinite, number of periods.

To accommodate for irreversible choice, we need to modify the definition of the index of each physical alternative in state $\omega^P \in \Omega^P$ as follows:

$$\mathcal{I}^P(\omega^P) \equiv \sup_{\pi,\tau} \frac{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s U_s^{\pi} | \omega^S\right]}{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s | \omega^s\right]}, \tag{13}$$

where τ is a stopping time, and where π is a rule specifying whether the DM focuses attention on the alternative, or irreversibly commits to it (i.e., chooses it). Similarly, modify the index of search $\mathcal{I}^S(\omega^S)$ in (15) by letting the rule π now specify not only whether the DM keeps searching or explores one of the alternatives introduced through search, but also whether she irreversibly commits to one of the alternatives that the new search brought to the CS.

Next, amend the definition of the index policy χ^* as follows. At each period $t \geq 0$, given the state $\mathcal{S}_t$ of the decision problem, the policy specifies to (a) search if $\mathcal{I}^S$ is greater than the index $\mathcal{I}^P$ of any physical alternative and the expected "retirement" value R of each alternative in the CS; (b) focus attention on a physical alternative in state ω^P if its index $\mathcal{I}^P$ is greater than its expected retirement value R, as well as the index of search, and both the index and the expected retirement value R of any other alternative in the CS; (c) choose (i.e., irreversibly commit to) an alternative in state ω^P if its retirement value $R(\omega^P)$ is greater than its index $\mathcal{I}^P(\omega^P)$, as well as the index of search and both the index and the expected retirement value of any other physical alternative in the CS.

We then have the following result (the proof is in the Appendix)

Theorem 2. *Suppose Condition 1 is satisfied for all $\xi \in \Xi$. The conclusions in Theorem 1 apply to the problem with irreversible choice under consideration. However, the stopping time τ^* in the characterization of the index of search in (9) is now the first time $s \geq 1$ at which $\mathcal{I}^S$, all the indexes of the alternatives brought in by search, and all retirement values of such alternatives fall below the value $\mathcal{I}^S(\omega^S)$ of the search index when search was launched.*

The result is established by considering a fictitious problem without irreversible choice in which, each time the DM focuses on a physical alternative in state ω^P, an "auxiliary alternative" with constant flow payoff equal to $R(\omega^P)$ is added to the CS and remains available in all subsequent periods, irrespectively of possible changes in the state of the physical alternative that generated it. The better-later-than-sooner property of Condition 1 guarantees that, if the DM ever selects one of these auxiliary alternatives, she necessarily picks the one corresponding to the latest activation of the physical alternative that generated it. This last property in turn implies that both (a) the non-perishability of the auxiliary alternatives and (b) the reversibility of choice in the fictitious problem play no role, which in turn implies that the optimal policy in the fictitious problem coincides with the one in the primitive problem.

3.5 Pandora's problem with endogenous boxes and gradual resolution of uncertainty

Equipped with the above results, we now revisit the extension of Weitzman's (1979) search problem introduced in Example 1. Recall that the problem is a generalization of the one in Section 2, in which (a) fully learning the value of each box (which, as in Weitzman's problem, is necessary to irreversibly commit to it) requires an arbitrary number of explorations M, with M stochastic and possibly equal to ∞, and (b) each exploration $s \leq M$ brings additional information both about the box's prize, u, and the number of explorations M necessary to learn the value u. For simplicity, and contrary to Weitzman, we assume here that there are no direct costs associated with the exploration of the various boxes (other than time), or with the expansion of the CS.[27]

More precisely, denote by $\omega^P = (\xi, \theta', n)$ the state of each box, represented by the box's primitive characteristics ξ (e.g., the prior distribution from which the box's prize, u, the number of explorations necessary to discover u, and the sequence of signals that lead to the discovery of u, are drawn), the history of signal realizations θ', and the number of times n the box has been explored. When a new box is added to the CS, its type ξ is drawn from a set Ξ from a known distribution $\mathcal{F}$. For simplicity, and consistently with the specification in Section 2, we assume that each expansion of the CS brings exactly one box.[28]

Let the reservation price of each box in state ω^P be equal to

$$\mathcal{I}^P(\omega^P) = \sup_{\tau>0} \frac{\mathbb{E}\left[\delta^{\phi}\left(1-\delta^{\tau-\phi}\right)\mathbf{1}_{\{\phi<\tau\}}u|\omega^P\right]}{1-\mathbb{E}\left[\delta^{\tau}|\omega^P\right]}, \tag{14}$$

where τ is a (stochastic) stopping-time and ϕ is the (stochastic) time at which the box's prize u is revealed. Similarly, let

$$\mathcal{I}^S(m) = \sup_{\tau>0} \frac{\delta \int_{\mathcal{A}(\mathcal{I}^S(m))}\left(\mathbb{E}\left[\delta^{\phi}\left(1-\delta^{\tau-\phi}\right)\mathbf{1}_{\{\phi<\tau\}}u|\xi\right]\right)\mathrm{d}\mathcal{F}(\xi)}{1-\int_{\mathcal{A}(\mathcal{I}^S(m))}\left(\mathbb{E}\left[\delta^{\tau}|\xi\right]\right)\mathrm{d}\mathcal{F}(\xi)} \tag{15}$$

[27] Introducing direct costs complicates the notation and the expressions for the reservation prices below, but is otherwise innocuous.

[28] This assumption is also not necessary, but simplifies the characterization of the reservation prices.

denote the reservation price of the option to expand the CS, where m denotes the number of times the CS has been expanded in the past, $\mathcal{A}(l) \equiv \{\xi \in \Xi : \mathcal{I}(\xi, \theta_0', 0) > l\}$ denotes the set of box types whose reservation price upon arrival exceeds l (the initial value θ_0' plays no role and can be taken arbitrarily), and τ, ϕ and u are as defined above.

Proposition 4. *In the sequential learning problem with an endogenous set of boxes and gradual resolution of uncertainty described above, the optimal policy is the same as in Proposition 1 but with the boxes' reservation prices given by (14) and the reservation price of the search option given by (15).*

We conclude by noting that the above generalization of Weitzman's model also accommodates for a generalization of the "eventual purchase" result of Section 2 to a setting in which the resolution of uncertainty is gradual. Consider the same environment described in Section 2, but now assume it takes M_i explorations (each with cost λ_i) to learn product i's value. Denote by $\mathcal{I}_i^k$ the reservation price of product i after it has been added to the CS and has been explored k times, and $\mathcal{I}_i^{\mathcal{S}}$ the search index after $i-1$ past searches, as defined in Section 2.[29]

Proposition 5. *For all $i \geq 1$, let $w_i = \min\{\mathcal{I}_i^{\mathcal{S}}, \mathcal{I}_i^1, \mathcal{I}_i^2, ..., \mathcal{I}_i^{M_i-1}, u_i\}$* be the "effective value" of product i. *The consumer purchases product i if, for all $j \neq i$, $w_j < w_i$ (and only if $w_i \geq w_j$ for all $j \neq i$).*

Similarly to its counterpart in Section 2, the result in Proposition 5 permits one to reduce the sequential learning problem with endogenous CS and gradual resolution of uncertainty to a static discrete-choice model. Contrary to the result in Section 2, however, the eventual purchase probabilities that one obtains from the "effective values" of Proposition 5 now depend not only on the endogenous sequence of expansions of the CS and the products' ultimate values u_i but also on the realizations of the signals the consumer receives about the values u_i after each exploration. Note that when the resolution of uncertainty about each alternative's value is gradual, the indices $\mathcal{I}_i^k$ reflect the rate at which information is revealed about the alternative (i.e., the speed of learning). This introduces a new tradeoff. As in Section 2, expansion of the CS is a substitute for exploration of the alternatives already in it. As a result, changes in the environment (e.g., in the inspection costs and/or in the products' values) that increase (alternatively, reduce) the relative desirability of expansion may lead the consumer to favor alternatives for which learning is faster (alternatively, slower).

[29] Note that, in this problem, product i's type $\xi = (F_i, \lambda_i)$ coincides with the distribution F_i over the consumer's value u_i for product i and the signals $\theta_i = (\theta_i^1, ..., \theta_i^{M_i})$ the consumer receives about u_i after each exploration, along with the cost λ_i of each exploration.

4 Extensions

In this section, we discuss how the results accommodate for a few simple extensions that may be relevant for applications.

Relative length of expansion. In order to allow for frictions in searching for new alternatives, we assume that, whenever the DM searches, she cannot focus on (that is, explore) any of the alternatives in the CS. In reality, the amount of time that each search occupies may differ from the amount of time that each exploration takes. For example, the online search for alternative providers of a given service may take seconds, but searching for a potentially suitable candidate for a given position make take longer than an interview. The results easily extend to a setting in which the number of periods that each search takes may differ from the number of periods that the exploration of each alternative occupies, with both times varying stochastically with the state. More generally, all of the results can be extended to a semi-Markov environment, where time is not slotted. Furthermore, because the length of time the exploration of each alternative takes can be arbitrary, by rescaling the payoffs and adjusting the discount factor appropriately, one can make the length of time for 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.

Multiple expansion possibilities. As illustrated in an example in the Online Appendix, if there are multiple options for search for which the outcome is correlated, an index policy cannot be guaranteed to be optimal.[30] Instead, the analysis readily extends to an environment in which there are multiple search possibilities with independent outcomes, by allowing for the possibility of multiple "search arms". For example, a researcher may choose in which field to search for a new project. A department with a single new faculty position may choose in which field to search for candidates. The analysis can also be extended to allow the results of search to include not just physical alternatives, but also new search possibilities.

No discounting. The proofs rely on the assumption that $\delta < 1$. As discussed above, an important special case of our analysis is an extension of Weitzman's problem in which the set of boxes is endogenous. Many applications of Weitzman's problem assume no discounting (i.e., $\delta = 1$). Our results extend to this case since, as noted in Olszewski and

[30]Such correlation arises naturally in an environment in which the DM can choose how much to invest in search, with different levels of investment corresponding to different "intensities" of search.

Weber (2015), Weitzman's problem with $\delta = 1$ is a special case of a multi-arm bandit problem with un-discounted "target processes", in which once an arm reaches a certain (target) state, payoffs no longer accrue. A well known result for such problems is that the finiteness they impose allows to take the limit as $\delta \to 1$ (see, e.g., Dumitriu, Tetali, and Winkler, 2003).

5 Concluding remarks

We introduce a model of sequential learning in which the decision maker alternates between exploring alternatives already in the consideration set and searching for new ones to explore in the future. The consideration set is thus constructed gradually over time in response to the information the decision maker collects. We characterize the optimal policy and study how the tradeoff between the exploration of existing alternatives and the expansion of the consideration set depend on the search technology. This tradeoff is conveniently summarized in a collection of indexes where the index for search is computed in recursive form accounting for future optimal decisions. The analysis also accommodates for certain irreversible decisions that admit as a special case an extension of Weitzman's (1979) "Pandora's boxes" problem in which the set of boxes is endogenous and the resolution of uncertainty about each box's value is gradual.

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

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.

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

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.

Edelman, B., Ostrovsky, M. and M. Schwarz (2007). "Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords," *American Economic Review*, 97(1), 242-259.

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

Epstein, R. and R. E. Robertson (2015). "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections," *Proceedings of the National Academy of Sciences*, 112(33), E4512-E4521.

Fershtman, D. and A. Pavan (2020). "Soft Affirmative Action and Minority Recruitment," mimeo, Northwestern and Tel-Aviv University.

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 indices," *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.

Gossner, O., Steiner, J. and C. Stewart (2019). "Attention, please!," mimeo.

Greminger, R. (2020). "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.

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.

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.

De los Santos, B., Hortaçsu, A. and M. R. Wildenbeest (2012). "Testing models of consumer search using data on web browsing and purchasing behavior," *American Economic Review*, 102(6), 2955-80.

Schneider, J. and C. Wolf (2019). "Switching forth and back: Problem solving under time pressure," mimeo.

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

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.

A Proofs

Proof of Proposition 1. The environment is a special case of the one studied in Section 3.4. Condition 1 is satisfied in this environment since $M_\xi = 1$ for all $\xi \in \Xi$ and all uncertainty is resolved upon the first inspection. As a result, Theorem 2 applies. It remains to verify that the policy described in the proposition coincides with the index policy of Theorem 2. This follows from observing that the reservation prices (1) are a special case of the indexes in (13) and, because the search technology is deteriorating, the reservation price of the option to expand the CS, $\mathcal{I}^S$, is a special case of the search index (9). ■

Proof of Proposition 2. Since product 0 corresponds to the outside option, one of the products is always purchased. Let $i \neq j$ be such that $w_i > w_j$. It suffices to show that product j will not be purchased.

Case 1: $j > i$. Suppose $w_j = \mathcal{I}_j^S$. Then $\min\{\mathcal{I}_i, u_i\} \geq w_i > \mathcal{I}_j^S$. This means that the consumer reads j's result only after clicking on i's result. Once she clicks on i's result, however, she will not read j's result, since $u_i > \mathcal{I}_j^S$. Next suppose $w_j = \mathcal{I}_j$. Then $\min\{\mathcal{I}_i, u_i\} \geq w_i > \mathcal{I}_j$. The consumer therefore clicks on j's result only after clicking on i's. But again, once she does so, she will not click on j's result, since $u_i > \mathcal{I}_j$. Finally, suppose $w_j = u_j$. Then even if she clicks on j's result, since $u_i \geq w_i > u_j$, the consumer will either recall a previous product or continue to search and find a better one. Hence product j will never be purchased, as claimed.

Case 2: $j < i$. Since $\mathcal{I}_i^S \geq w_i > w_j = \min\{\mathcal{I}_j, u_j, \mathcal{I}_j^S\}$, j cannot be purchased before i's result is read. Since the search indexes are declining, $\mathcal{I}_i^S \leq \mathcal{I}_j^S$. Because $w_i > w_j$, in turn this implies that $w_j = \min\{\mathcal{I}_j, u_j, \mathcal{I}_j^S\}$. Hence, $\min\{\mathcal{I}_i, u_i\} \geq \min\{\mathcal{I}_i, u_i, \mathcal{I}_i^S\} = w_i > w_j = \min\{\mathcal{I}_j, u_j\}$. Arguments analogous to those in Case 1 establish that j cannot be purchased. ■

Proof of Theorem 1. *Step 2.* We use the representation of the DM's payoff under the index rule in (8) to characterize how much the DM obtains from following the index policy χ^* from the outset rather than being forced to make a different decision in the first period and then reverting to χ^* from the next period onward. Such a characterization will permit us to establish in Step 3 the optimality of χ^* through dynamic programming.

Given the initial state $\mathcal{S}_0$, for any $\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0^P(\hat{\omega}^P) > 0\}$, denote by $\mathbb{E}\left[\tilde{u}|\omega^P\right]$ the immediate expected payoff from focusing on an alternative in state ω^P (the expectation is taken under the distribution H_{ω^P}) and by $\tilde{\omega}^P$ the new state of that alternative triggered by its exploration. Let

$$V^P(\omega^P|\mathcal{S}_0) \equiv (1-\delta)\mathbb{E}\left[\tilde{u}|\omega^P\right] + \delta\mathbb{E}\left[\mathcal{V}\left(\mathcal{S}_0 \backslash e(\omega^P) \vee e(\tilde{\omega}^P)\right)|\omega^P\right] \tag{16}$$

denote the DM's payoff from starting with exploring an alternative in state ω^P and then following the index policy χ^* from the next period onward. Similarly, let

$$V^S(\omega^S|\mathcal{S}_0) \equiv -(1-\delta)\mathbb{E}\left[\tilde{c}|\omega^S\right] + \delta\mathbb{E}\left[\mathcal{V}\left(S_0 \backslash e(\omega^S) \vee e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S)\right)|\omega^S\right] \tag{17}$$

denote the DM's payoff from expanding the CS, when the state of search is ω^s, and then following the index policy χ^* from the next period onward, where $\mathbb{E}\left[\tilde{c}|\omega^S\right]$ is the immediate expected cost from searching, $\tilde{\omega}^S$ is the new state of search after the first search is carried out, and $W^P(\tilde{\omega}^S)$ is the state of the new alternatives brought to the CS by search, with $\tilde{c}$ and $W^P(\tilde{\omega}^S)$ jointly drawn from the distribution H_{ω^S}.[31]

We introduce a fictitious "auxiliary option" which is available at all periods and yields a constant reward $M < \infty$ when chosen. Denote the state corresponding to this fictitious auxiliary option by ω^A_M, and enlarge Ω^P to include ω^A_M. Similarly, let $e(\omega^A_M)$ denote the state of the problem in which only the auxiliary option with fixed reward M is available. Since the payoff from the auxiliary option is constant at M, if $v \geq M$, then $\kappa(v|\mathcal{S}_0 \vee e(\omega^A_M)) = \kappa(v|\mathcal{S}_0)$. If, instead, $v < M$, then clearly $\kappa(v|\mathcal{S}_0 \vee e(\omega^A_M)) = \infty$. Hence, the representation in (8), adapted to the fictitious environment that includes the auxiliary option, implies that

$$\begin{aligned}\mathcal{V}(\mathcal{S}_0 \vee e(\omega^A_M)) = \textstyle\int_0^\infty \left(1 - \mathbb{E}\left[\delta^{\kappa(v)}|\mathcal{S}_0 \vee e(\omega^A_M)\right]\right) \mathrm{d}v = M + \textstyle\int_M^\infty \left(1 - \mathbb{E}\left[\delta^{\kappa(v)}|\mathcal{S}_0\right]\right) \mathrm{d}v \\ = \mathcal{V}(\mathcal{S}_0) + \textstyle\int_0^M \mathbb{E}\left[\delta^{\kappa(v)}|\mathcal{S}_0\right] \mathrm{d}v.\end{aligned} \tag{18}$$

The definition of χ^*, along with Conditions (16) and (17), then imply the following:

Lemma 1. *For any* (ω^S, ω^P, M),

$$\mathcal{V}(e(\omega^S) \vee e(\omega^A_M)) = \begin{cases} V^S(\omega^S|e(\omega^S) \vee e(\omega^A_M)) & \text{if } M \leq \mathcal{I}^S(\omega^S) \\ M > V^S(\omega^S|e(\omega^S) \vee e(\omega^A_M)) & \text{if } M > \mathcal{I}^S(\omega^S) \end{cases} \tag{19}$$

$$\mathcal{V}(e(\omega^P) \vee e(\omega^A_M)) = \begin{cases} V^P(\omega^P|e(\omega^P) \vee e(\omega^A_M)) & \text{if } M \leq \mathcal{I}^P(\omega^P) \\ M > V^P(\omega^P|e(\omega^P) \vee e(\omega^A_M)) & \text{if } M > \mathcal{I}^P(\omega^P). \end{cases} \tag{20}$$

[31]Note that $W^P(\tilde{\omega}^S)$ is a deterministic function of the new state $\tilde{\omega}^S = ((c_0, E_0), (c_1, E_1))$ of search. To see this, recall that, for any m, the function E_m counts how many alternatives have been added to the CS, for each possible state ω^P, as a result of the m'th search.

Proof of Lemma 1. First note that the index corresponding to the auxiliary option is equal to M. Hence, if $M \leq \mathcal{I}^S(\omega^S)$, χ^* prescribes to start with search. If, instead, $M > \mathcal{I}^S(\omega^S)$, χ^* prescribes to select the auxiliary option forever, yielding an expected (per period) payoff of M. To see why, in this case, $M > V^S(\omega^S|e(\omega^S) \vee e(\omega^A_M))$, observe that the payoff $V^S(\omega^S|e(\omega^S) \vee e(\omega^A_M))$ from starting with search and then following χ^* in each subsequent period is equal to $V^S(\omega^S|e(\omega^S) \vee e(\omega^A_M)) = \mathbb{E}\left[\sum_{s=0}^{\hat{\tau}-1} \delta^s U_s^{\hat{\pi}} + \frac{\delta^{\hat{\tau}}}{1-\delta} M|\omega^S\right]$, for some stopping and selection rules $\hat{\tau}, \hat{\pi}$. This follows from the fact that, once the DM, under χ^*, opts for the auxiliary option, he will continue to select that option in all subsequent periods. By definition of $\mathcal{I}^S(\omega^S)$,

$$M > \mathcal{I}^S(\omega^S) = \sup_{\pi,\tau} \frac{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s U_s^{\pi}|\omega^S\right]}{\mathbb{E}\left[\sum_{s=0}^{\tau-1} \delta^s|\omega^s\right]} \geq \frac{\mathbb{E}\left[\sum_{s=0}^{\hat{\tau}-1} \delta^s U_s^{\hat{\pi}}|\omega^S\right]}{\mathbb{E}\left[\sum_{s=0}^{\hat{\tau}-1} \delta^s|\omega^s\right]}.$$

Rearranging, $M\mathbb{E}\left[\sum_{s=0}^{\hat{\tau}-1} \delta^s|\omega^s\right] > \mathbb{E}\left[\sum_{s=0}^{\hat{\tau}-1} \delta^s U_s^{\hat{\pi}}|\omega^S\right]$. Therefore,

$$V^S(\omega^S|e(\omega^S) \vee e(\omega^A_M)) = \mathbb{E}\left[\sum_{s=0}^{\hat{\tau}-1} \delta^s U_s^{\hat{\pi}} + \frac{\delta^{\hat{\tau}} M}{1-\delta}|\omega^S\right] < M\mathbb{E}\left[\sum_{s=0}^{\hat{\tau}-1} \delta^s + \frac{\delta^{\hat{\tau}}}{1-\delta}|\omega^S\right] = M.$$

Similar arguments establish Condition (20). □

Next, for any initial state $\mathcal{S}_0$ of the decision problem, and any state $\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0(\hat{\omega}^P) > 0\}$ of the physical alternatives in the CS corresponding to $\mathcal{S}_0$, let $D^P(\omega^P|\mathcal{S}_0) \equiv \mathcal{V}(\mathcal{S}_0) - V^P(\omega^P|\mathcal{S}_0)$ denote the payoff differential between (a) starting by following the index rule χ^* right away and (b) focusing first on one of the physical alternatives in state ω^P and then following χ^* thereafter. Similarly, let $D^S(\omega^S|\mathcal{S}_0) \equiv \mathcal{V}(\mathcal{S}_0) - V^S(\omega^S|\mathcal{S}_0)$ denote the payoff differential between (c) starting with χ^* and (d) starting with search in state ω^S and then following χ^*. The next lemma relates these payoff differentials to the corresponding ones in a fictitious environment with the auxiliary option introduced above.[32]

Lemma 2. *Let $\mathcal{S}_0$ be the initial state of the decision problem, with $\omega^S \in \Omega^S$ denoting*

[32]In the statement of the lemma, $\mathcal{S}_0 \setminus e(\omega^S)$ is the state of a fictitious problem in which search is not possible, whereas $\mathcal{S}_0^P \setminus e(\omega^P)$ is the state of the CS obtained from $\mathcal{S}_0^P$ by subtracting an alternative in state ω^P.

the state of the search technology, as specified in $\mathcal{S}_0$. We have that[33]

$$D^S(\omega^S|\mathcal{S}_0) = \int_0^{\mathcal{I}^*(\mathcal{S}_0^P)} D^S(\omega^S|e(\omega^S) \vee e(\omega_v^A))\, d\mathbb{E}\left[\delta^{\kappa(v)}|\mathcal{S}_0 \setminus e(\omega^S)\right]. \tag{21}$$

Similarly, for any physical alternative in the CS of type $\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0^P(\hat{\omega}^P) > 0\}$,

$$D^P(\omega^P|\mathcal{S}_0) = \int_0^{max\{\mathcal{I}^*(\mathcal{S}_0^P \setminus e(\omega^P)), \mathcal{I}^S(\omega^S)\}} D^P(\omega^P|e(\omega^P) \vee e(\omega_v^A))\, d\mathbb{E}\left[\delta^{\kappa(v)}|\mathcal{S}_0 \setminus e(\omega^P)\right]. \tag{22}$$

Proof of Lemma 2. Using Condition (18), we have that, given the state $\mathcal{S}_0 \vee e(\omega_M^A)$ of the decision problem, and $\omega^S \in \Omega^S$,

$$\begin{aligned} D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_M^A)) = \mathcal{V}(\mathcal{S}_0) + \textstyle\int_0^M \mathbb{E}\left[\delta^{\kappa(v)}|\mathcal{S}_0\right] \mathrm{d}v + \mathbb{E}\left[\tilde{c}|\omega^S\right] \\ -\delta\mathbb{E}\left[\mathcal{V}(S_0 \backslash e(\omega^S) \vee e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S)) + \textstyle\int_0^M \mathbb{E}\left[\delta^{\kappa(v)}|S_0 \backslash e(\omega^S) \vee e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S))\right] \mathrm{d}v|\omega^S\right], \end{aligned} \tag{23}$$

where the equality follows from combining (17) with (18). Similarly,

$$\begin{aligned} D^S(\omega^S|e(\omega^S) \vee e(\omega_M^A)) = \mathcal{V}(e(\omega^S)) + \textstyle\int_0^M \mathbb{E}\left[\delta^{\kappa(v)}|e(\omega^S)\right] \mathrm{d}v + \mathbb{E}\left[\tilde{c}|\omega^S\right] \\ -\delta\mathbb{E}\left[\mathcal{V}(e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S))) + \textstyle\int_0^M \mathbb{E}\left[\delta^{\kappa(v)}|S_0 \backslash e(\omega^S) \vee e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S))\right] \mathrm{d}v|\omega^S\right]. \end{aligned} \tag{24}$$

Differentiating (23) and (24) with respect to M, using the independence across alternatives and search, and the property that $\kappa(v|\mathcal{S}^1 \vee \mathcal{S}^2) = \kappa(v|\mathcal{S}^1) + \kappa(v|\mathcal{S}^2)$, we have that

$$\frac{\partial D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_M^A))}{\partial M} = \mathbb{E}\left[\delta^{\kappa(M)}|S_0 \backslash e(\omega^S)\right] \frac{\partial D^S(\omega^S|e(\omega^S) \vee e(\omega_M^A))}{\partial M}. \tag{25}$$

That is, the improvement in $D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_M^A))$ that originates from a slight increase in the value of the auxiliary option M is the same as in a setting with only search and the auxiliary option, $D^S(\omega^S|e(\omega^S) \vee e(\omega_M^A))$, discounted by the expected time it takes (under the index rule χ^*) until there are no indices with value strictly higher than M, in an environment without search where the CS is the same as the one specified in $\mathcal{S}_0$.

[33]Recall that $\mathcal{I}^*(\mathcal{S}_0^P)$ is the largest index of the physical alternatives present in the CS under the state $\mathcal{S}_0$.

Similar arguments imply that, for any $\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0(\hat{\omega}^P) > 0\}$,

$$\frac{\partial D^P(\omega^P|\mathcal{S}_0 \vee e(\omega_M^A))}{\partial M} = \mathbb{E}\left[\delta^{\kappa(M)}|S_0\backslash e(\omega^S)\right] \frac{\partial D^P(\omega^P|e(\omega^P) \vee e(\omega_M^A))}{\partial M}. \tag{26}$$

Let $M^* \equiv \max\{\mathcal{I}^*(\mathcal{S}_0^P), \mathcal{I}^S(\omega^S)\}$. Integrating (25) over the interval $(0, M^*)$ of possible values for the auxiliary option and rearranging, we have that

$$\begin{aligned} D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_0^A)) &= D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_{M^*}^A)) - \int_0^{M^*} \mathbb{E}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^S))\right] \tfrac{\partial D^S(\omega^S|e(\omega^S) \vee e(\omega_v^A))}{\partial v} \mathrm{d}v \\ &= D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_{M^*}^A)) - \left[\mathbb{E}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^S))\right] D^S(\omega^S|e(\omega^S) \vee e(\omega_v^A))\right]_0^{M^*} \\ &\quad + \int_0^{M^*} D^S(\omega^S|e(\omega^S) \vee e(\omega_v^A))\mathrm{d}\mathbb{E}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^S))\right] \\ &= D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_{M^*}^A)) - D^S(\omega^S|e(\omega^S) \vee e(\omega_{M^*}^A)) \\ &\quad + \int_0^{M^*} D^S(\omega^S|e(\omega^S) \vee e(\omega_v^A))\mathbb{E}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^S))\right], \end{aligned}$$

where the second equality follows from integration by parts, whereas the third equality from the fact that $\mathbb{E}\left[\delta^{\kappa(M^*)}|S_0\backslash e(\omega^S)\right] = 1$ together with $\mathbb{E}\left[\delta^{\kappa(0)}|S_0\backslash e(\omega^S)\right] = 0$, as the DM can always receive her outside option. That the outside has value normalized to zero also implies that $D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_0^A)) = D^S(\omega^S|\mathcal{S}_0)$. It is also easily verified that $D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_{M^*}^A)) = D^S(\omega^S|e(\omega^S) \vee e(\omega_{M^*}^A))$.[34] Therefore, we have that

$$D^S(\omega^S|\mathcal{S}_0) = \int_0^{M^*} D^S(\omega^S|e(\omega^S) \vee e(\omega_v^A))\mathrm{d}\mathbb{E}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^S)\right]. \tag{27}$$

Similar arguments imply that

$$D^P(\omega^P|\mathcal{S}_0) = \int_0^{M^*} D^P(\omega^P|e(\omega^P) \vee e(\omega_v^A))\mathrm{d}\mathbb{E}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^P)\right]. \tag{28}$$

To complete the proof of Lemma 2, consider first the case where, given $\mathcal{S}_0$, χ^* specifies focusing on a physical alternative (i.e., $M^* \neq \mathcal{I}^S(\omega^S)$). Then Condition (21) in the lemma follows from (27) by noting that $M^* = \mathcal{I}^*(\mathcal{S}_0^P)$. Observe that, for any state $\omega^P \in \Omega^P$, if $M^* > \max\{\mathcal{I}^*(\mathcal{S}_0^P\backslash e(\omega^P)), \mathcal{I}^S(\omega^S)\}$ then $M^* = \mathcal{I}^P(\omega^P)$, in which case the integrand $D^P(\omega^P|e(\omega^P) \vee e(\omega_v^A))$ in (28) is equal to zero over the region $[0, \mathcal{I}(\omega^P)]$, and hence also over the interval $[0, \max\{\mathcal{I}^*(\mathcal{S}_0^P\backslash e(\omega^P)), \mathcal{I}^S(\omega^S)\}]$. That

[34] This follows immediately from the observation that $\mathcal{V}(\mathcal{S}_0 \vee e(\omega_{M^*}^A)) = \mathcal{V}(e(\omega^S) \vee e(\omega_{M^*}^A)) = M^*$, and similarly $\mathbb{E}\left[\mathcal{V}\left(\mathcal{S}_0\backslash e(\omega^S) \vee e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S) \vee e(\omega_{M^*}^A)\right)|\omega^S\right] = \mathbb{E}\left[\mathcal{V}\left(e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S) \vee e(\omega_{M^*}^A)\right)|\omega^S\right]$. Intuitively, under the index policy, any alternative with index strictly below M^* never receives any attention given the presence of an auxiliary option with payoff M^*.

is, in this case, Condition (22) clearly holds. Next, pick any state $\omega^P \in \Omega^P$ such that $\mathcal{I}^P(\omega^P) < M^*$. Condition (22) follows directly from (28) by noting that, in this case, $M^* = \max\{\mathcal{I}^*(\mathcal{S}_0^P \backslash e(\omega^P)), \mathcal{I}^S(\omega^S)\}$.

Next, consider the case where, given $\mathcal{S}_0$, χ^* specifies search (i.e., $M^* = \mathcal{I}^S(\omega^S)$). Then, for any $\omega^P \in \Omega^P$, $\max\{\mathcal{I}^*(\mathcal{S}_0^P \backslash e(\omega^P)), \mathcal{I}^S(\omega^S)\} = M^*$, in which case Condition (22) in the lemma follows directly from (28). That Condition (21) also holds follows from the fact that, in this case, $D^S(\omega^S|\mathcal{S}_0) = 0$ and the integrand $D^S(\omega^S|e(\omega^S) \vee e(\omega_v^A))$ in (27) is equal to zero over the entire region $[0, \mathcal{I}^S(\omega^S)]$. □

Step 3. Using the characterization of the payoff differentials in Lemma 2, we now establish that the average per-period payoff under the index policy χ^* satisfies the Bellman equation for the dynamic optimization problem under consideration. Let $\mathcal{V}^*(\mathcal{S}_0) \equiv (1-\delta)\sup_{\chi \in \mathcal{X}} \mathbb{E}^\chi \left[\sum_{t=0}^\infty \delta^t U_t | \mathcal{S}_0\right]$ denote the value function for the dynamic optimization problem.

Lemma 3. *For any state of the decision problem $\mathcal{S}_0$, with ω^S denoting the state of the search technology as specified under $\mathcal{S}_0$,*

1. *$\mathcal{V}(\mathcal{S}_0) \geq V^S(\omega^S|\mathcal{S}_0)$, with the inequality holding as an equality if and only if $\mathcal{I}^S(\omega^S) \geq \mathcal{I}^*(\mathcal{S}_0^P)$;*
2. *for any $\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0(\hat{\omega}^P) > 0\}$, $\mathcal{V}(\mathcal{S}_0) \geq V^P(\omega^P|\mathcal{S}_0)$ with the inequality holding as an equality if and only if $\mathcal{I}^P(\omega^P) = \mathcal{I}^*(\mathcal{S}_0^P) \geq \mathcal{I}^S(\omega^S)$.*

Hence, for any $\mathcal{S}_0$, $\mathcal{V}(\mathcal{S}_0) = \mathcal{V}^*(\mathcal{S}_0)$, and χ^* is optimal.

Proof of Lemma 3. First, use (19) to note that the integrand in (21) is non-negative for all $0 \leq v \leq \mathcal{I}^*(\mathcal{S}_0^P)$, and that the entire integral in (21) is equal to zero if and only $\mathcal{I}^*(\mathcal{S}_0^P) \leq \mathcal{I}^S(\omega^S)$. This establishes Condition 1 in the lemma. Similarly, use (20) to observe that for any $\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0^P(\hat{\omega}^P) > 0\}$, the integrand in (22) is non-negative for any $0 \leq v \leq \max\{\mathcal{I}^*(\mathcal{S}_0^P \backslash e(\omega^P)), \mathcal{I}^S(\omega^S)\}$, and that the entire integral in (22) is equal to zero if and only if $\mathcal{I}^P(\omega^P) \geq \max\{\mathcal{I}^*(\mathcal{S}_0^P \backslash e(\omega^P)), \mathcal{I}^S(\omega^S)\}$, which is the case if and only if $\mathcal{I}^P(\omega^P) = \mathcal{I}^*(\mathcal{S}_0^P) \geq \mathcal{I}^S(\omega^S)$. This establishes Condition 2 of the lemma.

Next, note that, jointly, Conditions 1 and 2 in the lemma imply that

$$\mathcal{V}(\mathcal{S}_0) = \max\left\{V^S(\omega^S|\mathcal{S}_0), \max_{\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0^P(\hat{\omega}^P) > 0\}} V^P(\omega^P|\mathcal{S}_0)\right\}.$$

Hence $\mathcal{V}$ solves the Bellman equation. That $\delta^T \mathbb{E}^\chi \left[\sum_{s=T}^{\infty} \delta^s U_s | \mathcal{S}\right] \to 0$ as $T \to \infty$ guarantees $\mathcal{V}(\mathcal{S}_0) = \mathcal{V}^*(\mathcal{S}_0)$, and hence the optimality of χ^*. $\square$

This completes the proof of the theorem. $\blacksquare$

Proof of Theorem 2. To ease the notation, assume the initial CS is empty. It will be evident from the arguments below that the optimality of χ^* does not hinge on this assumption. Consider first an environment where $M_\xi = 0$ for all ξ. It will also become evident from the arguments below that the result easily extends to environments where $M_\xi > 0$, as well as to environments where M_ξ is stochastic, and gradually learned over time.

Consider the following *auxiliary environment*, where all choices are *reversible*. Suppose that, whenever an alternative of type ξ is brought into the CS, an additional *auxiliary* alternative is also introduced into the CS, yielding a fixed flow payoff of $R(\xi, (\theta'_0, 0))$.[35] Next, suppose that, whenever a non-auxiliary alternative in state ω^P receives attention for the m-th time, a new auxiliary alternative is immediately introduced into the CS, yielding a fixed payoff of $R(\xi, (\theta', m))$, where $\omega^P = (\xi, (\theta', m))$ is the new state of the physical alternative after receiving attention for the m-th time, as defined in the main text. We say that an auxiliary alternative *corresponds to an (non-auxiliary) alternative* j if it has been introduced as the result of alternative j either being brought into the CS (through search) or receiving attention at some prior period. In this auxiliary environment, define the index of search as in (7), with the rule π specifying whether to keep searching or exploring one of the physical alternatives introduced through search, including the auxiliary alternatives brought in by the explorations of the physical alternatives introduced through search. For each physical alternative in state ω^P, define its new index as in (13), with the rule π in the definition of the index specifying for each period prior to stopping whether to focus on the alternative itself or to one of the auxiliary alternatives introduced as the result of the alternative's *future* explorations (importantly, π excludes any auxiliary alternative introduced in periods prior to the one in which the index is computed). Finally, let the index of any auxiliary alternative coincide with the alternative's retirement value, as specified by the function R.

It is easy to see that the same steps as in the proof of Theorem 1 imply that, in this auxiliary environment, the index policy based on the above new indices is optimal.[36] It

[35] Recall that $R(\xi, (\theta'_0, 0))$ is the retirement value of a physical alternative of type ξ that has never been explored.

[36] The proof must be adjusted to accommodate for the auxiliary alternatives introduced as the result

is also easy to see that the DM's problem in the auxiliary environment is a relaxation of the problem in the primitive environment in which (a) all decisions are reversible, and (b) physical alternatives can be retired also in states experienced in the past which may have changed as the result of further explorations. Hereafter, we argue that the DM's payoff in the primitive environment under the proposed index policy is the same as under the corresponding index policy in the auxiliary environment. To see this, first observe that, in the auxiliary environment, once the DM focuses attention on an auxiliary alternative, she continues to do so in all subsequent periods, since the indexes $R(\omega^P)$ of the auxiliary alternatives do not change. This implies that the reversibility of choice in the auxiliary environment plays no role. Next, observe that Condition 1 implies that, in the auxiliary environment, if the DM selects an auxiliary alternative, she always picks the one corresponding to the "newest" state of the corresponding non-auxiliary alternative, for the latest has the highest expected value R among all the auxiliary alternatives corresponding to the same physical alternative. This implies that the non-perishability of the older versions of the auxiliary alternatives in the auxiliary environment also plays no role. The same condition also guarantees that the policy π in the definition of the index of the physical alternatives in the auxiliary problem coincides with the one in (13) where the selection π is restricted to be over the exploration of the alternative under consideration and the retirement of the latter in its most recent state.

Finally, note that the proof immediately extends to settings in which $M_\xi > 0$ by assuming that, in the auxiliary environment, an auxiliary alternative is introduced into the CS only when its corresponding physical alternative has been explored more than M_ξ times, with M_ξ possibly stochastic and learned over time (in this latter case, the state θ' of a physical alternative may also contain information about M_ξ). ■

Proof of Proposition 4. The result follows from Theorem 2, by noting that the perceived value of the alternative, u, given the current information about it, is precisely the expected flow value of choosing an alternative $R(\omega^P)$. The reservation prices (14) are a special case of (13) in this setting, and, since the search technology is deteriorating, (15) is a special case of (9). ■

Proof of Proposition 5. The proof follows from arguments analogous to those in the proof of Proposition 2, and is therefore omitted. ■

---

of the DM exploring the physical alternatives. Since all the steps are virtually the same, the proof is omitted.